

Mathematik I

J. Hellmich

Stuttgart
2024 / 25

Autor: Dr. Jürgen Hellmich

*Ich höre und vergesse,
ich sehe und erinnere mich,
ich tue es und verstehe.*
(Konfuzius)

Lineare Algebra, Analysis

© Dr. Jürgen Hellmich

Alle Rechte vorbehalten, auch die der fotomechanischen Wiedergabe und der Speicherung in elektronischen Medien. Die elektronische Speicherung und fotomechanische Wiedergabe ist den Hörern der Lehrveranstaltung *Lineare Algebra, Analysis* an der DHBW Stuttgart im Semester 2024/25 ausschließlich zur Begleitung der Vorlesung gestattet.

Stand: 30. 11. 2024

Textsatz: \LaTeX

Bilder: Asymptote, Dia, GeoGebra, POV-Ray, PSTricks

Inhaltsverzeichnis

1	Grundlagen	1
1.1	Aussagenlogik	1
1.2	Mengen	7
2	Relationen und Funktionen	19
2.1	Allgemeine Eigenschaften von Relationen	19
2.2	Klassifikation von Relationen	21
2.3	Äquivalenzrelationen	21
2.4	Verkettung von Relationen und die inverse Relation	24
2.5	Funktionen	25
3	Zahlentheorie	35
3.1	Teilbarkeitstheorie ganzer Zahlen	35
3.2	Rechnen Modulo p	40
3.3	RSA-Verschlüsselung	45
3.4	Chinesischer Restsatz	47
4	Kombinatorik	51
4.1	Die Urnenmodelle	51
5	Vektorräume	61
5.1	\mathbb{R}^2 und \mathbb{R}^3 als Vektorraum	61
5.2	Die komplexen Zahlen	76
5.3	\mathbb{C}^n als Vektorraum	89
5.4	Der allgemeine Vektorraumbegriff*	90
5.5	Vektorräume mit Norm und Skalarprodukt	92
6	Matrizen	99
6.1	Lineare Gleichungssysteme und das GAUSS-Verfahren	99
6.2	Die Matrix zum LGS	105
6.3	Das Schema zum GAUSS-Verfahren	107
6.4	Lineare Unabhängigkeit	114
6.5	Basis und Dimension	124
6.6	Matrizen als lineare Abbildungen	135
6.7	Direkte Zerlegung eines Vektorraums	138
6.8	Die Dimensionsformel	141
6.9	Die inverse Matrix	143

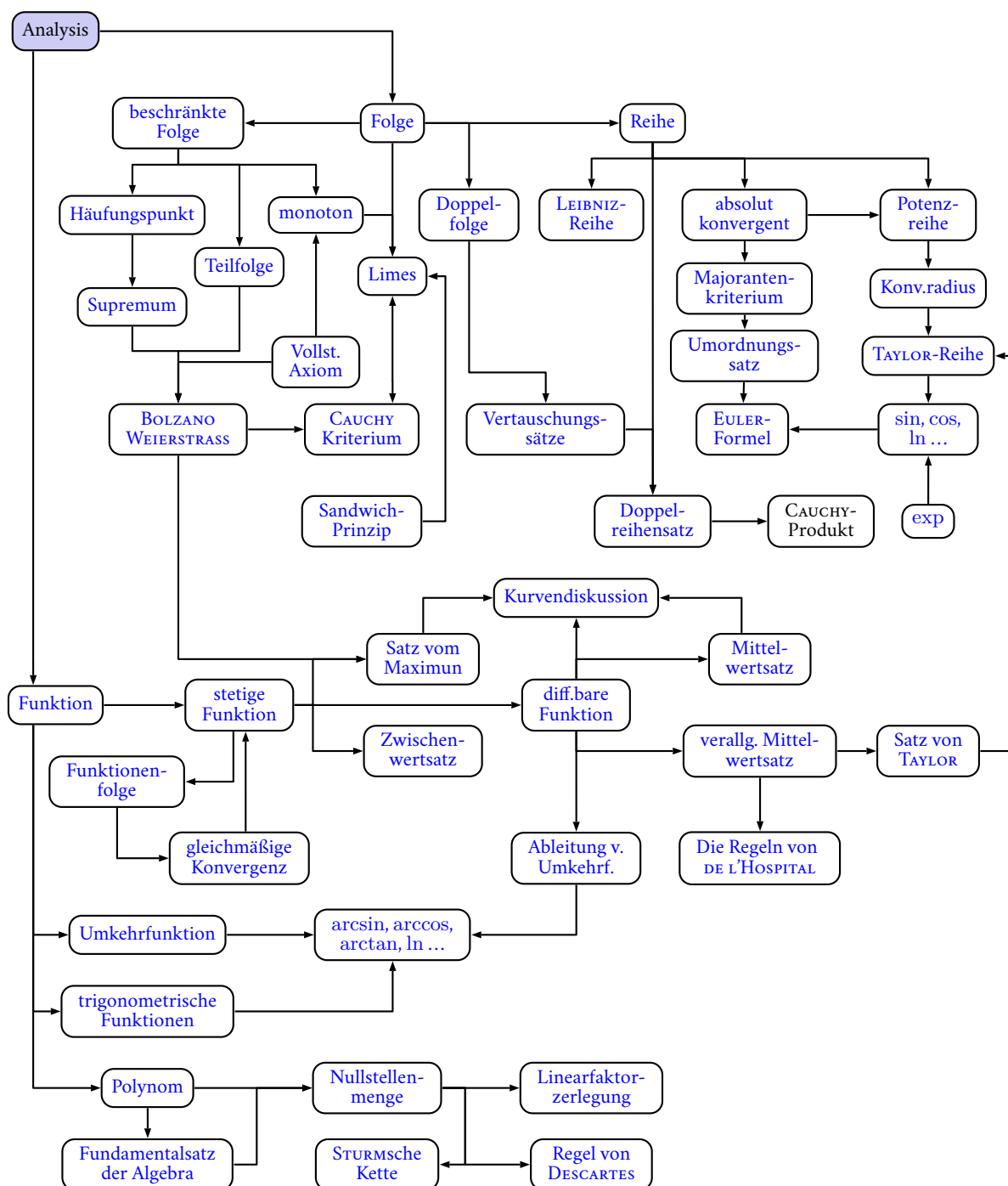
6.10	Die adjungierte Matrix	147
6.11	Koordinatentransformation	154
6.12	Determinanten	159
7	Eigenwerttheorie	177
7.1	Spektrum und Eigenvektoren	177
7.2	Selbstadjungierte lineare Abbildungen	179
7.3	Funktionalkalkül	184
7.4	Normale Abbildungen	185
8	Hauptachsentransformation	189
8.1	Kegelschnitte	189
8.2	Quadratische Formen	196
8.3	Beispiele für Quadriken	199
9	Rekurrenzgleichungen	205
9.1	Lineare Rekurrenzgleichungen 1. Ordnung	206
9.2	Lineare Rekurrenzgleichungen 2. Ordnung	208
10	Folgen und Reihen	213
10.1	Der Grenzwertbegriff	215
10.2	Reihen	242
11	Funktionen	267
11.1	Stetige Funktionen	267
11.2	Differentialrechnung	286
11.3	Ableitung von Potenzreihen	301
11.4	Ableitung von Umkehrfunktionen	308
11.5	Stammfunktionen	316
11.6	Die Taylor-Entwicklung	321
11.7	Das Newton-Verfahren	334
11.8	Polynome	340
11.9	Kurvendiskussion	366
12	Integralrechnung	381
12.1	Das Flächenproblem	381
12.2	Integrationstechniken	389
12.3	Anwendungen	395
12.4	Schaubilder der elementaren Funktionen	405
13	Lösungen	1-L
13.1	Aussagenlogik	1-L
13.2	Mengen	3-L
13.3	Relationen	5-L
13.4	Zahlentheorie	6-L

13.5	\mathbb{R}^2 und \mathbb{R}^3 als Vektorraum	8-L
13.6	Vektorräume	10-L
13.7	Matrizen	17-L
13.8	Eigenwerttheorie	33-L
13.9	Kegelschnitte	41-L
13.10	Quadratische Formen	44-L
13.11	Rekurrenzgleichungen	51-L
13.12	Folgen und Reihen	54-L
13.13	Funktionen	64-L
13.14	Kurvendiskussion	78-L
Nomenclature		1-Sy
Index		1-In

Lineare Algebra



Analysis



Zusammenfassung Ein *Vektorraum* (V, \mathbb{K}) über \mathbb{K} besteht aus einer Menge V , deren Elemente $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ als *Vektoren* bezeichnet werden und einem *Körper* \mathbb{K} , den sogenannten *Skalaren* t, s, r, \dots . Darüber hinaus gibt es eine *Addition* $V \times V \rightarrow V$, $[\mathbf{x}, \mathbf{y}] \mapsto \mathbf{x} + \mathbf{y} \in V$ für zwei Vektoren \mathbf{x} und \mathbf{y} , sowie eine *Skalarmultiplikation* $\mathbb{K} \times V \rightarrow V$, $[t, \mathbf{x}] \mapsto t \cdot \mathbf{x} \in V$ zwischen einem Skalar t und einem Vektor \mathbf{x} , die den gewohnten Rechenregeln folgen (5.4.1).

Beispiele:

- i) $(\mathbb{K}^n, \mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{F}_2$. Die Vektoren sind, wenn nichts anderes gesagt wird, Spaltenvektoren $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$ und $\mathbf{y} = [y_1, y_2, \dots, y_n]^t$, (p. 105)

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \pm \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 \pm y_1 \\ x_2 \pm y_2 \\ \vdots \\ x_n \pm y_n \end{bmatrix}, \quad t \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} tx_1 \\ tx_2 \\ \vdots \\ tx_n \end{bmatrix}.$$

- ii) $(\mathcal{P}, \mathbb{C})$, $(\mathcal{P}_n, \mathbb{C})$, der Vektorraum der komplexen Polynome \mathcal{P} , bzw. der Polynome mit einem Grad $\leq n$. Hier ist die Addition zweier Vektoren p und q und die Skalarmultiplikation von p mit einer Zahl $t \in \mathbb{C}$, die Addition von Funktionen: $(p + q)(x) := p(x) + q(x)$, bzw. die Multiplikation einer Funktion mit einer Zahl: $(tp)(x) := tp(x)$.
- iii) (M_n, \mathbb{K}) die Menge der $n \times n$ -Matrizen über dem Körper \mathbb{K} . Die Addition und Skalarmultiplikation ist komponentenweise definiert, wie bei Vektoren aus \mathbb{K}^n :

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \pm \begin{bmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} \pm b_{11} & \dots & a_{1n} \pm b_{1n} \\ a_{21} \pm b_{21} & \dots & a_{2n} \pm b_{2n} \\ \vdots & & \vdots \\ a_{n1} \pm b_{n1} & \dots & a_{nn} \pm b_{nn} \end{bmatrix},$$

$$t \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} ta_{11} & \dots & ta_{1n} \\ ta_{21} & \dots & ta_{2n} \\ \vdots & & \vdots \\ ta_{n1} & \dots & ta_{nn} \end{bmatrix}.$$

Eine *Metrik* auf dem Vektorraum (V, \mathbb{K}) ist eine Abbildung $d: V \times V \rightarrow \mathbb{R}$, mit den Eigenschaften

- i) $d(\mathbf{x}, \mathbf{y}) \geq 0$, $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$ (Definitheit)
- ii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (Symmetrie)
- iii) $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (Dreiecksungleichung)

Ein solcher Raum heißt *metrischer Raum*. Über die Metrik kann man in V Abstände messen.

Eine *Norm* auf dem Vektorraum (V, \mathbb{K}) ist eine Abbildung $\|\cdot\|: V \rightarrow \mathbb{R}_0^+$, mit den Eigenschaften (5.5.1)

- i) $\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ (Definitheit)

- ii) $\|t\mathbf{x}\| = |t|\|\mathbf{x}\|$
 iii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (Dreiecksungleichung)

Ein Vektorraum V mit einer Norm heißt *normierter Raum*, ist er endlichdimensional, auch *Banachraum*. Durch die Norm kann man auf V Längen messen. Durch jede Norm auf V wird über $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ eine kanonisch zugeordnete Metrik erzeugt.

Eine *Skalarprodukt* auf einem Vektorraum (V, \mathbb{C}) ist eine Abbildung $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{C}$, mit den Eigenschaften (5.5.2)

- i) $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$, $\langle \mathbf{x} | \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ (Definitheit)
 ii) $\langle \mathbf{x} | \mathbf{y} \rangle = \overline{\langle \mathbf{y} | \mathbf{x} \rangle}$ (Antisymmetrie)
 iii) $\langle s\mathbf{x} + t\mathbf{y} | \mathbf{z} \rangle = s\langle \mathbf{x} | \mathbf{z} \rangle + t\langle \mathbf{y} | \mathbf{z} \rangle$ (Sesquilinearität)

Ein solcher Raum heißt *Hilbertraum*, wenn V endlichdimensional ist. Ein Skalarprodukt auf (V, \mathbb{R}) hat, bis auf die Eigenschaft ii), die durch $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle$ zu ersetzen ist, dieselben Eigenschaften. Ein Skalarprodukt erlaubt es auf V über

$$\langle \mathbf{x} | \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos(\alpha)$$

Winkel zu messen. Dabei ist $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ die vom Skalarprodukt kanonisch erzeugte Norm, deren Dreiecksungleichung aus der CAUCHY-SCHWARZschen Ungleichung folgt (5.5.3):

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

Zwei Vektoren \mathbf{x} und \mathbf{y} heißen *orthogonal*, $\mathbf{x} \perp \mathbf{y}$, falls $\langle \mathbf{x} | \mathbf{y} \rangle = 0$ gilt.

Beispiele:

- i) Auf $(\mathbb{C}^n, \mathbb{C})$ bzw. $(\mathbb{R}^n, \mathbb{R})$ wird durch

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^n x_k \overline{y_k}, \quad \|\mathbf{x}\| = \sqrt{\sum_{k=1}^n |x_k|^2}, \quad \text{bzw.} \quad \langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^n x_k y_k, \quad \|\mathbf{x}\| = \sqrt{\sum_{k=1}^n x_k^2}$$

das übliche Skalarprodukt mit der kanonisch zugeordneten Norm definiert.

- ii) Auf $(\mathbb{C}^n, \mathbb{C})$ wird durch

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_{k=1}^n |x_k|^p}, \quad p \geq 1, \quad \|\mathbf{x}\|_\infty = \max\{|x_k| \mid k = 1, \dots, n\},$$

die Familie der L^p -Normen definiert, die außer für $p = 2$, nicht von einem Skalarprodukt stammen. (5.5.8, 5.5.14)

Ein *Teilraum* T eines Vektorraums (V, \mathbb{K}) ist eine Teilmenge von V , so daß (T, \mathbb{K}) ein Vektorraum ist (5.4.2):

$$\forall \mathbf{x}, \mathbf{y} \in T \quad \forall s, t \in \mathbb{K} \quad s\mathbf{x} + t\mathbf{y} \in T.$$

Beispiele:

- i) V und $\{0\}$ sind (die trivialen) Teilräume eines Vektorraums (V, \mathbb{K}) .
- ii) Für zwei Teilräume S, T von (V, \mathbb{K}) ist der Schnitt $S \cap T$ wieder ein Teilraum (während die Vereinigung $S \cup T$ normalerweise keiner ist).
- iii) Für eine Menge $\mathcal{B} \subseteq V$ ist die *lineare Hülle* von \mathcal{B} , (6.5.1)

$$\text{lh } \mathcal{B} = \{ t_1 \mathbf{b}_1 + t_2 \mathbf{b}_2 + \cdots + t_n \mathbf{b}_n \mid t_1, \dots, t_n \in \mathbb{K}, \mathbf{b}_1, \dots, \mathbf{b}_n \in \mathcal{B}, n \in \mathbb{N} \},$$

ein Teilraum von V . Für zwei Teilräume S, T von (V, \mathbb{K}) bezeichnet

$$S \vee T = \text{lh } S \cup T$$

den *Aufspann* von S und T (manchmal bezeichnet man diesen Teilraum auch durch $S + T$ (6.7.1)). Sind S und T zwei Ursprungsgeraden, mit den Richtungsvektoren \mathbf{u} und \mathbf{v} , dann ist $S \vee T$ die Ursprungsebene, die von \mathbf{u} und \mathbf{v} erzeugt wird.

- iv) Für eine lineare Abbildung $A : V \rightarrow W$ sind $\ker A$ und $\text{im } A$ Teilräume von V bzw. von W .

Eine Teilmenge \mathcal{B} eines Vektorraums V heißt *linear unabhängig*, wenn jede Nullkombination aus Vektoren von \mathcal{B} trivial ist. D.h., für jede Wahl von Vektoren $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ aus \mathcal{B} folgt aus $t_1 \mathbf{b}_1 + t_2 \mathbf{b}_2 + \cdots + t_n \mathbf{b}_n = \mathbf{0}$ immer $t_1 = \cdots = t_n = 0$. Andernfalls heißt \mathcal{B} *linear abhängig*. (6.4.1)

Eine Teilmenge \mathcal{B} eines Vektorraums V heißt *Basis* von V wenn sie linear unabhängig und erzeugend ist. (6.5.1)

Beispiele:

- i) Die kanonischen Basisvektoren für \mathbb{K}^n : (6.5.9)

$$\mathcal{E} = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ n \end{bmatrix} \right\} = \{ \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \}.$$

- ii) Die Menge $\mathcal{M} = \{ p_k \mid k \in \mathbb{N}_0 \}$ der Monome $p_k(x) = x^k$, bilden eine Basis für den Vektorraum \mathcal{P} der Polynome. (6.4.4)
- iii) In einem Vektorraum der Dimension n ist jede linear unabhängige Menge aus n Vektoren eine Basis für V .

Jeder Vektorraum (V, \mathbb{K}) hat eine Basis \mathcal{B} . Enthält diese nur endlich viele Elemente, so heißt V *endlichdimensional*, andernfalls *unendlichdimensional*. Die Anzahl der Elemente von \mathcal{B} heißt *Dimension* von V : $\dim V$. (6.5.6)

Jeder Vektor $\mathbf{x} \in V$ hat genau eine Entwicklung

$$\mathbf{x} = \sum_{k=1}^n x_i \mathbf{b}_i$$

bzgl. einer Basis \mathcal{B} . Der Spaltenvektor $\mathbf{x}_{\mathcal{B}} = [x_1, \dots, x_n]_{\mathcal{B}}^t$ der Koeffizienten von \mathbf{x} heißt *Basisdarstellung* von \mathbf{x} in der Basis \mathcal{B} . (6.5.2) Die Bestimmung der Koeffizienten x_i erfordert normalerweise das Lösen eines inhomogenen linearen Gleichungssystems, meist mittels GAUSS-Verfahren (6.1). Handelt es sich bei $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ um eine *Orthonormalbasis* (ONB), d. h., gilt $\langle \mathbf{b}_i | \mathbf{b}_j \rangle = \delta_{ij}$, so gewinnt man die Koeffizienten einfach durch $x_i = \langle \mathbf{x} | \mathbf{b}_i \rangle$ (6.5.15)

$$\mathbf{x} = \sum_{k=1}^n \langle \mathbf{x} | \mathbf{b}_k \rangle \mathbf{b}_k, \quad \mathbf{x}_{\mathcal{B}} = \mathbf{B}^* \mathbf{x} = [\langle \mathbf{x} | \mathbf{b}_1 \rangle, \langle \mathbf{x} | \mathbf{b}_2 \rangle, \dots, \langle \mathbf{x} | \mathbf{b}_n \rangle]_{\mathcal{B}}^t,$$

mit der Transformationsmatrix $\mathbf{B} := [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n]$. (p. 157)

Eine *lineare Abbildung* $A : V \rightarrow W$ von dem Vektorraum (V, \mathbb{K}) in den Vektorraum (W, \mathbb{K}) ist eine Funktion mit Definitionsbereich V , die mit der Vektorraumstruktur verträglich ist, d. h., mit folgender Eigenschaft: Für alle $\mathbf{x}, \mathbf{y} \in V$, $s, t \in \mathbb{K}$ gilt (6.6)

$$A(s\mathbf{x} + t\mathbf{y}) = sA(\mathbf{x}) + tA(\mathbf{y}).$$

Meist schreibt man $A\mathbf{x}$ statt $A(\mathbf{x})$. Ist $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ eine Basis von V , so ist A durch die Bilder der Basisvektoren $A\mathbf{b}_k$, $k = 1, \dots, n$, bereits vollständig bestimmt:

$$A(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \dots + x_n\mathbf{b}_n) = x_1A\mathbf{b}_1 + x_2A\mathbf{b}_2 + \dots + x_nA\mathbf{b}_n.$$

Wichtige Kenngrößen von A sind der *Kern* $\ker A$ und das *Bild* $\operatorname{im} A$ von A :

$$\ker A = \{\mathbf{x} \in V \mid A\mathbf{x} = \mathbf{0}\}, \quad \operatorname{im} A = \{\mathbf{y} \in W \mid \exists_{\mathbf{x} \in V} A\mathbf{x} = \mathbf{y}\}.$$

Für einen endlichdimensionalen Vektorraum V sind sie über die *Dimensionsformel* verknüpft (6.8.4):

$$\dim V = \dim \ker A + \dim \operatorname{im} A.$$

Für $V = \mathbb{C}^n$ und $W = \mathbb{C}^m$ wird eine lineare Abbildung $A : V \rightarrow W$ meist durch eine $m \times n$ -Matrix $A = [a_{ij}]_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$ beschrieben:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n].$$

Dabei sind die *Spaltenvektoren* $\mathbf{a}_j = [a_{1j}, a_{2j}, \dots, a_{mj}]^t$ die Bilder $A\mathbf{e}_j$ der kanonischen Basisvektoren \mathbf{e}_j ($j = 1, \dots, n$). Für einen Vektor $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$ ist

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{bmatrix} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n = \sum_{j=1}^n x_j \mathbf{a}_j.$$

Die Hintereinanderausführung $B \circ A$ zweier linearer Abbildungen $A : V \rightarrow W$ und $B : W \rightarrow X$, normalerweise einfach BA geschrieben, ist eine lineare Abbildung von V nach X .

Für die $k \times n$ -Matrix $A = [\mathbf{a}_1, \dots, \mathbf{a}_n] : \mathbb{K}^n \rightarrow \mathbb{K}^k$ und die $m \times k$ -Matrix $B = [\mathbf{b}_1, \dots, \mathbf{b}_k] : \mathbb{K}^k \rightarrow \mathbb{K}^m$ ist $BA : \mathbb{K}^n \rightarrow \mathbb{K}^m$ die $m \times n$ -Matrix (p. 137)

$$BA = [B\mathbf{a}_1, \dots, B\mathbf{a}_n].$$

Beispiele:

- i) Die identische Abbildung $\text{id} : \mathbb{K}^n \rightarrow \mathbb{K}^n, \mathbf{x} \mapsto \mathbf{x}$ ist durch die sogenannte *Einheitsmatrix* $\mathbb{1}$ gegeben und die *Nullabbildung* $0 : \mathbb{K}^n \rightarrow \mathbb{K}^n, \mathbf{x} \mapsto \mathbf{0}$ durch die *Nullmatrix* 0 :

$$\mathbb{1} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad 0 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

- ii) Die ebene Drehung D_α um den Winkel α ist

$$D_\alpha = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

- iii) Die Projektion $P_{\mathbf{n}}$ auf die Richtung des normierten Vektors $\mathbf{n} = [n_1, n_2, \dots, n_k]^t$ wird durch $P_{\mathbf{n}}\mathbf{x} = \langle \mathbf{x} | \mathbf{n} \rangle \mathbf{n}$ beschrieben und hat die Matrix

$$P_{\mathbf{n}} = \begin{bmatrix} n_1 n_1 & n_1 n_2 & \dots & n_1 n_k \\ n_2 n_1 & n_2 n_2 & \dots & n_2 n_k \\ \vdots & \vdots & \ddots & \vdots \\ n_k n_1 & n_k n_2 & \dots & n_k n_k \end{bmatrix}.$$

Die Projektion auf den zu \mathbf{n} orthogonalen Teilraum ist $\mathbb{1} - P_{\mathbf{n}}$.

- iv) Ein *lineares Gleichungssystem (LGS)*

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

wird durch die Koeffizientenmatrix $A = [a_{ij}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, durch den Lösungsvektor $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$ und durch die *Inhomogenität* $\mathbf{y} = [y_1, y_2, \dots, y_m]^t$ zu der handlichen

Vektorgleichung $A\mathbf{x} = \mathbf{y}$.

Das *homogene* LGS $A\mathbf{x} = \mathbf{0}$ hat $\ker A$ als Lösungsmenge, und das inhomogene

$$L = \{ \mathbf{x}_1 + \mathbf{x}_0 \mid \mathbf{x}_0 \in \ker A \}.$$

Dabei ist \mathbf{x}_1 eine Lösung der inhomogenen Gleichung: $A\mathbf{x}_1 = \mathbf{y}$. Daher ist auch $\mathbf{x}_1 + \mathbf{x}_0$ eine Lösung:

$$A(\mathbf{x}_1 + \mathbf{x}_0) = A\mathbf{x}_1 + A\mathbf{x}_0 = \mathbf{y} + \mathbf{0} = \mathbf{y}.$$

Für eine lineare Abbildung $A = [\mathbf{a}_1, \dots, \mathbf{a}_n] : V \rightarrow V$ auf dem n -dimensionalen Vektorraum (V, \mathbb{K}) sind die folgenden Aussagen äquivalent:

- i) A ist invertierbar.
- ii) A ist injektiv.
- iii) A ist surjektiv.
- iv) $\ker A = \{\mathbf{0}\}$.
- v) Die Menge $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ der Spaltenvektoren von A ist linear unabhängig.
- vi) $\det A := \det(\mathbf{a}_1, \dots, \mathbf{a}_n) \neq 0$.

Die *inverse Matrix* A^{-1} von A läßt sich gegebenenfalls durch das erweiterte GAUSS-Verfahren als Lösung B der Matrixgleichung $AB = \mathbb{1}$ berechnen (p. 144):

$$\begin{array}{|cccc|cccc|} \hline \mathbf{a}_{11} & \mathbf{a}_{12} & \cdots & \mathbf{a}_{1n} & 1 & 0 & \cdots & 0 \\ \mathbf{a}_{21} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_{n1} & \mathbf{a}_{n2} & \cdots & \mathbf{a}_{nn} & 0 & 0 & \cdots & 1 \\ \hline \end{array} \xrightarrow[\text{Verf.}]{\text{GAUSS}} \begin{array}{|cccc|cccc|} \hline 1 & 0 & \cdots & 0 & \mathbf{b}_{11} & \mathbf{b}_{12} & \cdots & \mathbf{b}_{1n} \\ 0 & 1 & \cdots & 0 & \mathbf{b}_{21} & \mathbf{b}_{22} & \cdots & \mathbf{b}_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \mathbf{b}_{n1} & \mathbf{b}_{n2} & \cdots & \mathbf{b}_{nn} \\ \hline \end{array}$$

Die *Determinante* \det auf dem Vektorraum $(\mathbb{K}^n, \mathbb{K})$ ist eine *normierte, alternierende Multilinearform*, d. h., eine Abbildung von $\mathbb{K}^n \times \mathbb{K}^n \times \dots \times \mathbb{K}^n \rightarrow \mathbb{K}$, so daß $[\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n] \mapsto \det(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)$ in jeder Komponente \mathbf{a}_i linear ist, die Vertauschung zweier Komponenten \mathbf{a}_i und \mathbf{a}_j für $i \neq j$ zu einem Vorzeichenwechsel führt und $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$ gilt. Durch $\det(A) := \det(\mathbf{a}_1, \dots, \mathbf{a}_n)$, für $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$, führen wir die Determinante für quadratische Matrizen ein (6.12.1). Durch die Koordinatendarstellung $\mathbf{a}_i = \sum_{k=1}^n a_{ki} \mathbf{e}_k$ läßt sich die Summendarstellung der Determinante gewinnen ((6.44))

$$\det(A) = \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \sum_{\pi \in \mathcal{S}_n} \text{sgn}(\pi) \cdot a_{\pi(1)1} a_{\pi(2)2} \cdots a_{\pi(n)n}.$$

Sie hat für konkrete Anwendungen weniger Bedeutung, aber sie stellt über das *Pfadbild* (p. 162) alternative Berechnungsmöglichkeiten bereit. Die alternierende Eigenschaft der Determinante sorgt dafür, daß sie bei elementaren Spaltenumformungen $\mathbf{a}_i \mapsto \mathbf{a}_i + t\mathbf{a}_j$ ihren Wert nicht ändert. Wegen $\det(A) = \det(A^t)$ (6.12.7), gilt das auch für Zeilenumformungen. Dadurch läßt sich eine Determinante werterhaltend in eine Determinante mit einer oberen (oder unteren)

Dreiecksmatrix umformen, so daß ihr Wert (gemäß dem Pfadbild) einfach als Produkt der Diagonalelemente abgelesen werden kann. Das ist das GAUSS-Verfahren zur Berechnung von Determinanten (6.12.8), das bei großen Matrizen angewendet wird. Ein anderes Verfahren liefert der LAPLACESche Entwicklungssatz, den man am besten an einem Beispiel versteht (6.12.10):

Wir entwickeln nach der dritten Zeile und dann nach der dritten Spalte:

$$\begin{aligned}
 \det \begin{bmatrix} 3 & 2 & 1 & 4 \\ 3 & 3 & 2 & 0 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 7 & 4 \end{bmatrix} &= +3 \cdot \det \begin{bmatrix} 3 & 2 & 1 & 4 \\ 3 & 3 & 2 & 0 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 7 & 4 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 3 & 2 & 1 & 4 \\ 3 & 3 & 2 & 0 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 7 & 4 \end{bmatrix} \\
 &= 3 \cdot \det \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 0 \\ 3 & 7 & 4 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 3 & 1 & 4 \\ 3 & 2 & 0 \\ 2 & 7 & 4 \end{bmatrix} \\
 &= 3 \cdot \left(4 \cdot \det \begin{bmatrix} 3 & 2 \\ 3 & 7 \end{bmatrix} + 4 \cdot \det \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \right) - 2 \cdot \left(4 \cdot \det \begin{bmatrix} 3 & 2 \\ 2 & 7 \end{bmatrix} + 4 \cdot \det \begin{bmatrix} 3 & 1 \\ 3 & 2 \end{bmatrix} \right) \\
 &= 3 \cdot (4 \cdot (21 - 6) + 4 \cdot (4 - 3)) - 2 \cdot (4 \cdot (21 - 4) + 4 \cdot (6 - 3)) = 32.
 \end{aligned}$$

Das Vorzeichen $(-1)^{i+j}$ der Unterdeterminante M_{ij} , (die aus $\det(A)$ durch Streichen der i -ten Zeile und der j -ten Spalte entsteht) findet man, in dem man in der Matrix A die linke obere Position in Gedanken mit einem $+$ markiert und dann die restlichen Felder des Quadrats schachbrettartig abwechselnd mit $-$ und $+$ versieht:

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

An der Position (i, j) dieser *Vorzeichenmatrix* steht das Vorzeichen von M_{ij} .

Aus der Eindeutigkeit alternierender Multilinearformen, bis auf eine multiplikative Konstante (6.12.2), erhält man sehr elegant den *Determinanten-Produktsatz* (6.12.4):

$$\det(AB) = \det(A) \det(B).$$

Die zu einer $m \times n$ -Matrix $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ adjungierte $n \times m$ -Matrix A^* ist durch

$$A^* := \overline{A}^t = \begin{bmatrix} \mathbf{a}_1^* \\ \mathbf{a}_2^* \\ \vdots \\ \mathbf{a}_n^* \end{bmatrix}$$

definiert (6.10.1). D. h., sie entsteht aus A durch Transponierung und Übergang zu konjugiert komplexen Einträgen $\overline{a_{ji}}$. Für reelle Matrizen gilt natürlich $A^* = A^t$.

Die zentrale Eigenschaft von A^* zeigt sich im Skalarprodukt (6.10.2):

$$\langle \mathbf{x} | A\mathbf{y} \rangle = \langle A^*\mathbf{x} | \mathbf{y} \rangle$$

gilt für alle $\mathbf{y} \in \mathbb{K}^n$ und alle $\mathbf{x} \in \mathbb{K}^m$, aus der man leicht $(AB)^* = B^*A^*$ gewinnen kann.

Wichtige Klassen quadratischer Matrizen sind die *selbstadjungierten Matrizen*, also Matrizen A , die mit A^* übereinstimmen (7.2.1)

$$A = A^*,$$

die *unitären* (6.10.4)

$$A^* = A^{-1}$$

und die *normalen Matrizen*, die durch die Bedingung

$$A^*A = AA^*$$

charakterisiert werden (7.4.1). Selbstadjungierte und unitäre Matrizen sind insbesondere normal.

Eigenvektoren $\mathbf{0} \neq \mathbf{x} \in V$ einer linearen Abbildung $A : V \rightarrow V$ sind Vektoren, die durch A lediglich gestreckt werden, die also der sogenannten *Eigenwertgleichung*

$$A\mathbf{x} = \lambda\mathbf{x}$$

genügen (7.1). Der Streckungsfaktor λ heißt *Eigenwert* von A zum Eigenvektor \mathbf{x} . Die Menge $\text{sp}(A)$ aller Eigenwerte heißt *Spektrum* von A . Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig (7.1.2).

Um für einen n -dimensionalen Vektorraum V zu einem Lösungsverfahren der Eigenwertgleichung zu kommen, in der ja sowohl \mathbf{x} , als auch λ unbekannt sind, wird die Bedingung $A\mathbf{x} = \lambda\mathbf{x}$ so lange äquivalent umgeformt, bis man eine Bedingung an die Eigenwerte allein gefunden hat (p. 177):

$$\exists \mathbf{0} \neq \mathbf{x} \in V \ A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow \exists \mathbf{0} \neq \mathbf{x} \in V \ (A - \lambda \mathbb{1})\mathbf{x} = \mathbf{0} \Leftrightarrow \ker A - \lambda \mathbb{1} \neq \{\mathbf{0}\} \Leftrightarrow \det(A - \lambda \mathbb{1}) = 0.$$

Bei der Funktion $\lambda \mapsto \det(A - \lambda \mathbb{1}) =: \chi(\lambda)$ handelt es sich um das sogenannte *charakteristische Polynom* von A (ein Polynom vom Grade n). Zumindest für komplexe Vektorräume hat es immer n Lösungen (von denen einige durchaus gleich sein können). Hat man die Eigenwerte λ einer Abbildung A gefunden, dann geschieht das Berechnen der zugehörigen Eigenvektoren \mathbf{x} normalerweise durch Lösen des linearen Gleichungssystems $(A - \lambda \mathbb{1})\mathbf{x} = \mathbf{0}$ mit Hilfe des GAUSS-Verfahrens.

Für jede normale Matrix, insbesondere also auch für jede unitäre und jede selbstadjungierte Matrix, gibt es immer eine ONB aus Eigenvektoren (7.2.4, 7.4.2). Das Spektrum einer normalen Matrix kann eine beliebige Menge aus maximal n verschiedenen komplexen Zahlen sein. Für unitäre Matrizen besteht das Spektrum aus komplexen Zahlen vom Betrag 1 (daher der Name *unitär*) (7.4.4), während das Spektrum einer selbstadjungierten Matrix immer aus reellen Zahlen besteht, selbst wenn A komplexe Einträge hat (7.2.2).

1 Grundlagen

1.1 Aussagenlogik

In der Aussagenlogik befassen wir uns mit Aussagen, für die es nur die beiden Wahrheitswerte *wahr* (w) und *falsch* (f) gibt. Aussagen können dabei durchaus umgangssprachlich ausgedrückt werden, müssen sich dabei aber auf Objekte beziehen, die sich einer objektiven Überprüfung unterziehen lassen. Die Aussage *Mir geht es gut* ist zur Zeit der Niederschrift möglicherweise wahr, aber das könnte sich von Tag zu Tag ändern. Sie läßt sich nicht verlässlich in den Kontext der Mathematik einbetten. Die Aussage *Im Jahre 2013 haben 70% aller Menschen braune Augen* ist ebenfalls möglicherweise wahr, aber zumindest können wir davon ausgehen, daß sie einen eindeutigen Wahrheitswert hat – auch wenn wir den voraussichtlich nicht erfahren werden, da wir schwerlich alle Menschen auf ihre Augenfarbe hin untersuchen können. Dagegen sind Aussagen wie

- $A := 2$ ist die einzige gerade Primzahl
- $B :=$ Alle ungeraden Zahlen haben die Form $2n + 1$
- $G :=$ Jede gerade Zahl größer als 2 ist die Summe zweier Primzahlen

präzise in der Sprache der Mathematik formuliert. Aussage A ist offensichtlich wahr, und durch kurzes Nachdenken erkennt man auch die Wahrheit von B. Aussage G dagegen, obwohl einfach zu verstehen, erschließt sich nicht sofort als wahr oder falsch. Testet man sie an einigen Zahlen, ($20 = 13 + 7$, $200 = 163 + 37 = 197 + 3$, $55698022500 = 22367000459 + 33331022041$, ...) dann wird man vermutlich nur Beispiele finden, die der Wahrheit von G nicht widersprechen. Um den Wahrheitswert von G als *falsch* zu identifizieren, wäre ja nur eine gerade Zahl nötig, die sich nicht durch zwei Primzahlen darstellen läßt. Eine solche Zahl ist noch nicht gefunden worden. Bei G handelt es sich um die sogenannte *GOLDBACHsche Vermutung* (1742), die bisher nicht entschieden ist.

Unsere Aufgabe besteht im Moment nicht darin, solche Aussagen zu beweisen, sondern darin Werkzeuge bereitzustellen, die es gestatten, Verknüpfungen zwischen Aussagen zu bilden und zu untersuchen. Wir gehen dabei von elementaren Aussagen A, B, ... aus, die, was die Regeln der Aussagenlogik angeht, als atomar, also ohne weitergehende innere Struktur angesehen werden. Dabei stellen wir jede Aussage durch ihre Wahrheitstabelle dar. Für eine Aussage A hat sie die simple Form

A
w
f

(1.1)

und gibt einfach die möglichen Wahrheitswerte von A wieder. Für Verknüpfungen von Aussagen erweitern wir sie durch die Wahrheitswerte der beteiligten Aussagen. Eine einstellige Verknüpfung ist der Übergang von A zur *Negation* $\neg A$ (wir verwenden auch die Schreibweise \bar{A} , die manchmal eine kompaktere Darstellung erlaubt). Dabei ist $\neg A$ genau dann wahr, wenn A falsch ist und genau dann falsch, wenn A wahr ist. In der Wahrheitstabelle läßt sich das übersichtlich darstellen:

A	$\neg A$
w	f
f	w

(1.2)

Für zwei Aussagen A und B stehen uns die elementaren Verknüpfungen *und*, *oder* und *folgt* zur Verfügung, die durch die sog. *Junktoren* \wedge , \vee und \Rightarrow dargestellt werden: $A \wedge B$, $A \vee B$, $A \Rightarrow B$. Die Wahrheitstabelle zeigt die Definitionen:

A	B	$A \wedge B$	$A \vee B$	$A \Rightarrow B$
w	w	w	w	w
w	f	f	w	f
f	w	f	w	w
f	f	f	f	w

(1.3)

Die *Konjunktion* A und B , oder kurz $A \wedge B$, ist genau dann wahr, wenn A und wenn B wahr ist. Die *Disjunktion* A oder B , oder kurz $A \vee B$, ist genau dann wahr, wenn A wahr ist, oder wenn B wahr ist, oder wenn beide wahr sind. Es handelt sich bei diesem *oder* also nicht um das umgangssprachliche, ausschließende oder, das in der Mathematik als *entweder oder* bezeichnet würde. Die *Implikation* A folgt B , A impliziert B , A ist Voraussetzung für B , oder kurz $A \Rightarrow B$, entspricht nicht genau dem alltäglichen Sprachgebrauch dieser Wendungen. Wenn aus A die Aussage B folgen soll, gehen wir normalerweise davon aus, daß zwischen A und B ein kausaler Zusammenhang besteht. Das behauptet $A \Rightarrow B$ nicht, auch wenn das in der Anwendung üblicherweise der Fall ist. Wenn wir die heute als falsch erkannte Aussage $A := \text{Der Mond besteht aus grünem Käse}$ und für B die oben verwendete über die ungeraden Zahlen wählen, hat die Implikation $A \Rightarrow B$ den Wahrheitswert *wahr*, auch wenn sicher kein erkennbarer Zusammenhang zwischen A und B besteht. Auf den ersten Blick mag auch befremdlich erscheinen, daß ein Wahrheitswert f für A und w für B den Wert w für $A \Rightarrow B$ nach sich zieht. Zusammen mit den Werten f für A , f für B und w für $A \Rightarrow B$ kann man das als die Formalisierung dafür ansehen, daß aus etwas Falschem beliebige Wahrheitswerte gefolgert werden können. Es ist der Grund dafür, daß ein Beweis mit etwas unbestreitbar Wahrem starten und bei der zu beweisenden Aussage enden muß — eine Tatsache, die einem die Schwierigkeit des Beweisens vor Augen führt, da es keine allgemeingültigen Regeln dafür gibt, bei welcher wahren Aussage man beginnen soll. Folgert man dagegen aus der Behauptung etwas Wahres, dann kann man nicht sicher sein, daß man dabei nicht doch von einer falschen Aussage ausgegangen ist und bei anderer Schlußweise auf etwas Falsches gestoßen wäre. Anders sieht es aus, wenn man in den Folgerungen auf eine falsche Aussage trifft, denn dann kann die Behauptung nicht wahr gewesen sein.

Eine weitere Bemerkung zu den verwendeten Begriffen *folgern*, *schließen*, *beweisen* ist angebracht. Wir haben sie in unseren bisherigen Argumenten benutzt, als wäre es klar, was genau darunter zu verstehen ist. Wollten wir das aber präzisieren, dann müßten wir den Vorgang des

Schließens formalisieren, um ihn auf eine solide Grundlage zu stellen. Wollten wir das hier in Angriff nehmen, so würden wir uns aber schnell in den Abgründen der Grundlagenmathematik verlieren (formale Sprachen, Prädikatenlogik ...). Wir beschäftigen uns sozusagen nur mit der globalen Logik des Schließens, ohne dabei auf dessen Mechanismus Bezug zu nehmen, der erst bei einer konkreten Belegung durch Aussagen ins Spiel kommt. Für $A := p$ ist eine Primzahl und $B :=$ Der größte gemeinsame Teiler von p mit jeder natürlichen Zahl ist 1 oder p kann das z. B. folgendermaßen aussehen: p sei eine Primzahl und $n \in \mathbb{N}$ eine beliebige natürliche Zahl. Da der größte gemeinsame Teiler t von p und n insbesondere ein Teiler von p ist und p nach der Definition einer Primzahl nur die Teiler 1 und p besitzen kann, muß t entweder 1 oder p sein. Dabei sind wir von der Wahrheit von A ausgegangen und haben auf die Wahrheit von B geschlossen. Das ist ein *direkter Beweis*, der durch folgende aussagenlogische Formel wiedergegeben werden kann:

$$A \wedge (A \Rightarrow B) \Rightarrow B. \quad (1.4)$$

Sehen wir uns ihre Wertetabelle an:

A	B	$A \Rightarrow B$	$A \wedge (A \Rightarrow B)$	$A \wedge (A \Rightarrow B) \Rightarrow B$
w	w	w	w	w
w	f	f	f	w
f	w	w	f	w
f	f	w	f	w

(1.5)

Formel (1.4) hat eine besondere Eigenschaft, nämlich, daß sie bei jeder Belegung von A und B mit Wahrheitswerten immer den Wert w ergibt. Eine solche Aussage wird als *Tautologie* bezeichnet. Tautologien beschreiben logische Strukturen, von denen einige wichtige Beweistechniken widerspiegeln. Bevor wir weitere Tautologien kennenlernen, führen wir den Begriff *Äquivalenz von Aussagen* ein:

$$A \Leftrightarrow B := (A \Rightarrow B) \wedge (B \Rightarrow A). \quad (1.6)$$

Die Wahrheitstabelle zeigt, daß $A \Leftrightarrow B$ genau dann w ergibt, wenn A und B denselben Wahrheitswert aufweisen.

A	B	$A \Rightarrow B$	$B \Rightarrow A$	$A \Leftrightarrow B$
w	w	w	w	w
w	f	f	w	f
f	w	w	f	f
f	f	w	w	w

(1.7)

In obigem Beispiel für A und B gilt sogar $A \Leftrightarrow B$: $A \Rightarrow B$ haben wir schon gezeigt. Bleibt die Umkehrung $B \Rightarrow A$: Dafür sei t ein Teiler von p , d. h. insbesondere gilt $t \leq p$. Der größte gemeinsame Teiler von t und p ist also t selbst. Daher kann t nur 1 oder p sein. Das bedeutet, daß p eine Primzahl ist.

Die Tautologie $A \vee \bar{A}$, oder $A \vee \neg A$, bezeichnet man als *Satz vom ausgeschlossenen Dritten*:

A	\bar{A}	$A \vee \bar{A}$
w	f	w
f	w	w

(1.8)

Satz von der Kontraposition:

A	B	$A \Rightarrow B$	$\overline{B} \Rightarrow \overline{A}$	$(A \Rightarrow B) \Leftrightarrow (\overline{B} \Rightarrow \overline{A})$
w	w	w	w	w
w	f	f	f	w
f	w	w	w	w
f	f	w	w	w

(1.9)

Setzen wir das in die Formel (1.4) für den direkten Beweis ein, so erhalten wir die Formel für den Widerspruchsbeweis

$$A \wedge (\overline{B} \Rightarrow \overline{A}) \Rightarrow B. \quad (1.10)$$

Das ist folgendermaßen zu verstehen: Wir gehen von der Aussage A aus und zeigen die Implikation $\overline{B} \Rightarrow \overline{A}$. Sie kann nur dann w liefern, wenn \overline{B} und \overline{A} beide den Wert w haben, oder wenn \overline{B} den Wert f hat (vergl. Tabelle (1.3)). Der erste Fall kann nicht eintreten, denn er würde $A \wedge \overline{A}$ mit den Wert w ausstatten, was nie der Fall sein kann. Dann bleibt nur f für \overline{B} , also w für B . Oder, etwas weniger formal: Anstatt direkt die Aussage B aus A zu folgern ist es manchmal einfacher, das Gegenteil von B anzunehmen und daraus das Gegenteil von A zu folgern. Da wir von der wahren Aussage A ausgegangen sind, kann \overline{B} nicht gelten, denn A und \overline{A} können nicht beide wahr sein. Also muß B wahr sein. Mit unserem Beispiel für A und B könnte das folgendermaßen ablaufen. Wir wollen $A \Rightarrow B$ zeigen. Wir gehen von \overline{B} aus, d. h. wir nehmen an, es gäbe eine natürliche Zahl n , die mit p einen größten gemeinsamen Teiler t hat, der von 1 und p verschieden ist. Da t dann ein echter Teiler von p ist, kann p keine Primzahl sein, im Widerspruch zur vorausgesetzten Wahrheit von A . Also muß B wahr sein, denn \overline{B} führt zu einem Widerspruch.

1.1.1 A Zeigen Sie, daß die Formeln $(A \Rightarrow B) \Leftrightarrow \overline{A} \vee B$ und $A \wedge \overline{A} \Rightarrow B$ Tautologien sind. Wie läßt sich die zweite interpretieren?

1.1.2 A DE MORGANSche Regeln: Zeigen Sie

- i) $\overline{A \wedge B} \Leftrightarrow \overline{A} \vee \overline{B}$
- ii) $\overline{A \vee B} \Leftrightarrow \overline{A} \wedge \overline{B}$
- iii) $(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$
- iv) $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$

1.1.3 A Nach Vernehmung der Verdächtigen Mark, Robert und John erklärt Holmes seinem Freund Dr. Watson: Wenn Robert der Täter ist oder wenn John der Täter ist, so ist das Alibi von Mark echt. Sind aber Mark oder John unschuldig, so ist Robert der Täter. Wenn John schuldig ist, so war auch Mark an dem Verbrechen beteiligt. Dr. Watson kombiniert und lächelt. Er weiß jetzt, wer schuldig ist.

1.1.4 Aussageformen Für jedes x einer Menge sei eine Aussage $A(x)$ gegeben. Dann heißt A eine *Aussageform*.

Z. B. sei P die Aussageform, die durch $P(x) := x \text{ ist eine Primzahl}$ auf der Menge $X = \mathbb{N}$ der natürlichen Zahlen definiert ist. Dann ist $P(2)$ die wahre Aussage *2 ist eine Primzahl* und $P(4)$ die falsche *4 ist eine Primzahl*.

Aussageformen lassen sich quantifizieren. Dafür gibt es die *Quantoren*

\exists	Existenzquantor
\forall	Allquantor

Sie lassen sich auf Aussageformen anwenden:

$\exists_{x \in X} P(x)$	Es gibt ein x aus X , für das $P(x)$ gilt
$\forall_{x \in X} P(x)$	Für alle x aus X gilt $P(x)$

Auf diese Weise lassen sich aus Aussageformen neue Aussagen gewinnen.

1.1.5 Beispiel

i) $X = \mathbb{N}$, $D(n) := \text{Die Summe aller natürlichen Zahlen bis } n \text{ ist } \frac{1}{2} n(n+1)$.

Meist wird das kurz durch $1+2+\dots+n = \frac{1}{2} n(n+1)$ ausgedrückt, ohne die Aussageform D explizit anzugeben. Sie wird sozusagen mitgedacht. Damit läßt sich die Aussage

$$A := \forall_{n \in \mathbb{N}} D(n) = \forall_{n \in \mathbb{N}} 1 + 2 + \dots + n = \frac{1}{2} n(n+1)$$

bilden. Die Aussage A als wahr zu erkennen bedeutet dann, die Gleichung(en) $1 + 2 + \dots + n = \frac{1}{2} n(n+1)$ zu beweisen.

ii) $X = \mathbb{N}$, $R(n) := \sqrt{n} \text{ ist eine rationale Zahl}$

Wir können $B := \forall_{n \in \mathbb{N}} R(n)$, oder $C := \exists_{n \in \mathbb{N}} \overline{R(n)}$ bilden.

B behauptet, daß die Wurzel aus allen natürlichen Zahlen als Bruch geschrieben werden kann. C ist wahr, wenn es (wenigstens) eine natürliche Zahl gibt, deren Wurzel kein Bruch ist. C gilt offensichtlich genau dann, wenn B falsch ist. Das bedeutet $C \equiv \overline{B}$. Wir werden sehen, daß B falsch ist, denn schon $\sqrt{2}$ ist kein Bruch. An diesem Beispiel erkennt man aber, wie quantifizierte Aussagen zu verneinen sind:

$$\overline{\forall_{x \in X} P(x)} \Leftrightarrow \exists_{x \in X} \overline{P(x)}, \quad (1.11)$$

$$\overline{\exists_{x \in X} P(x)} \Leftrightarrow \forall_{x \in X} \overline{P(x)}, \quad (1.12)$$

nämlich einfach dadurch, daß die Quantoren vertauscht und die quantifizierten Aussagen $P(x)$ negiert werden.

iii) Quantoren lassen sich auch kombinieren:

$$Pr := \forall_{n \in \mathbb{N}} \exists_{\substack{m \in \mathbb{N} \\ m > n}} P(m)$$

liest sich folgendermaßen: Für jede natürliche Zahl n gibt es eine natürliche Zahl m , die größer als n ist und für die m eine Primzahl ist. Oder kurz gesagt: Es gibt unendlich viele Primzahlen.

$$\overline{\text{Pr}} = \exists_{n \in \mathbb{N}} \overline{\exists_{\substack{m \in \mathbb{N} \\ m > n}} P(m)} = \exists_{n \in \mathbb{N}} \forall_{\substack{m \in \mathbb{N} \\ m > n}} \overline{P(m)}$$

behauptet, daß es eine natürliche Zahl gibt, so daß alle größeren keine Primzahlen sind, d. h., daß es nur eine endlich Anzahl von Primzahlen gibt. Wir werden sehen, daß diese Aussage falsch ist.

1.2 Mengen

1.2.1 Definition (CANTOR 1895) *Unter einer Menge verstehen wir jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.*

Die einfachsten Mengen lassen sich durch Angabe ihrer Elemente wiedergeben:

$W_1 := \{1, 2, 3, 4, 5, 6\}$, die möglichen Ergebnisse beim Würfeln mit einem Würfel.

$F := \{\text{rot, grün, blau}\}$, die drei Grundfarben.

$W_2 := \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),$
 $(2, 2), (2, 3), (2, 4), (2, 5), (2, 6),$
 $(3, 3), (3, 4), (3, 5), (3, 6),$
 $(4, 4), (4, 5), (4, 6),$
 $(5, 5), (5, 6),$
 $(6, 6)\}$,

die möglichen Ergebnisse beim Werfen zweier gleichfarbiger Würfel.

Wenn eine Menge zu viele Elemente hat, als daß man sie aufzählen könnte, sie aber einfach genug gebaut ist, verwendet man die Aufzählung mitunter trotzdem. Man zählt die Elemente soweit auf, bis das Bildungsgesetz deutlich wird (natürlich nur, wenn man es kennt) und ersetzt den Rest durch Auslassungspunkte.

$G := \{2, 4, 6, 8, 10, \dots\}$ ist, die Menge aller geraden Zahlen.

$H := \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots\}$ ist die Menge aller Stammbrüche, etc.

Diese Methode stößt natürlich schnell an ihre Grenzen. Bevor wir gleich leistungsfähigere kennenlernen, legen wir die übliche Notation fest:

$m \in M$	Das Element m gehört zur Menge M , m ist ein Element von M
\in	gehört zu ..., ist ein Element von ...
\notin	gehört nicht zu ..., ist kein Element von ...

1.2.2 Definition *Eine Menge A ist Teilmenge einer Menge B , wenn jedes Element von A auch ein Element von B ist. Dafür schreiben wir $A \subseteq B$, oder $B \supseteq A$. Zwei Mengen A und B heißen gleich, wenn A und B dieselben Elemente enthalten. Wir schreiben $A \subset B$, oder $B \supset A$, wenn A eine Teilmenge von B ist und die beiden Mengen nicht gleich sind.*

1.2.3 Lemma *Für zwei Mengen A und B gilt genau dann $A = B$, wenn $A \subseteq B$ und $B \subseteq A$ gilt.*

Beweis. Ist $A = B$, so ist nichts zu zeigen, denn $A \subseteq B$ und $B \subseteq A$ ist auch erfüllt, wenn die beiden Mengen gleich sind.

Gehen wir also von $A \subseteq B$ und $B \subseteq A$ aus und nehmen $A \neq B$ an, d. h., wir machen einen Widerspruchsbeweis. Da die Situation symmetrisch in A und B ist, können wir, ohne die Allgemeinheit unserer Argumentation einzuschränken, davon ausgehen, daß es ein Element a in

A gibt, das nicht zu B gehört. Das ergibt aber sofort einen Widerspruch dazu, daß laut Voraussetzung $A \subseteq B$ gilt, d. h., daß jedes Element aus A , insbesondere auch a , ein Element von B ist. Deshalb ist unsere Annahme zu verwerfen, und es gilt $A = B$. \square

Bemerkung: Es wird sich in kommenden Beweisen immer wieder die Situation ergeben, daß eine Ausgangslage symmetrisch bzgl. zweier Größen ist, so daß die folgende Argumentation nicht unzulässig eingeschränkt wird, wenn einer von beiden eine bestimmte Eigenschaft zugewiesen wird, die die andere genauso gut haben könnte. Dafür hat sich die Redewendung *ohne Beschränkung der Allgemeinheit* und deren Abkürzung *o. B. d. A.* durchgesetzt, die wir künftig auch verwenden werden.

Viele Mengen A sind Teilmengen gegebener Mengen B . Oft zeichnen sich die Elemente von A durch eine bestimmte Eigenschaft P aus, die (normalerweise) nicht allen Elementen von B zukommt. P ist eine *Aussageform*, die sich auf die Elemente von B bezieht. Für jedes $a \in B$ ist $P(a)$ eine Aussage. Die Menge A enthält dann alle Elemente a aus B , für die $P(a)$ wahr ist. Das wird durch die Schreibweise

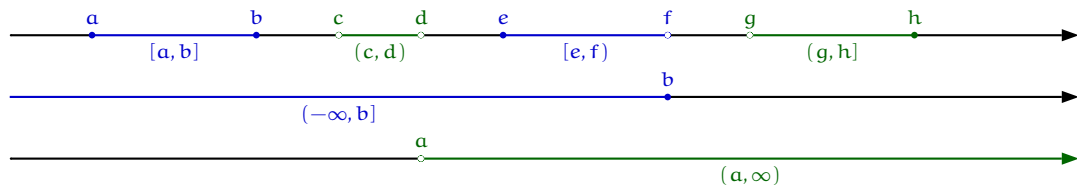
$$A = \{ a \in B \mid P(a) \} \quad (1.13)$$

wiedergegeben. Um Beispiele anführen zu können, nehmen wir vorerst die natürlichen Zahlen \mathbb{N} , die ganzen Zahlen \mathbb{Z} , die rationalen Zahlen \mathbb{Q} und die reellen Zahlen \mathbb{R} als gegeben an. Die geraden und die ungeraden Zahlen in \mathbb{N} lassen sich dann folgendermaßen darstellen:

$$G := \{ n \in \mathbb{N} \mid \exists_{k \in \mathbb{N}} n = 2k \}, \quad U := \{ n \in \mathbb{N} \mid \exists_{k \in \mathbb{N}} n = 2k - 1 \}.$$

Eine Teilmenge von \mathbb{R} der Art $\{ x \in \mathbb{R} \mid a \leq x \leq b \}$ heißt *Intervall*. Man vereinbart dafür die suggestive Schreibweise $[a, b]$. Es gibt die folgenden Intervalle:

$$\begin{aligned} [a, b] &:= \{ x \in \mathbb{R} \mid a \leq x \leq b \}, & (a, b) &:= \{ x \in \mathbb{R} \mid a < x < b \}, \\ [a, b) &:= \{ x \in \mathbb{R} \mid a \leq x < b \}, & (a, b] &:= \{ x \in \mathbb{R} \mid a < x \leq b \}, \\ (-\infty, b] &:= \{ x \in \mathbb{R} \mid x \leq b \}, & [a, \infty) &:= \{ x \in \mathbb{R} \mid a \leq x \}, \\ (-\infty, b) &:= \{ x \in \mathbb{R} \mid x < b \}, & (a, \infty) &:= \{ x \in \mathbb{R} \mid a < x \}. \end{aligned}$$



Natürlich gibt es auch $(-\infty, \infty)$, aber das ist \mathbb{R} . $[a, b]$ heißt *abgeschlossenes* und (a, b) *offenes Intervall*. $[a, b)$ und $(a, b]$ werden *halboffen* genannt. Diese Intervalle sind *beschränkt*, im Gegensatz zu $[a, \infty)$ und $(-\infty, b]$, die man *unbeschränkt* nennt. $[a, b]$ ist die Menge aller Punkte zwischen a und b , einschließlich der Randpunkte a und b . Dagegen enthält (a, b) die Randpunkte nicht und $[a, b)$ enthält zwar den unteren Randpunkt a , nicht aber den oberen b , usw. Insbesondere muß $[a, \infty)$ halboffen sein, denn ∞ ist kein Punkt aus \mathbb{R} .

Bemerkung: Für offene und halboffene Intervalle sind auch die Schreibweisen $]a, b[$, $]a, b]$, $[a, b[$ usw. gebräuchlich. Wir bevorzugen die Version mit runden Klammern, da das nach Ansicht des Autors die Lesbarkeit deutlich erhöht. Man vergleiche etwa $((-\infty, a] \cup (b e^{\mu((a,b])}, b]) \cap (3, 10^3)^c$ mit $(]-\infty, a] \cup]b e^{\mu([a,b]), b}) \cap]3, 10^3[^c$.

Die sog. *Dreieckszahlen* sind durch $D := \{n \in \mathbb{N} \mid \exists k \in \mathbb{N} \, n = 1 + 2 + \dots + k\}$ definiert, die MERSENNE-Zahlen durch $M := \{n \in \mathbb{N} \mid \exists k \in \mathbb{N} \, n = 2^k - 1\}$ und die FERMAT-Zahlen durch $F := \{n \in \mathbb{N} \mid \exists k \in \mathbb{N} \, n = 2^{2^k} + 1\}$. Das Beispiel der letzten beiden Mengen zeigt, daß eine etwas schlankere Notation wünschenswert ist. Die Menge M wird durch die Form $2^k - 1$ seiner Elemente eindeutig festgelegt. Daher sollte man diese Menge einfach durch Angabe seiner Konstruktionsvorschrift wiedergeben können. Dafür vereinbart man die Schreibweise

$$M = \{2^k - 1 \mid k \in \mathbb{N}\}.$$

Man gibt die Elemente einer Menge im ersten Teil durch eine Formel an und legt nach dem senkrechten Strich ihre Variablen fest. In dieser Allgemeinheit geht diese Vorschrift über die Beschreibung von Teilmengen hinaus. Man kann sie auch dafür verwenden, um neue Mengen aus schon gegebenen zu bilden. Der Formelteil muß dabei nicht mehr Elemente einer bereits gegebenen Menge erzeugen. Wir machen das an einem Beispiel klar. Stellen wir uns vor, wir wollen die rationalen Zahlen \mathbb{Q} definieren und können dabei auf die ganzen und die natürlichen Zahlen zurückgreifen. Das Symbol $\frac{p}{q}$ für einen Bruch können wir für die Definition nicht verwenden, denn es ist ja gerade unsere Aufgabe, diesem Symbol einen mathematischen Sinn zu verleihen. Trotzdem werden wir uns davon leiten lassen, was wir über Brüche wissen. Da wir uns im Moment nicht mit den Rechenregeln der rationalen Zahlen befassen, sondern nur mit der Menge an sich, ist das Einzige, was wir beachten müssen, daß unsere Definition keinen Bruch ergibt, der sich durch Kürzen in einen anderen verwandeln läßt. Gemäß der Mengendefinition darf \mathbb{Q} nämlich nur verschiedene Elemente enthalten (CANTOR spricht von *wohlunterschiedenen Objekten*). Wir müssen daher die Situation vermeiden, in der p und q einen gemeinsamen Teiler haben, der von 1 verschieden ist. Das können wir leicht erreichen, wenn wir den Begriff *größter gemeinsamer Teiler* $\text{ggT}(p, q)$ von p und q verwenden (siehe S. 36). Dann lautet unsere Bedingung einfach $\text{ggT}(p, q) = 1$. Als Definition der Menge der rationalen Zahlen erhalten wir

$$\mathbb{Q} := \{[p, q] \mid p \in \mathbb{Z} \wedge q \in \mathbb{N} \wedge \text{ggT}(p, q) = 1\}. \quad (1.14)$$

Mit dem neutralen $[p, q]$ haben wir uns bewußt von der üblichen Schreibweise $\frac{p}{q}$ abgegrenzt. Auf dieser Menge gibt es nämlich noch keinerlei Rechengesetze, die man sich bei dem Ausdruck $\frac{p}{q}$ meist dazudenkt. Diese Rechengesetze sind der zweite Schritt, nachdem die Menge als solche vorhanden ist. Sie prägen ihr eine mathematische Struktur auf, die sie erst zu dem macht, was wir unter den rationalen Zahlen verstehen.

Die Menge

$$\mathbb{P} := \{p \in \mathbb{N} \mid p \geq 2, \, n \mid p \Rightarrow n = 1 \vee n = p\} \quad (1.15)$$

ist leicht als die Menge aller Primzahlen zu identifizieren.

1.2.4 Definition Für zwei Mengen A und B wird die Menge

$$A \times B := \{[a, b] \mid a \in A \wedge b \in B\} \quad (1.16)$$

als Produktmenge von A und B bezeichnet. Allgemeiner ist für Mengen A_1, A_2, \dots, A_n die Produktmenge folgendermaßen definiert:

$$A_1 \times A_2 \times \dots \times A_n := \{ [a_1, a_2, \dots, a_n] \mid a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n \}. \quad (1.17)$$

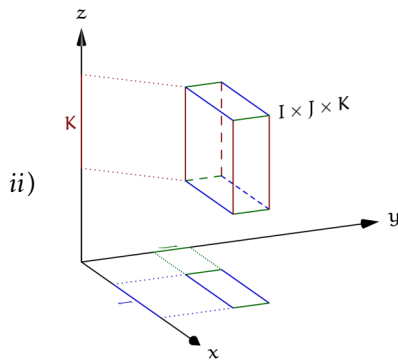
Bei der Definition (1.17) haben wir von einer gängigen Übereinkunft Gebrauch gemacht, gemäß der die Aussage $a_1 \in A_1 \wedge a_2 \in A_2 \wedge \dots \wedge a_n \in A_n$ durch eine einfache Aufzählung der einzelnen Bedingungen $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$ abgekürzt werden darf.

Bemerkung: Normalerweise besteht keine Gefahr, das Tupel $[a, b]$ mit dem Intervall $[a, b] \subset \mathbb{R}$ zu verwechseln, da aus dem Kontext heraus klar wird, was gemeint ist.

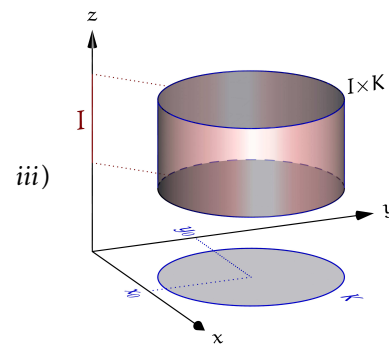
1.2.5 Beispiel

i) Für $A_1 := \mathbb{R}, \dots, A_n := \mathbb{R}$ ergibt sich $\mathbb{R}^n := \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{n \text{ mal}}$ als die Menge

$$\mathbb{R}^n = \{ [x_1, x_2, \dots, x_n] \mid x_1, x_2, \dots, x_n \in \mathbb{R} \}. \quad (1.18)$$



I, J und K sind Intervalle in \mathbb{R} .
Dann ist $I \times J \times K \subseteq \mathbb{R}^3$ ein Quader.



$I \subseteq \mathbb{R}$ ist ein Intervall, K die Kreisscheibe
 $K := \{ [x, y] \mid (x - x_0)^2 + (y - y_0)^2 \leq r^2 \} \subset \mathbb{R}^2$ mit Radius r und Zentrum $[x_0, y_0]$. Dann ist $I \times K$ ein Zylinder.

1.2.6 Definition A und B seien Teilmengen einer Menge X . Wir definieren die leere Menge $\emptyset := \{ \}$ als die Menge, die kein Element enthält. Dann lassen sich die folgenden Mengenoperationen uneingeschränkt einführen.

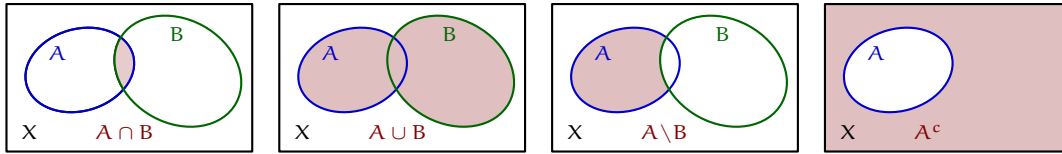
$$A \cap B := \{ x \in X \mid x \in A \wedge x \in B \}, \quad \text{Durchschnitt} \quad (1.19)$$

$$A \cup B := \{ x \in X \mid x \in A \vee x \in B \}, \quad \text{Vereinigung} \quad (1.20)$$

$$A^c := \{ x \in X \mid \neg x \in A \}, \quad \text{Komplement} \quad (1.21)$$

$$A \setminus B := \{ x \in X \mid x \in A \wedge x \notin B \}. \quad \text{Mengendifferenz} \quad (1.22)$$

A und B heißen disjunkt, falls $A \cap B = \emptyset$ gilt. Wir schreiben $x \notin A$ statt $\neg x \in A$.



Offensichtlich ist X^c die leere Menge. Die enge Verwandtschaft von \cap , \cup und c mit den logischen Junktoren \wedge , \vee und \neg spiegelt sich in den DE MORGANSchen Regeln für Mengen wieder:

1.2.7 Satz Für Teilmengen A , B und C einer Menge X gelten folgende Rechenregeln:

$$(A \cap B)^c = A^c \cup B^c \quad (1.23)$$

$$(A \cup B)^c = A^c \cap B^c \quad (1.24)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C) \quad (1.25)$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \quad (1.26)$$

Beweis. Wir verwenden die DE MORGANSchen Regeln der Logik (Übung 1.1.2) und die Tatsache, daß für zwei Mengen M und N Gleichheit gilt, falls $x \in M \Leftrightarrow x \in N$ erfüllt ist.

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

Das zeigt bereits die erste Gleichung. Wir verwenden $M^{cc} := (M^c)^c = M$, um aus ihr die zweite auszurechnen:

$$(A \cup B)^c = (A^{cc} \cup B^{cc})^c = ((A^c)^c \cup (B^c)^c)^c = ((A^c \cap B^c)^c)^c = A^c \cap B^c.$$

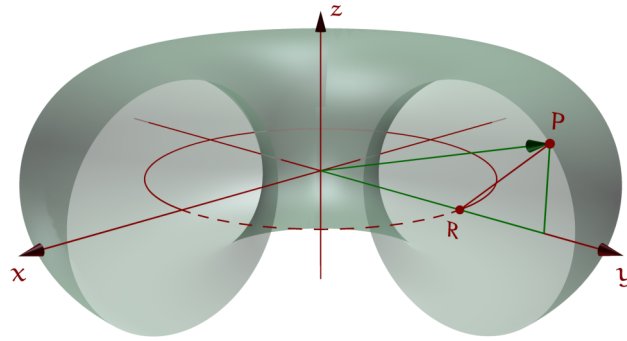
Zur dritten Gleichung:

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

Die letzte Gleichung kann man auf dieselbe Weise ableiten, oder mit Hilfe der bisherigen Ergebnisse berechnen:

$$\begin{aligned} (A \cup B) \cap C &= (A^{cc} \cup B^{cc}) \cap C^{cc} = (A^c \cap B^c)^c \cap C^{cc} = ((A^c \cap B^c) \cup C^c)^c \\ &= ((A^c \cup C^c) \cap (B^c \cup C^c))^c = (A^c \cup C^c)^c \cap (B^c \cup C^c)^c \\ &= (A \cap C) \cup (B \cap C). \end{aligned} \quad \square$$

1.2.8 A Finden Sie für den Torus eine Mengendarstellung. (Ein Torus ist ein Ring, der entsteht, wenn ein Kreis mit Radius r , der in der xz -Ebene liegt und dessen Mittelpunkt sich auf der x -Achse in der Entfernung $R > r$ vom Ursprung befindet, einmal um die z -Achse herumgedreht wird.)



Finden Sie für $R = 6$ und $r = 4$ einen Punkt auf dem Torus.

1.2.9 Die natürlichen Zahlen Wir haben bisher zur Illustration des Mengenbegriffs Beispiele herangezogen, die uns in einem strengen Aufbau der Mathematik noch gar nicht zur Verfügung stehen würden. Eine einführende Vorlesung ist jedoch nicht der geeignete Ort, um eine formale Grundlegung der Mathematik vorzunehmen. Deshalb werden wir uns auch weiterhin die Freiheit herausnehmen, die Mengen \mathbb{Z} , \mathbb{Q} und \mathbb{R} als gegeben anzusehen. Mit derselben Begründung könnten wir das auch mit den natürlichen Zahlen \mathbb{N} tun. Weil der Aufwand jedoch nicht allzu groß ist, wollen wir hier einmal demonstrieren, wie die Mengentheorie dazu eingesetzt werden kann, etwas scheinbar Vertrautes, wie die Zahlen $1, 2, 3, \dots$ einzuführen. Diesen Abschnitt kann man aber auch getrost überschlagen. Wer auf sein bisheriges Verständnis der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, 4, \dots\}$ vertraut, kann nach diesem Abschnitt unbeschadet weiterlesen. Wir stellen VON NEUMANN'S Idee zur Konstruktion der natürlichen Zahlen vor. Wir starten mit der einzigen Menge, die keine weiteren Voraussetzungen benötigt, mit der leeren Menge \emptyset . Sie wird die Rolle der 0 übernehmen (auch wenn wir die natürlichen Zahlen ohne die Null einführen werden). Das nächste Element ist die Menge, die die leere Menge enthält $\{\emptyset\}$. Für sie vereinbaren wir die Abkürzung 1. Es ist klar, wie das weitergehen könnte. Als 2 wählen wir die Menge $\{\{\emptyset\}\}$, als 3 dann $\{\{\{\emptyset\}\}\}$ usw. Das ist tatsächlich eine Möglichkeit, doch VON NEUMANN wählte einen anderen Weg:

$$0 := \emptyset, \quad 1 := \{\emptyset\}, \quad 2 := \{\emptyset, \{\emptyset\}\}, \quad 3 := \{\emptyset, \{\emptyset, \{\emptyset\}\}\}, \quad \dots$$

Bevor wir das allgemeine Schema vorstellen, sehen wir uns erst einmal an, wie die 3 aus der 2 entstanden ist:

$$3 = \{\emptyset, \{\emptyset, \{\emptyset\}\}\} = \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} = 2 \cup \{2\}.$$

Das läßt ein allgemeines Verfahren erkennen: Haben wir das Element n bereits konstruiert, dann definieren wir für seinen Nachfolger n' :

$$n' := n \cup \{n\}. \quad (1.27)$$

Wir führen nun die *natürlichen Zahlen* \mathbb{N} als die Menge aller Nachfolger von 1 ein. Dieser Schritt ist nicht so harmlos, wie er scheint. Unsere Konstruktion stellt uns immer nur endlich viele natürliche Zahlen zur Verfügung, während die Vorstellung sie zu einer Menge zu vereinigen erfordert, daß wir sie in einem geeigneten Sinne alle vorliegen haben. Wenn wir

uns einmal damit abgefunden haben, daß Zahlen durch Mengen beschrieben werden, dann ist die Konstruktion eigentlich ganz überzeugend: Die Anzahl der Elemente in den Mengen $1, 2, 3, \dots$ stimmt mit unserer Vorstellung dieser Zahlen überein. Es gilt $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, $n = \{0, 1, 2, \dots, n-1\}$. Eine Zahl ist immer eine Teilmenge ihres Nachfolgers, denn $n \subset n \cup \{n\} = n'$. Die Addition auf \mathbb{N} müssen wir *rekursiv* definieren. D. h., wir können nicht sofort sagen, was wir unter $n + m$ verstehen wollen, sondern müssen schrittweise vorgehen, indem wir erst einmal $n + 1$ definieren und dann, unter der Voraussetzung, daß wir schon $n + k$ eingeführt haben, sagen was wir unter $n + k + 1$ (genau genommen unter $n + (k + 1)$) verstehen wollen. Konkret bedeutet das: $n + 1 := n'$ und $n + k' := (n + k)'$. Die Idee dahinter ist natürlich, daß nach der Festlegung von $n + 1$, nach k Schritten auch $n + k$ festgelegt ist. Genauso führen wir die Multiplikation ein: $n \cdot 1 := n$, $n \cdot k' = n \cdot (k + 1) := n \cdot k + n$ (diese Konstruktion ist ein Spezialfall des sog. Induktionsprinzips, das weiter unten vorgestellt wird). Die Zahlen $n + k$ bezeichnen wir ebenfalls als Nachfolger von n . Eine Zahl m ist demnach genau dann ein Nachfolger von n , wenn $n \subset m$ gilt. Das bedeutet, daß die \subseteq -Relation hier das gewohnte \leq auf \mathbb{N} modelliert. \subset ist dann mit $<$ zu übersetzen. Damit haben wir die natürlichen Zahlen nur mit Hilfe der Mengentheorie eingeführt, auch wenn dabei einige formale Begründungen unterschlagen wurden. Wir haben die natürlichen Zahlen ohne die Null definiert. Mitunter ist es aber wünschenswert die Menge \mathbb{N} um die Null zu erweitern. Wir schreiben dafür $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

1.2.10 Das Induktionsprinzip Für eine Teilmenge M von \mathbb{N} seien die folgenden beiden Eigenschaften erfüllt:

$$1 \in M, \quad (\text{I-1})$$

$$n \in M \Rightarrow n + 1 \in M. \quad (\text{I-2})$$

Dann besagt das *Induktionsprinzip*, daß daraus bereits $M = \mathbb{N}$ folgt. Anschaulich besagt es die einsichtige Eigenschaft, daß M mit 1 ihren Nachfolger 2, dann dessen Nachfolger 3 usw. enthalten muß und daß auf diese Weise eben alle Elemente von \mathbb{N} erreicht werden.

Dieses Prinzip läßt sich, je nach Einführung von \mathbb{N} , beweisen, oder als Teil des Axiomensystems der natürlichen Zahlen ansehen (dahinter stecken die sog. PEANO-Axiome, die wir hier aber nicht weiter erörtern werden, da wir die natürlichen Zahlen ab jetzt als intuitiv gegeben ansehen wollen). Wir nehmen den zweiten Standpunkt ein, da das Prinzip einsichtig ist.

1.2.11 Das Prinzip der vollständigen Induktion Wir stellen uns vor, daß nach einer bestimmten Vorschrift für jede natürliche Zahl $n \in \mathbb{N}$ eine Aussage A_n gegeben ist. Ein typisches Beispiel lautet etwa: A_n : Die Summe der ersten n Zahlen ist $\frac{1}{2} n(n + 1)$. Unsere Aufgabe besteht darin, diese Aussagen auf ihren Wahrheitsgehalt hin zu untersuchen, im besten Fall also, diese unendlich vielen Aussagen zu beweisen. Daß das nicht durch viel Fleiß zu bewältigen ist, indem man sich die Aussagen der Reihe nach vornimmt, ist einsichtig. In unserem Leben können wir nur endlich viele Aufgaben erledigen, so daß wir auf diese Weise niemals *alle* Aussagen beweisen können. Hier hilft uns das Induktionsprinzip weiter. Wir definieren zu diesem Zweck die Menge

$$M := \{ n \in \mathbb{N} \mid A_n \text{ ist wahr} \} \subseteq \mathbb{N}.$$

Alle Aussagen A_n sind sicher dann wahr, wenn $M = \mathbb{N}$ gilt. Wir müssen daher nur (I-1) und (I-2) des Induktionsprinzips für M nachweisen, um auf $M = \mathbb{N}$ und daraus auf die Wahrheit aller

Aussagen A_n schließen zu können. Im Einzelnen bedeutet das: Wir müssen A_1 beweisen, denn dann gilt $1 \in M$. Dann müssen wir $n \in M \Rightarrow n+1 \in M$ zeigen. Das bedeutet, wir nehmen die Aussage A_n als wahr an ($n \in M$) und müssen mit Hilfe dieser Information auf die Wahrheit von A_{n+1} schließen ($n+1 \in M$). Entscheidend ist, daß wir nun nicht jedes einzelne A_n beweisen müssen, sondern nur noch den Mechanismus, der die Wahrheit von A_n an A_{n+1} weiterreicht, sicherzustellen haben. Ist das nämlich gelungen, dann brauchen wir uns um die Wahrheit von A_2 keine Gedanken mehr machen, da aus der (bewiesenen) Wahrheit von A_1 sofort die von A_2 , dann die von A_3 usw. folgt.

1.2.12 Satz (Vollständige Induktion) *Für jedes $n \in \mathbb{N}$ sei eine Aussage A_n gegeben. Alle Aussagen sind wahr, falls die folgenden beiden Eigenschaften erfüllt sind.*

i) A_1 ist wahr. (Induktionsanfang)

ii) Ist A_n wahr, dann auch A_{n+1} . (Induktionsschritt)

Den Beweis haben wir oben über das Induktionsprinzip bereits erbracht. Führen wir das Verfahren einmal am obigen Beispiel vor. Die Behauptung A_n lautet

$$1 + 2 + \cdots + n = \frac{1}{2}n(n+1). \quad (1.28)$$

i) Der Induktionsanfang (meist einfach durch „ $n = 1$ “ gekennzeichnet):

A_1 behauptet $1 = \frac{1}{2} \cdot 1 \cdot 2$, was offensichtlich wahr ist.

ii) Der Induktionsschritt (durch „ $n \rightarrow n+1$ “ abgekürzt):

Wir dürfen (und müssen) A_n verwenden, d. h. wir nehmen die Aussage $1+2+\cdots+n = \frac{1}{2}n(n+1)$ für bare Münze. Wir kennzeichnen ihre Verwendung im Beweis an geeigneter Stelle durch IV (für Induktionsvoraussetzung). Mit ihrer Hilfe müssen wir auf A_{n+1} : $1+2+\cdots+n+1 = \frac{1}{2}(n+1)(n+2)$ schließen. Es ist mitunter hilfreich, wenn man sich die Aussage A_{n+1} einmal explizit aufschreibt, um zu sehen, was denn eigentlich gezeigt werden muß. Bisher ist allerdings noch nichts passiert, wir haben sozusagen nur die Werkzeuge bereitgelegt. Starten wir:

$$\begin{aligned} 1 + 2 + \cdots + n + 1 &= 1 + 2 + \cdots + n + n + 1 \stackrel{\text{IV}}{=} \frac{1}{2}n(n+1) + n + 1 \\ &= (n+1)\left(\frac{1}{2}n + \frac{2}{2}\right) = \frac{1}{2}(n+1)(n+2). \end{aligned}$$

Wir beginnen auf der linken Seite der Behauptung und formen sie soweit um, daß die Induktionsvoraussetzung erkennbar wird. Nachdem wir diese angewandt haben, ist der Rest nur noch einfaches Rechnen, um zur rechten Seite von A_{n+1} zu gelangen. Man kann sich denken, daß vollständige Induktion beileibe nicht immer nach diesem simplen Schema ablaufen kann. Dazu ist dieses Werkzeug zu allgemein angelegt. Trotzdem sind die Grundzüge eines typischen Induktionsbeweises erkennbar. Die Schwierigkeit im Induktionsschritt besteht meist darin, herauszufinden, wie A_n in der Aussage A_{n+1} angewendet werden kann. Dafür gibt es kein Schema, man muß die Idee normalerweise für jede Induktion neu entwickeln.

1.2.13 Der binomische Lehrsatz

1.2.14 Definition Für $n, k \in \mathbb{N}_0$ und $n \geq k$ werden durch

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \quad (1.29)$$

die sogenannten Binomialkoeffizienten (gesprochen: n über k) definiert. Dabei ist

$$n! := 1 \cdot 2 \cdot \dots \cdot (n-1) \cdot n, \quad 0! := 1, \quad (1.30)$$

die Fakultät von n (gesprochen: n Fakultät).

1.2.15 A Zeigen Sie, daß für $1 \leq k \leq n$ die Gleichung

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \quad (1.31)$$

gilt. Beweisen Sie damit, daß alle Binomialkoeffizienten natürliche Zahlen sind.

1.2.16 Satz (Binomischer Lehrsatz) Es gilt

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}. \quad (1.32)$$

Mit der Notation, die uns bisher zur Verfügung steht, hätten wir die Behauptung des Satzes eigentlich folgendermaßen aufschreiben müssen:

$$\begin{aligned} (a+b)^n &= \binom{n}{0} b^n + \binom{n}{1} a b^{n-1} + \binom{n}{2} a^2 b^{n-2} + \binom{n}{3} a^3 b^{n-3} + \dots \\ &\quad \dots + \binom{n}{k} a^k b^{n-k} + \dots + \binom{n}{n-1} a^{n-1} b + \binom{n}{n} a^n. \end{aligned}$$

Diese Schreibweise ist auf die Dauer zu schwerfällig. Dabei steckt ein einfacher Mechanismus hinter dieser Summe: Alle Summanden entstehen aus einer einzigen Formel, nämlich $\binom{n}{k} a^k b^{n-k}$, indem der Reihe nach der *Summationsindex* k durch $0, 1, 2, \dots, n-1$ und n ersetzt wird und die dadurch entstandenen Ausdrücke aufsummiert werden. Genau das beschreibt das Summenzeichen \sum :

$$\sum_{k=m}^n a_k := a_m + a_{m+1} + a_{m+2} + \dots + a_{n-1} + a_n. \quad (1.33)$$

Dieselbe Aufgabe übernimmt das Produktzeichen \prod für Produkte

$$\prod_{k=m}^n a_k := a_m \cdot a_{m+1} \cdot a_{m+2} \cdot \dots \cdot a_{n-1} \cdot a_n. \quad (1.34)$$

Wir werden gleich sehen, daß das Summenzeichen nicht nur eine bequeme Abkürzung ist, sondern daß man damit wirklich rechnen kann.

Beweis. Der Beweis wird mittels vollständiger Induktion geführt. Für $n = 1$ gilt, wegen $\binom{1}{0} = \binom{1}{1} = 1$:

$$\sum_{k=0}^1 \binom{1}{k} a^k b^{1-k} = \binom{1}{0} b + \binom{1}{1} a = (a+b)^1.$$

Das zeigt die Behauptung für $n = 1$.

Der Induktionsschritt $n \rightarrow n+1$:

$$\begin{aligned} (a+b)^{n+1} &= (a+b)(a+b)^n \stackrel{IV}{=} (a+b) \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} a^{\overbrace{k+1}^{=: \ell}} b^{n-k} + \sum_{k=0}^n \binom{n}{k} a^k b^{n+1-k} \\ &= \sum_{\ell=1}^{n+1} \binom{n}{\ell-1} a^\ell b^{n-(\ell-1)} + \sum_{k=0}^n \binom{n}{k} a^k b^{n+1-k} \\ &= \binom{n}{n} a^{n+1} + \sum_{\ell=1}^n \binom{n}{\ell-1} a^\ell b^{n+1-\ell} + \sum_{k=1}^n \binom{n}{k} a^k b^{n+1-k} + \binom{n}{0} b^{n+1} \\ &= a^{n+1} + \sum_{k=1}^n \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n+1-k} + b^{n+1} \\ &\stackrel{(1.31)}{=} \binom{n+1}{n+1} a^{n+1} b^0 + \sum_{k=1}^n \binom{n+1}{k} a^k b^{n+1-k} + \binom{n+1}{0} a^0 b^{n+1} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}. \quad \square \end{aligned}$$

1.2.17 Bemerkung Zur konkreten Berechnung eines Binoms wird der Satz meist nicht verwendet. Dazu bietet sich ein Verfahren an, das mit dem sogenannten PASCALSchen Dreieck verbunden ist. Stellen wir uns vor, wir benötigen tatsächlich den Ausdruck $(a+b)^6$. Nach dem binomischen Lehrsatz (1.32) müßten wir dafür

$$(a+b)^6 = \binom{6}{6} a^6 + \binom{6}{5} a^5 b + \binom{6}{4} a^4 b^2 + \binom{6}{3} a^3 b^3 + \binom{6}{2} a^2 b^4 + \binom{6}{1} a b^5 + \binom{6}{0} b^6$$

berechnen. In dieser Formel erfordern nur die Binomialkoeffizienten $\binom{6}{k}$ wirkliche Rechenarbeit, denn die Abfolge der Potenzen $a^k b^{6-k}$ gehorcht einer starren Regel:

$$(a+b)^6 = a^6 + ? \cdot a^5 b + ? \cdot a^4 b^2 + ? \cdot a^3 b^3 + ? \cdot a^2 b^4 + ? \cdot a b^5 + b^6.$$

Wenn wir die Koeffizienten auf andere Weise, als durch Berechnung von $\binom{6}{k}$, erhalten könnten, ließe sich das Ergebnis ohne großen Aufwand hinschreiben. Das PASCALSche Dreieck ist dafür das Mittel der Wahl. Das Bildungsgesetz läßt sich leicht erraten. Jede neue Zeile entsteht aus der vorhergehenden, indem man jeweils zwei benachbarte Einträge addiert und das Ergebnis

2 Relationen und Funktionen

2.1 Allgemeine Eigenschaften von Relationen

2.1.1 Definition Gegeben seien zwei Mengen A und B . Dann ist eine (zweistellige) Relation R eine Teilmenge von $A \times B$:

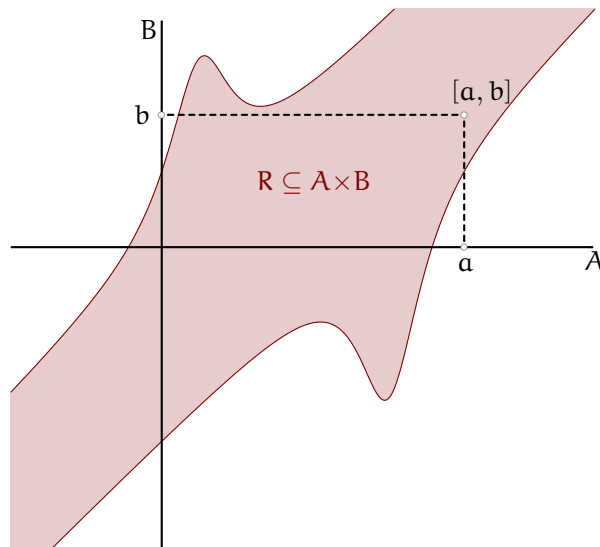
$$R \subseteq A \times B.$$

Wir sagen, ein Element $a \in A$ steht in der Relation R zu einem Element $b \in B$, wenn $[a, b] \in R$ gilt. Dafür wird aRb , manchmal auch $R(a, b)$ geschrieben.

Sind allgemeiner n Mengen A_1, A_2, \dots, A_n gegeben, so ist eine n -stellige Relation R eine Teilmenge von $A_1 \times A_2 \times \dots \times A_n$:

$$R \subseteq A_1 \times A_2 \times \dots \times A_n.$$

Wir sagen, die Elemente $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$ stehen in der Relation R , falls $[a_1, a_2, \dots, a_n] \in R$ gilt. Dafür schreiben wir $R(a_1, a_2, \dots, a_n)$.



Dieser Ansatz erscheint zunächst sehr abstrakt. Tatsächlich bildet er aber auf einfache Weise das ab, was wir machen, wenn wir Dinge in eine Beziehung (Relation) zueinander setzen. Stellen wir uns als Beispiel den Fuhrpark einer (kleinen) Autovermietung vor. Praktischerweise werden wir den Bestand an Fahrzeugen in Form einer Tabelle darstellen. Dann wird das Fahrzeug mit

Fahrz.Nr.	Fabrikat	Kennzeichen	Farbe	Baujahr
1	BMW	X - Y 123	blau	2008
2	VW	Z - U 456	rot	2009
3	Fiat	V - W 100	gelb	2003

Tabelle 2.1 *Fuhrpark*

der Fahrzeugnummer 2 durch den Datensatz [2, VW, Z - U 456, rot, 2009] dargestellt. Das ist ein Element der Produktmenge

$$\mathbb{N} \times \text{Autos} \times \text{Kennzeichen} \times \text{Farben} \times \mathbb{N}_{>1999},$$

die aus den natürlichen Zahlen \mathbb{N} für die Fahrzeugnummer, der Menge aller lieferbaren Autos, der Menge der von der Straßenverkehrsbehörde zu vergebenden Autokennzeichen, der Menge der gängigen Farben und den natürlichen Zahlen größer als 1999 gebildet wird. Wir können den Fuhrpark also auch als eine Teilmenge *Fuhrpark* dieser Produktmenge auffassen:

$$[2, \text{VW}, \text{Z - U 456}, \text{rot}, 2009] \in \mathbb{N} \times \text{Autos} \times \text{Kennzeichen} \times \text{Farben} \times \mathbb{N}_{>1999}.$$

[3, Mercedes, A - BC 1, pink, 2001] ist zwar auch ein Element der Produktmenge, gehört aber offensichtlich nicht zur Relation *Fuhrpark*.

Wenn wir uns das Beispiel weiterdenken, bieten sich noch andere Relationen an, z. B. die Relationen *Kunden* und *Verleih*:

Kunden.Id.	Name	Vorname
001	Lustig	Peter
002	Riese	Adam
007	Bond	James

Tabelle 2.2 *Kunden*

Kunden.Id.	Fahrz.Nr.
007	1
007	3
001	2

Tabelle 2.3 *Verleih*

Die Relation $\text{Verleih}(007, 3) \in \mathbb{N}_0^3 \times \mathbb{N}$ bedeutet dann, daß der gelbe Fiat an den Kunden James Bond ausgeliehen wurde.

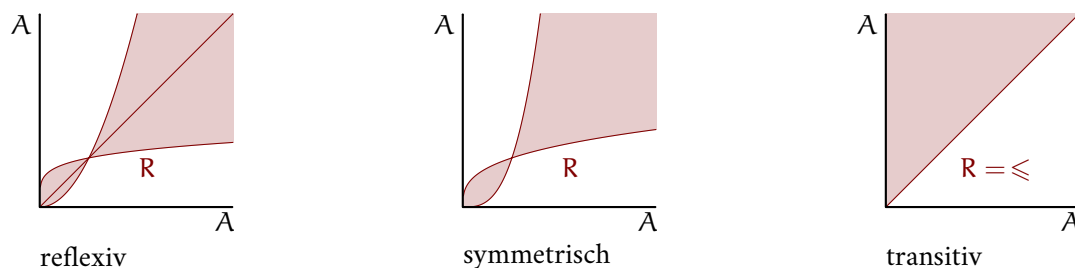
2.2 Klassifikation von Relationen

2.2.1 Definition Eine Relation $R \subseteq A \times B$ heißt

<i>homogen</i>	$:\Leftrightarrow$	$A = B$,
<i>reflexiv</i>	$:\Leftrightarrow$	R ist homogen und f. a. $a \in A$ ist $[a, a] \in R$,
<i>symmetrisch</i>	$:\Leftrightarrow$	R ist homogen und aus $[a, b] \in R$ folgt $[b, a] \in R$,
<i>antisymmetrisch</i>	$:\Leftrightarrow$	R ist homogen und aus $[a, b] \in R$ und $[b, a] \in R$ folgt $a = b$,
<i>transitiv</i>	$:\Leftrightarrow$	R ist homogen und aus $[a, b] \in R$, $[b, c] \in R$ folgt $[a, c] \in R$,
<i>linkstotal</i>	$:\Leftrightarrow$	f. a. $a \in A$ gibt es ein $b \in B$, so daß $[a, b] \in R$,
<i>rechtstotal</i>	$:\Leftrightarrow$	f. a. $b \in B$ gibt es ein $a \in A$, so daß $[a, b] \in R$,
<i>funktional</i>	$:\Leftrightarrow$	aus $[a, b] \in R$ und $[a, c] \in R$ folgt $b = c$,
<i>injektiv</i>	$:\Leftrightarrow$	aus $[a, b] \in R$ und $[c, b] \in R$ folgt $a = c$,
<i>bijektiv</i>	$:\Leftrightarrow$	f. a. $b \in B$ gibt es genau ein $a \in A$, so daß $[a, b] \in R$.

Rechtstotale Relationen werden meist als surjektiv bezeichnet.

Das Adjektiv *bijektiv* wird üblicherweise nur für funktionale Relationen gebraucht. Man überlegt sich leicht: Eine Relation ist genau dann bijektiv, wenn sie sowohl injektiv als auch surjektiv ist.



Die für uns wichtigsten Relationen sind die *Ordnungsrelationen*, die *Äquivalenzrelationen* und natürlich die funktionalen Relationen, denn sie definieren *Funktionen*. Erstere ist eine homogene, antisymmetrische, reflexive und transitive Relation auf einer Menge A . Sie ist nach der gewohnten \leq -Relation auf \mathbb{R} modelliert. Deshalb schreibt man auch meist $a \leq b$ für eine solche Relation und nicht $[a, b] \in \leq$. Die abstrakten Eigenschaften sind in \mathbb{R} sofort einsichtig: Die Reflexivität bedeutet $a \leq a$, die Antisymmetrie, daß $a \leq b$ und $b \leq a$ nur für $a = b$ möglich ist und die Transitivität, daß aus $a \leq b$ und $b \leq c$ auch $a \leq c$ folgen muß.

2.3 Äquivalenzrelationen

Das Rechnen modulo p wird oft als Gelegenheit genutzt, um Äquivalenzrelationen einzuführen. Dabei werden zwei ganze Zahlen a und b als *gleich modulo einer natürlichen Zahl p* angesehen, wenn sich b von a um ein ganzzahliges Vielfaches von p unterscheidet: $b = a + t \cdot p$, $t \in \mathbb{Z}$, geschrieben als $a \equiv b \pmod{p}$ (vergl. Seite 40). Wenn wir diese Relation vorübergehend mit R bezeichnen, dann ist $R \subseteq \mathbb{Z} \times \mathbb{Z}$ eine homogene, reflexive, symmetrische und transitive Relation. Die Reflexivität ist wegen $a = a + 0 \cdot p$ erfüllt, die Symmetrie, weil aus $a = b + t \cdot p$

natürlich $b = a + (-t) \cdot p$ folgt. Die Transitivität ist erfüllt, da aus $a = b + t \cdot p$ und $b = c + s \cdot p$ sofort $a = c + (s + t) \cdot p$ folgt. Eine Relation auf einer beliebigen Menge nach diesem Muster heißt Äquivalenzrelation. Für sie wird oft das Symbol \sim verwendet.

2.3.1 Definition Eine Relation $\sim \subseteq A \times A$ auf einer Menge A mit den Eigenschaften

i) $a \sim a$ f. a. $a \in A$, (Reflexivität)

ii) $a \sim b \Rightarrow b \sim a$, (Symmetrie)

iii) $a \sim b \wedge b \sim c \Rightarrow a \sim c$, (Transitivität)

heißt Äquivalenzrelation auf A . Die Menge $[a]_{\sim} := \{ b \in A \mid a \sim b \}$ heißt Äquivalenzklasse der Relation \sim zum Repräsentant a . Für $a \sim b$ sagen wir: a ist zu b äquivalent.

Meist schreiben wir für eine Äquivalenzklasse einfach $[a]$, wenn es klar ist, welche Äquivalenzrelation gemeint ist. Der Repräsentant a einer Äquivalenzklasse ist normalerweise keineswegs eindeutig. Es ist $[a] = [a']$ für zwei verschiedene Elemente a und a' möglich – und zwar genau dann, wenn $a \sim a'$ gilt, also wenn die Repräsentanten äquivalent sind. Das wollen wir uns jetzt überlegen. Zunächst zeigen wir, daß aus $a \sim a'$ die Gleichheit von $[a]$ und $[a']$ folgt: Wegen $a \sim a'$ liegt a' in $[a]$. Für ein $b \in [a']$ muß $b \sim a'$ erfüllt sein. Wegen der Transitivität folgt dann aus $a' \sim a$ auch $b \sim a$, also $b \in [a]$. Jedes Element b von $[a']$ liegt daher in $[a]$. Damit haben wir $[a'] \subseteq [a]$ gezeigt. Da unsere Argumentation bezüglich a und a' symmetrisch ist, haben wir auch $[a] \subseteq [a']$ und daher $[a] = [a']$.

Für die andere Richtung gehen wir von $[a] = [a']$ aus. Das bedeutet aber, da a' wegen der Reflexivität in $[a']$ liegt, daß a' ein Element von $[a]$ ist, was per Definition $a \sim a'$ bedeutet. Damit ist alles gezeigt.

Eine unmittelbare Folgerung ist, daß zwei verschiedene Äquivalenzklassen $[a]$ und $[b]$ disjunkt sein müssen. Gäbe es nämlich ein $c \in [a] \cap [b]$, dann hieße das $a \sim c$ und $c \sim b$, woraus wegen der Transitivität sofort $a \sim b$ und, nach unseren Überlegungen, der Widerspruch $[a] = [b]$ entstehen würde.

Die Menge aller Äquivalenzklassen besteht aus lauter disjunkten Mengen, deren Vereinigung ganz A ergibt. Denn jedes Element $a \in A$ gehört zu $[a]$ und damit zu $\bigcup_{b \in A} [b] \subseteq A$. Das bedeutet $A \subseteq \bigcup_{b \in A} [b]$, oder $A = \bigcup_{b \in A} [b]$ (man beachte, daß in dieser Vereinigung viele Mengen gleich sind). Die Äquivalenzklassen bilden eine sogenannte *disjunkte Überdeckung* von A . Darunter verstehen wir folgendes: Es gibt eine Indexmenge I und für jedes $i \in I$ eine Teilmenge A_i von A , so daß $A_i \cap A_j = \emptyset$ für $i \neq j$ und $\bigcup_{i \in I} A_i = A$ gilt. Eine solche Überdeckung schreiben wir kurz in der Form $\{ A_i \mid i \in I \}$. Als Indexmenge I wählen wir in unserem Kontext die Menge der verschiedenen Äquivalenzklassen $I := \bigcup_{a \in A} [a]$ (warum definieren wir nicht einfach $I := \{ [a] \mid a \in A \}$?) und für die disjunkten Mengen die Elemente von I (also die Äquivalenzklassen) selbst. Jetzt können wir unsere obigen Überlegungen etwas verfeinern: Jedes $a \in A$ gehört zu $[a]$ und damit zu einer der Äquivalenzklassen aus I . Das bedeutet $A \subseteq \bigcup_{[b] \in I} [b] \subseteq A$, also $A = \bigcup_{[b] \in I} [b]$.

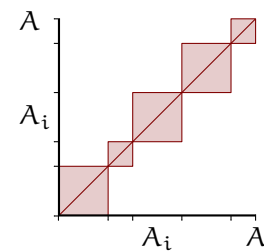
Damit haben wir die Hälfte des folgenden Satzes bereits bewiesen.

2.3.2 Satz Für jede Äquivalenzrelation \sim auf einer Menge A bilden die verschiedenen Äquivalenzklassen $[a]_{\sim}$, $a \in A$, eine disjunkte Überdeckung von A . Umgekehrt definiert jede disjunkte Überdeckung $\{A_i \mid i \in I\}$ von A auf folgende Weise eine Äquivalenzrelation: $a \sim b \Leftrightarrow \exists i \in I, a, b \in A_i$. Die zugehörigen Äquivalenzklassen sind dabei durch die Mengen A_i gegeben.

Beweis. Wir gehen von einer disjunkten Überdeckung $\{A_i \mid i \in I\}$ von A aus und müssen zeigen, daß durch $a \sim b \Leftrightarrow \exists i \in I, a, b \in A_i$ eine Äquivalenzrelation definiert wird. Die Reflexivität folgt aus der Tatsache, daß jedes $a \in A$ in genau einer der Mengen A_i liegen muß, da diese Mengen ja ganz A überdecken und paarweise disjunkt sind. Die Symmetrie ist offensichtlich. Bleibt die Transitivität: Aus $a \sim b$ und $b \sim c$ folgt zunächst, daß es genau ein $i \in I$ mit der Eigenschaft $a, b \in A_i$ gibt. Dann bedeutet aber $b \sim c$, daß auch c zu A_i gehören muß. Es folgt $a, c \in A_i$, also $a \sim c$. Die Bedingungen von 2.3.1 sind damit nachgewiesen. Die Äquivalenzklassen $[a]$ sind laut Definition $[a] = \{b \in A \mid a \sim b\}$. Andererseits liegt a in genau einer Menge A_i . Alle zu a äquivalenten Elemente b liegen gemäß der Definition der Äquivalenzrelation aber ebenfalls in A_i . Das zeigt $[a] \subseteq A_i$. Andererseits gilt für jedes $b \in A_i$ natürlich $a, b \in A_i$, also $a \sim b$, oder $b \in [a]$. Das hat die fehlende Inklusion $A_i \subseteq [a]$ zur Folge. \square

Im Lichte dieses Satzes sollte die symbolische Darstellung einer Äquivalenzrelation wie in der nebenstehenden Skizze aussehen.

Meist will man mit einer Äquivalenzrelation Eigenschaften der Ausgangsmenge A in die Menge der Äquivalenzklassen übertragen. Etwa bei dem Beispiel $x \sim y \Leftrightarrow x = y \pmod p$ auf der Menge $A := \mathbb{Z}$ will man die Grundrechenarten \pm und \cdot auf der Menge der Äquivalenzklassen zur Verfügung haben. Dafür schreibt man dann A/\sim und spricht vom *Schnitt der Menge A nach der Relation \sim* . Es haben sich aber auch andere Schreibweisen für Konstruktionen dieser Art eingebürgert, wie wir gleich sehen werden. In unserem Beispiel haben wir die Rechenoperationen \pm und \cdot auf den Äquivalenzklassen $[x]$ und $[y]$ zu definieren. Das machen wir folgendermaßen:



$$[x] + [y] := [x + y], \quad (2.1)$$

$$[x] - [y] := [x - y], \quad (2.2)$$

$$[x] \cdot [y] := [x \cdot y]. \quad (2.3)$$

Diese Definitionen haben wir mit Hilfe der Repräsentanten x und y von $[x]$ bzw. $[y]$ formuliert. Da diese durch die Äquivalenzklassen aber nicht eindeutig festgelegt sind, müssen wir uns davon überzeugen, daß die Definition nicht in sich widersprüchlich ist. Man sagt, wir müssen prüfen, daß die Rechenoperationen *wohldefiniert* sind. Wir wissen, daß die Äquivalenzklassen $[x]$ und $[y]$ auch durch die zu x bzw. y äquivalenten Zahlen x' bzw. y' erzeugt werden: $[x] = [x']$, $[y] = [y']$. Für die linke Seite von (2.1) könnten wir auch $[x'] + [y']$ schreiben. Wir müssen also zeigen, daß die rechte Seite $[x + y]$ nicht von der speziellen Wahl von x und y abhängt. Das ist sicher dann der Fall, wenn wir $[x + y] = [x' + y']$ nachweisen können. Jetzt verwenden wir die Definition der Äquivalenzrelation: $x \sim x' \Leftrightarrow \exists s \in \mathbb{Z} x' = x + sp$, $y \sim y' \Leftrightarrow \exists t \in \mathbb{Z} y' = y + tp$. Es folgt $x' \pm y' = x \pm y + (s \pm t)p \sim x \pm y$ und $x' \cdot y' = x \cdot y + (sy + tx + st)p \sim x \cdot y$.

Damit haben wir die Wohldefiniertheit gleich für alle Rechenoperationen (2.1) – (2.3) gezeigt. Natürlich läßt sich jetzt auch $[x]^n := [x^n]$ für $n \in \mathbb{N}$ definieren. In der Zahlentheorie auf Seite 40 führen wir letztlich diese Rechengesetze für das Rechnen modulo p ein, auch wenn wir dort nicht diesen formalen Aufwand treiben (wir vereinbaren die Schreibweise $x =_p y$ für $x = y \bmod p$ und zeigen wie oben, daß sich diese neue Gleichheit mit den Grundrechenarten \pm und \cdot verträgt). Die Äquivalenzklasse $[0]$ besteht einfach aus allen ganzzahligen Vielfachen von p . Das ist eine Menge, die man suggestiv durch $p\mathbb{Z}$ bezeichnet. Sie legt die Äquivalenzrelation fest, denn $x \sim y \Leftrightarrow x - y \in p\mathbb{Z}$. Deshalb hat sich für die Menge der Äquivalenzklassen mit den eingeführten Rechenoperationen das aussagekräftige Symbol $\mathbb{Z}/p\mathbb{Z}$ durchgesetzt.

2.4 Verkettung von Relationen und die inverse Relation

2.4.1 Definition Für zwei Relationen $R \subseteq A \times B$ und $S \subseteq B \times C$ definiert

$$S \circ R := \{ [a, c] \in A \times C \mid \exists b \in B [a, b] \in R \wedge [b, c] \in S \} \quad (2.4)$$

die Verkettung der Relationen R und S .

$$R^{-1} := \{ [b, a] \in B \times A \mid [a, b] \in R \} \quad (2.5)$$

heißt inverse Relation von R .

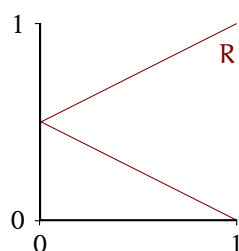
In dieser Allgemeinheit ist es durchaus möglich, daß $S \circ R$ die sogenannte *Nullrelation* $O := \emptyset \subseteq A \times C$ ist. Da für $[b, c] \in S$ nicht alle Elemente aus B vorkommen müssen, kann es sein, daß für kein Paar $[a, c] \in A \times C$ ein gemeinsames $b \in B$ mit der Eigenschaft $[a, b] \in R$ und $[b, c] \in S$ gefunden werden kann.

Die Rolle von R^{-1} zeigt sich bei der Verkettung mit R :

$$\begin{aligned} R^{-1} \circ R &= \{ [a, a'] \in A \times A \mid \exists b \in B [a, b] \in R \wedge [b, a'] \in R^{-1} \} \\ &= \{ [a, a'] \in A \times A \mid \exists b \in B [a, b] \in R \wedge [a', b] \in R \}. \end{aligned}$$

Offensichtlich ist mit jedem $[a, b] \in R$ das Paar $[a, a]$ in $R^{-1} \circ R$. Diese Menge kann aber weitaus mehr Elemente als diese sogenannten *Diagonalelemente* enthalten. Etwa für $A = B = \mathbb{R}$ und der \leq -Relation R ist R^{-1} die \geq -Relation. Daher ist $R^{-1} \circ R = \{ [a, a'] \in \mathbb{R} \times \mathbb{R} \mid \exists b \in \mathbb{R} a \leq b \wedge a' \leq b \} = \mathbb{R} \times \mathbb{R}$, denn für je zwei reelle Zahlen a und a' gibt es immer eine reelle Zahl b , die beide übertrifft.

Man sollte also nicht zu viel hinter der Bezeichnung *Inverse* vermuten, wie z. B., daß $R^{-1} \circ R$ eine Teilmenge der *identischen Relation* (oder *Gleichheitsrelation*) $I_A := \{ [a, a] \in A \times A \mid a \in A \}$ ist. Dafür muß aus $[a, b] \in R$ und $[a', b] \in R$ immer $a = a'$ folgen,



d. h., R hat dafür injektiv zu sein. Soll auch noch $R \circ R^{-1} \subseteq I_B$ gelten, so muß aus $[b, a] \in R^{-1}$ und $[b', a] \in R^{-1}$, also aus $[a, b] \in R$ und $[a, b'] \in R$ immer $b = b'$ folgen. Das bedeutet, daß R eine funktionale Relation ist. Beide Eigenschaften sind daher nur für injektive Funktionen (s.u.) erfüllt. Man könnte meinen, daß eine injektive Relation automatisch auch funktional sein muß, aber die Relation R aus der nebenstehenden Skizze zeigt, daß das nicht der Fall sein muß.

2.5 Funktionen

Für eine funktionale Relation $\emptyset \neq f \subseteq A \times B$ definieren wir die Mengen

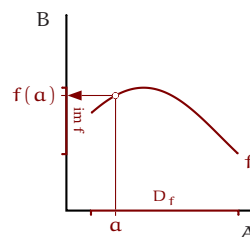
$$D_f := \{ a \in A \mid \exists b \in B [a, b] \in f \}, \quad (2.6)$$

$$\text{im } f := \{ b \in B \mid \exists a \in A [a, b] \in f \}. \quad (2.7)$$

$D_f \subseteq A$ heißt *Definitionsbereich* und $\text{im } f$ *Bild* von f . Solche Relationen bezeichnen wir als *Funktionen* und vereinbaren statt $[a, b] \in f$ oder $a \mapsto b$ die üblichen Schreibweisen $f : A \rightarrow B$, $a \mapsto f(a) = b$, (streng genommen ist diese Schreibweise dem Fall $D_f = A$ vorbehalten) oder ausführlicher $f : A \supseteq D_f \rightarrow B$, $a \mapsto f(a)$. Das Bild von f ist jetzt $\text{im } f = \{ b \in B \mid \exists a \in A b = f(a) \}$.

Wieso beschreiben funktionale Relationen genau das, was wir unter einer Funktion verstehen?

Die gängige Vorstellung von einer Funktion ist doch, daß aus einem Vorrat von Werten a aus einer Menge A , der *Urbildmenge*, einer geeigneten Vorschrift folgend, Bildelemente $f(a)$ aus einer Zielmenge B gebildet werden. In die Sprache der Relationen übersetzt heißt das: Jedes $a \in A$ darf mit nicht mehr als einem Element $b \in B$ in Relation f stehen. Das ist genau der Inhalt der Definition einer funktionalen Relation: Aus $[a, b] \in f$ und $[a, c] \in f$ folgt $b = c$. In der üblichen Notation heißt das: Aus $f(a) = b$ und $f(a) = c$ folgt $b = c$. Die Teilmenge f von $A \times B$ können wir jetzt durch $f = \{ [a, f(a)] \mid a \in D_f \}$ wiedergeben. f ist demnach die Menge, die man den *Graphen* der Funktion f nennt. Vom Standpunkt der Relationen aus identifiziert man also den Graphen einer Funktion mit der Funktion selbst. Aus diesem Grund werden wir oft von einer Funktion f sprechen und dabei den Definitionsbereich und die Abbildungsvorschrift als gegeben annehmen.



Wenn die Mengen A und B aus dem Kontext klar ersichtlich sind, gestatten wir es uns aber auch, einfach von *der Funktion* $f(x)$ zu sprechen, obwohl das natürlich streng genommen falsch ist, denn $f(x)$ ist der sogenannte Funktionswert von f und nicht die Funktion. Was damit gemeint ist, zeigt folgendes Beispiel: Für die Funktion f auf \mathbb{R} , die ausführlich durch $f : \mathbb{R} \supseteq \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$, $x \mapsto \frac{x}{1-x}$ definiert ist, können wir dann einfach von der Funktion $f(x) = \frac{x}{1-x}$ sprechen. Wir vertrauen dabei darauf, daß der Leser aus dem Kontext heraus schon erkennen wird, daß es sich hier um eine reelle Funktion handelt, deren Definitionsbereich, mit Ausnahme der 1, alle reellen Zahlen umfasst, und die durch die Rechenvorschrift $\frac{x}{1-x}$ eindeutig festgelegt ist. Sollte dabei ein kleinerer Definitionsbereich D_f , als der maximal mögliche $\mathbb{R} \setminus \{1\}$ gewünscht sein, so muß dieser natürlich noch explizit angegeben werden. Trotz der augenscheinlichen Bequemlichkeit dieser Konvention, werden wir aber versuchen, sie wirklich nur in so übersichtlichen Situationen anzuwenden, wie in diesem Beispiel.

2.5.1 Rechnen mit Funktionen Für zwei Funktionen $f : X \supseteq D_f \rightarrow Y$ und $g : Y \supseteq D_g \rightarrow Z$ läßt sich, falls das Bild $\text{im } f$ im Definitionsbereich D_g von g enthalten ist, durch

$$g \circ f : X \supseteq D_f \rightarrow Z, \quad x \mapsto g(f(x)) \quad (2.8)$$

eine Funktion gewinnen, die als *Verkettung*, oder *Hintereinanderausführung* von g und f bezeichnet wird. Für $f(x) := x^2 + 1$ und $g(y) := \frac{1}{y}$ ist $g \circ f(x) = g(f(x)) = \frac{1}{x^2 + 1}$. Gemäß unserer

Übereinkunft ist der Definitionsbereich dieser Funktionen jeweils der maximal mögliche, da nichts anderes bestimmt wurde. Das heißt $D_f = \mathbb{R}$ und $D_g = \mathbb{R} \setminus \{0\}$. Wenn wir $f(x)$ zu $x^2 - 1$ abändern, lassen sich g und f nicht mehr verketteten, denn $0 \in \text{im } f$, so daß im f nicht mehr in D_g enthalten ist. Die Funktionsvorschrift $g(f(x)) = \frac{1}{x^2 - 1}$ ergibt aber durchaus eine vernünftige Funktion, mit dem natürlichen Definitionsbereich $\mathbb{R} \setminus \{-1, 1\}$. Unsere Version der Verkettung zweier Funktionen ist also zu restriktiv. Wenn wir im $f \subseteq D_g$ zu im $f \cap D_g \neq \emptyset$ abschwächen, läßt sich $g(f(x))$ für x aus einer möglicherweise echten Teilmenge von D_f bilden. Genauer gesagt läßt sich dieser Funktionswert für alle x aus der Menge $\{x \in X \mid f(x) \in D_g\} \subseteq D_f$ bilden.

2.5.2 Definition Für eine Funktion $f : X \supseteq D_f \rightarrow Y$ und eine Teilmenge $B \subseteq Y$ heißt

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\} \quad (2.9)$$

Urbild(menge) von B unter der Funktion f . Für eine Teilmenge $A \subseteq D_f$ heißt die Funktion $f|_A : X \supseteq A \rightarrow Y$, $A \ni x \mapsto f(x)$ die Einschränkung von f auf A . Die Funktion $\text{id} : X \rightarrow X$, $x \mapsto x$ wird identische Funktion genannt.

Natürlich ist $f^{-1}(Y) = D_f$ und $f^{-1}(\emptyset) = \emptyset$. Die etwas unglücklich gewählte Notation $f^{-1}(A)$ hat normalerweise nichts mit einer Umkehrfunktion zu tun (die es beileibe nicht für alle Funktionen gibt, siehe Abschnitt 2.5.5), oder gar mit der Funktion $\frac{1}{f}$. Aus dem Kontext heraus ist aber meist klar, was gemeint ist, denn f^{-1} wird auf eine Menge angewandt und nicht auf das Element einer Menge.

Für die Verkettung zweier Funktionen f und g verlangen wir also nur noch, daß im $f \cap D_g \neq \emptyset$ gilt und haben dadurch den natürlichen Definitionsbereich $D_{g \circ f} := f^{-1}(D_g)$ für $g \circ f$.

Wenn die Zielmenge Y für zwei Funktionen $f : X \supseteq D_f \rightarrow Y$, $g : X \supseteq D_g \rightarrow Y$ eine Addition und eine Multiplikation, gegebenenfalls auch eine Division zur Verfügung stellt, dann lassen sich zu jeder dieser Rechenoperationen auch Funktionen bilden:

$$(f \pm g)(x) := f(x) \pm g(x), \quad D_{f \pm g} = D_f \cap D_g, \quad (2.10)$$

$$(f \cdot g)(x) := f(x) \cdot g(x), \quad D_{f \cdot g} = D_f \cap D_g, \quad (2.11)$$

$$\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}, \quad D_{\frac{f}{g}} = D_f \cap (D_g \setminus g^{-1}(\{0\})), \quad (2.12)$$

$$f^n(x) := (f(x))^n, \quad D_{f^n} = D_f, \quad (n \in \mathbb{N}_0), \quad (2.13)$$

$$f^{-n}(x) := \frac{1}{(f(x))^n}, \quad D_{f^{-n}} = D_f \setminus f^{-1}(\{0\}), \quad (n \in \mathbb{N}). \quad (2.14)$$

Im Lichte dieser Definition sollte man jetzt mit der Übersetzung der Gleichung $\sin^2 + \cos^2 = 1$ keine Probleme haben. Es gilt $\sin^2(x) + \cos^2(x) = 1$, oder ganz ausgeschrieben: $(\sin(x))^2 + (\cos(x))^2 = 1$. Da das aber sehr umständlich aussieht, schreibt man auch $\sin(x)^2 + \cos(x)^2 = 1$ (vergl. (11.14)). Nebenbei gesagt sind Gleichungen dieser Art der Grund dafür, daß wir auch für die bekannten Funktionen \sin , \cos , \tan , \ln etc. strikt an der Funktionsschreibweise festhalten, also $\sin(x)$, $\cos(x)$ usw. statt $\sin x$, $\cos x$ schreiben. So müssen wir nie erklären, daß wir mit $\sin(x)^2$ natürlich $(\sin(x))^2$ meinen und nicht etwa $\sin(x^2)$. In der kürzeren Notation kann

$\sin x^2$ nur $\sin(x^2)$ bedeuten, $\sin(x)^2$ müßte immer durch die schwerfällige Version $(\sin x)^2$ ersetzt werden. Spätestens wenn man Konstruktionen wie $\sin(x+1)$ benötigt, läßt sich die Kurznotation nicht mehr durchhalten, da sonst $\sin x + 1$ nicht von $(\sin x) + 1$ zu unterscheiden wäre.

2.5.3 A Zeigen Sie: Für die Urbildmengen einer Funktion $f : X \rightarrow Y$, $D_f = X$, gelten die Rechenregeln:

$$\begin{aligned} f^{-1}(A \cap B) &= f^{-1}(A) \cap f^{-1}(B), & f^{-1}(A \cup B) &= f^{-1}(A) \cup f^{-1}(B), \\ f^{-1}(A^c) &= f^{-1}(A)^c, & f^{-1}(A \setminus B) &= f^{-1}(A) \setminus f^{-1}(B). \end{aligned}$$

2.5.4 A Prüfen Sie nach, daß die Verkettung der funktionalen Relationen g und f wieder eine funktionale Relation ist, die mit der Verkettung der zu g und f gehörenden Funktionen übereinstimmt.

2.5.5 Die Umkehrfunktion

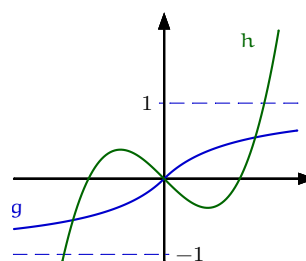
2.5.6 Definition Eine Abbildung $f : X \supseteq D_f \rightarrow Y$ von der Menge X in die Menge Y heißt *injektiv*, falls aus $f(x_1) = f(x_2)$ immer $x_1 = x_2$ folgt. Sie heißt *surjektiv*, falls das Bild $\text{im } f := \{y \in Y \mid \exists x \in X \ y = f(x)\} =: f(X)$ die ganze Menge Y ist. f heißt *bijektiv*, falls f injektiv und surjektiv ist.

Wir haben mit dieser Definition die Eigenschaften injektiv, surjektiv und bijektiv einer funktionalen Relation in die gängige Sprache der Funktionen übertragen.

Die Injektivität einer Abbildung $f : X \rightarrow Y$ wird von Studenten gerne in der Form *f ist injektiv*, wenn jedem x aus X genau ein y aus Y zugeordnet wird wiedergegeben. Das hört sich überzeugend an, ist aber falsch. Diese Eigenschaft muß eine injektive Abbildung sicher haben, denn diese Eigenschaft muß *jede* Abbildung aufweisen – andernfalls wäre sie nämlich keine (sondern?). Die Injektivität verlangt aber darüber hinaus, daß jedes der Bilder $y \in f(X)$ von *jeweils genau einem* $x \in X$ stammt. Diese Eigenschaft wird z. B. von $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^3$ erfüllt, nicht aber von $g : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$. Aus $f(x_1) = f(x_2)$, also $x_1^3 = x_2^3$ folgt nämlich $0 = x_1^3 - x_2^3 = (x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2)$. Das ist nur für $x_1^2 + x_1x_2 + x_2^2 = 0$, oder für $x_1 = x_2$ möglich. Im ersten Fall läßt sich $x_1^2 + 2x_1x_2 + x_2^2 = (x_1 + x_2)^2 = x_1x_2 \geq 0$ folgern. Damit ist $x_1^2 + x_1x_2 + x_2^2$ eine Summe aus nicht negativen Zahlen, die nur für $x_1 = x_2 = 0$ Null ergeben kann. In jedem Fall muß $x_1 = x_2$ gelten. g ist wegen $g(2) = 4 = g(-2)$ nicht injektiv.

Die Surjektivität einer Abbildung $f : X \rightarrow Y$ bedeutet, daß alle Elemente von Y Bilder sind: $\text{im } f = Y$. Die Funktion f von oben ist surjektiv, denn für jedes $y \in \mathbb{R}$ ist $f(\sqrt[3]{y}) = y$. Die Surjektivität läßt sich für eine einzelne Abbildung erzwingen, wenn sie nicht erfüllt sein sollte, indem man die Zielmenge Y gegebenenfalls verkleinert und durch $\text{im } f$ ersetzt.

Meistens ist das aber gar nicht wünschenswert. Für die Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \frac{x}{1+|x|}$ würde es z. B. bedeuten, sie als surjektive



Funktion von \mathbb{R} nach $(-1, 1)$ anzusehen. Für mehrere Funktionen ist das meist sowieso nicht möglich, da sie verschiedene Bildmengen in Y haben können. Von den beiden Funktionen g und $h : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3 - x$ ist g injektiv, aber nicht surjektiv und h surjektiv, aber nicht injektiv.

Für eine injektive Funktion $f : X \supseteq D_f \rightarrow Y$ gibt es für jedes Bildelement $y \in \text{im } f$ genau ein $x \in D_f$, für das $y = f(x)$ gilt. Daher wird durch die Teilmenge $\{[y, x] \in Y \times X \mid [x, y] \in f\}$ von $Y \times X$ wieder eine funktionale Relation, also eine Funktion von Y nach X definiert. Für zwei Elemente $[y, x]$ und $[y, w]$ dieser Menge gilt nämlich $[x, y] \in f$ und $[w, y] \in f$, d. h. $y = f(x)$ und $y = f(w)$. Da f injektiv ist, muß $x = w$ folgen. Das zeigt die funktionale Eigenschaft der Relation. Die zugehörige Funktion wird *Umkehrfunktion* von f genannt und mit f^{-1} bezeichnet. Diese Schreibweise ist für jemanden, der die Potenzrechengesetze kennt, etwas irreführend. Eigentlich verstehen wir unter f^{-1} , gemäß (2.14), die Funktion $\frac{1}{f}$. Aber genau die ist hier *nicht* gemeint. Die Schreibweise rührt daher, daß f^{-1} die Inverse von f bezüglich der Verkettung von Funktionen ist: $f^{-1} \circ f(x) = f^{-1}(f(x)) = x$ für alle $x \in D_f$. $f^{-1} \circ f$ ist demnach die identische Funktion $\text{id} : X \rightarrow X, x \mapsto x$, eingeschränkt auf D_f . Die Gleichung $f^{-1} \circ f = \text{id}$ erinnert an die Inversenbildung bei der Multiplikation gewöhnlicher Zahlen und ist der Grund für die Schreibweise f^{-1} . Trotz der Mehrdeutigkeit hat sich diese Notation für die Umkehrfunktion durchgesetzt, weil aus dem Zusammenhang normalerweise erkennbar ist, ob $\frac{1}{f}$, oder die Umkehrfunktion gemeint ist. An einem einfachen Beispiel wird das klar: Für $f(x) := x^3$ haben wir oben geklärt, daß es sich um eine injektive und surjektive, also eine bijektive Funktion von \mathbb{R} nach \mathbb{R} handelt. Um ihre Umkehrfunktion zu finden, müssen wir die Zuordnung $x \mapsto x^3 = y$ von f umkehren. Das bedeutet aber nichts weiter, als die Gleichung $y = x^3$ nach x aufzulösen. Wir erhalten natürlich $x = \sqrt[3]{y}$. Daher ist $f^{-1}(y) = \sqrt[3]{y}$ und $f^{-1} \circ f(x) = f^{-1}(x^3) = \sqrt[3]{x^3} = x = \text{id}(x)$, während $\frac{1}{f(x)} = \frac{1}{x^3}$ in der Verkettung mit f die Abbildung $x \mapsto \frac{1}{x^9}$ ergäbe. Die Existenz der n -ten Wurzeln überlegen wir uns in Beispiel 2.5.9.

Wie sehen der Definitionsbereich und das Bild von f^{-1} aus? Gemäß (2.6) und (2.7) haben wir

$$\begin{aligned} D_{f^{-1}} &= \{y \in Y \mid \exists x \in X [y, x] \in f^{-1}\} = \{y \in Y \mid \exists x \in X [x, y] \in f\} = \text{im } f, \\ \text{im } f^{-1} &= \{x \in X \mid \exists y \in Y [y, x] \in f^{-1}\} = \{x \in X \mid \exists y \in Y [x, y] \in f\} = D_f. \end{aligned}$$

Für die Umkehrfunktion von $f : X \supseteq D_f \rightarrow \text{im } f \subseteq Y$ gilt daher, wie erwartet, $f^{-1} : Y \supseteq \text{im } f \rightarrow D_f \subseteq X$, denn f^{-1} ordnet ja jedem Bildelement $y = f(x) \in \text{im } f$ sein Urbildelement $x \in D_f$ zu.

2.5.7 A $f : X \supseteq D_f \rightarrow \text{im } f \subseteq Y$ sei eine injektive Funktion, die nach unseren Überlegungen die Umkehrfunktion $f^{-1} : Y \supseteq \text{im } f \rightarrow D_f \subseteq X$ hat. Skizzieren Sie den Graph von f und von f^{-1} als Teilmenge von $X \times Y$. Was fällt Ihnen dabei auf?

2.5.8 A $f : X \supseteq D_f \rightarrow \text{im } f \subseteq Y$ sei eine injektive Funktion. Zeigen Sie, daß $f^{-1} \circ f = \text{id}_{D_f}$ und $f \circ f^{-1} = \text{id}_{\text{im } f}$ gilt.

2.5.9 Beispiel Die identische Funktion $\text{id} : \mathbb{R} \rightarrow \mathbb{R}$ ist offensichtlich bijektiv. Die Potenzfunktion $\text{id}^n : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \text{id}^n(x) = x^n$ ist nur für ungerade n ebenfalls bijektiv ($D_{\text{id}^n} := \mathbb{R}$). Für gerades n müssen wir den Definitionsbereich von \mathbb{R} auf $D_{\text{id}^n} := \mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$ einschränken, um die Bijektivität zu gewährleisten: $\text{id}^n : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$. Diese Behauptungen wollen

wir jetzt zeigen. Dafür benötigen wir eine Verallgemeinerung der dritten binomischen Formel $x^2 - y^2 = (x - y)(x + y)$:

$$\begin{aligned} x^n - y^n &= (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + x^2y^{n-3} + xy^{n-2} + y^{n-1}) \\ &= (x - y) \sum_{k=1}^n x^{n-k}y^k. \end{aligned} \quad (2.15)$$

Wir rechnen die Gleichung einfach nach:

$$\begin{aligned} (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + x^2y^{n-3} + xy^{n-2} + y^{n-1}) \\ = x^n + x^{n-1}y + x^{n-2}y^2 + \dots + x^3y^{n-3} + x^2y^{n-2} + xy^{n-1} \\ - x^{n-1}y - x^{n-2}y^2 - \dots - x^3y^{n-3} - x^2y^{n-2} - xy^{n-1} - y^n = x^n - y^n. \end{aligned}$$

Die Idee für (2.15): Das Polynom $p(x) := x^n - y^n$ hat die Nullstelle $x = y$. Daher muß die Polynomdivision von $p(x)$ mit $x - y$ ohne Rest aufgehen (vergl. 11.8.11).

Wir können $n > 1$ annehmen, da für $n = 1$ nichts zu zeigen ist. Für die Funktion $x \mapsto \text{id}^n(x) = x^n$ müssen wir beweisen, daß aus $x \neq y \in D_{\text{id}^n}$ immer $\text{id}^n(x) \neq \text{id}^n(y)$, also $x^n \neq y^n$ folgt, oder, was dazu äquivalent ist: Aus $x^n = y^n$ muß $x = y$ folgen (vergl. (1.9)).

Aus $x^n = y^n$ ergibt sich also $x - y = 0$, oder $q := x^{n-1} + x^{n-2}y + \dots + xy^{n-2} + y^{n-1} = 0$. Falls $x = y$ gelten sollte, sind wir fertig. Wir nehmen $x \neq y$ an. Bei geradem n gilt $x \geq 0$ und $y \geq 0$. Daher ist q eine Summe aus nicht negativen Zahlen, die nur dann 0 ergeben kann, wenn jeder einzelne Summand verschwindet. Insbesondere muß demnach $x^{n-1} = 0$ sein. Wäre $x \neq 0$, so könnten wir diese Gleichung auf beiden Seiten wiederholt durch x teilen, bis wir doch bei $x = 0$ angekommen wären. Also muß $x = 0$ gelten. Genauso können wir dann von $y^{n-1} = 0$ auf $y = 0$ schließen, also auf $x = y$, im Widerspruch zur Annahme. Für gerades n haben wir die Behauptung demnach gezeigt.

Jetzt sei n ungerade und $x \neq y$. Da $x^n = y^n$ gilt, können x und y nicht verschiedene Vorzeichen haben (bei geradem n war dieser Schluß nicht möglich, weshalb in den Voraussetzungen für diesen Fall eben $x, y \geq 0$ zu fordern war). Da $n - 1$ gerade ist, ist $x^{n-1} \geq 0$. $x^{n-2}y$ ist aber ebenfalls nicht negativ, denn, falls $x < 0$ sein sollte, gilt das auch für y , so daß $x^{n-2} < 0$ und $y < 0$ zu $x^{n-2}y > 0$ führt. Auf diese Weise können wir alle Summanden von q als nicht negativ nachweisen. Daher läßt sich, wie oben vorgeführt, auf $x = y = 0$ schließen, im Widerspruch zur Annahme.

Nun, da die Injektivität nachgewiesen ist, wissen wir nach Übung 2.5.8, daß es eine Umkehrfunktion $(\text{id}^n)^{-1} : \text{im } \text{id}^n \rightarrow D_{\text{id}^n}$ gibt, die durch Auflösen der Gleichung $x^n = y$ nach x bestimmt werden kann. Für gerades n ist $\text{im } \text{id}^n = \mathbb{R}_0^+$, für ungerades ist $\text{im } \text{id}^n = \mathbb{R}$. Das scheint offensichtlich zu sein, können wir aber im Moment noch nicht zeigen. Dafür fehlt uns noch der Stetigkeitsbegriff 11.1.1 und der Zwischenwertsatz 11.1.10. Die Umkehrfunktion ist die n -te Wurzel: $(\text{id}^n)^{-1}(y) =: \sqrt[n]{y}$. Näheres dazu in 11.1.12.

Wir könnten jetzt der Meinung sein, daß wir bei der Berechnung der Umkehrfunktion auf eine bekannte Funktion, nämlich die n -te Wurzel, gestoßen sind. Tatsächlich haben wir sie aber eben erst *definiert*, als Ergebnis der allgemeinen Überlegungen zur Existenz von Umkehrfunktionen.

2.5.10 Permutationen Eine bijektive Abbildung π der Menge $\mathbb{N}_n := \{1, \dots, n\}$ heißt *Permutation* der Zahlen $1, 2, \dots, n$. Wir schließen den trivialen Fall $n = 1$ aus und nehmen ab jetzt $n > 1$ an. Eine Funktion auf der endlichen Menge \mathbb{N}_n läßt sich vollständig durch eine Aufzählung ihrer Funktionswerte $[\pi(1), \pi(2), \dots, \pi(n)]$ wiedergeben. So ist durch $[2, 9, 3, 4, 1, 8, 6, 7, 5]$ die Permutation π mit den Funktionswerten $\pi(1) = 2, \pi(2) = 9, \pi(3) = 3, \dots, \pi(8) = 7$ und $\pi(9) = 5$ bestimmt. Die Inverse π^{-1} wird dann durch $[5, 1, 3, 4, 9, 7, 8, 6, 2]$ beschrieben. Wir identifizieren eine Permutation π mit der geordneten Aufzählung $[\pi(1), \pi(2), \dots, \pi(n)]$ ihrer Bilder. Für die angegebenen Beispiele schreiben wir also $\pi = [2, 9, 3, 4, 1, 8, 6, 7, 5]$ und $\pi^{-1} = [5, 1, 3, 4, 9, 7, 8, 6, 2]$. Wenn mit Permutationen gerechnet werden muß, ist diese Notation etwas knapp. Dann verwendet man gerne die ausführlichere Schreibweise $\pi = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 9 & 3 & 4 & 1 & 8 & 6 & 7 & 5 \end{bmatrix}$. So läßt sich etwa die Inverse leicht angeben, indem man dieses Schema von unten nach oben in aufsteigender Reihenfolge durchgeht: Man sucht die 1 und kann darüber ihr Urbild 5 ablesen. Damit weiß man $\pi^{-1}(1) = 5$, dann $\pi^{-1}(2) = 1$ usw. Genauso einfach läßt sich die Verkettung, also die Hintereinanderausführung zweier Permutationen bilden. Für $\sigma := \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 2 & 8 & 1 & 9 & 6 & 3 & 4 & 7 \end{bmatrix}$

ist $\sigma \circ \pi = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 7 & 8 & 1 & 5 & 4 & 6 & 3 & 9 \end{bmatrix}$. Aus diesen Darstellungen ist $1 \xrightarrow{\pi} 2 \xrightarrow{\sigma} 2$, oder $\sigma(\pi(1)) = \sigma(2) = 2$ leicht abzulesen, dann $2 \xrightarrow{\pi} 9 \xrightarrow{\sigma} 7$, also $\sigma(\pi(2)) = 7$ usw.

Die Menge \mathcal{S}_n aller Permutationen auf \mathbb{N}_n heißt *symmetrische Gruppe*. Mit der Hintereinanderausführung \circ als Verknüpfung zweier Permutationen hat sie nämlich die folgenden *Gruppeneigenschaften*: Für alle $\pi, \sigma, \tau \in \mathcal{S}_n$ gilt $\pi \circ \sigma \in \mathcal{S}_n$, $\pi \circ (\sigma \circ \tau) = (\pi \circ \sigma) \circ \tau$ (Assoziativität), die identische Abbildung id auf \mathbb{N}_n ist das neutrale Element der Verknüpfung, d. h., es gilt $\pi \circ \text{id} = \text{id} \circ \pi = \pi$ für alle $\pi \in \mathcal{S}_n$, und für jedes $\pi \in \mathcal{S}_n$ gibt es die Inverse $\pi^{-1} \in \mathcal{S}_n$, also eine Abbildung mit der Eigenschaft $\pi \circ \pi^{-1} = \pi^{-1} \circ \pi = \text{id}$.

Normalerweise lassen wir das Verknüpfungszeichen \circ weg, d. h., wir schreiben statt $\pi \circ \sigma$ einfach $\pi\sigma$. Auch sprechen wir lieber von dem Produkt $\pi\sigma$ der Permutationen π und σ , als von der Hintereinanderausführung.

Eine *Transposition* ist eine Permutation, bei der genau zwei Elemente vertauscht werden. Wird das Element i mit $j > i$ vertauscht, so bezeichnen wir die zugehörige Transposition mit t_{ij} . Für t_{i+1} schreiben wir einfach t_i . Es ist eine *Transposition nächster Nachbarn*, die durch den kleineren der beiden vollständig festgelegt ist. Jede Transposition t_{ij} läßt sich durch eine ungerade Anzahl, nämlich $2(j - i) - 1$ Transpositionen nächster Nachbarn gewinnen. Denn wir brauchen $j - i$ solcher Transpositionen, um i mit $i + 1$, dann mit $i + 2$ usw. und schließlich i mit j zu vertauschen. Anschließend benötigen wir noch einmal $j - i - 1$, um j mit $j - 1$, dann mit $j - 2$ usw. und endlich mit $i + 1$ zu vertauschen:

$$[i, i + 1, \dots, j - 1, j] \rightarrow [i + 1, i + 2, \dots, j - 1, j, i] \rightarrow [j, i + 1, \dots, j - 1, i].$$

Eine *Inversion* bei einer Permutation π ist ein Zahlenpaar $[\pi(i), \pi(i + k)]$, mit den Eigenschaften $k \geq 1$ und $\pi(i) > \pi(i + k)$. Das sind also Zahlen, die nach der Permutation nicht mehr in aufsteigender Reihenfolge vorliegen. Die Permutation $[3, 2, 1, 4, 5, 7, 9, 8, 6]$ etwa hat die Inversionen $[3, 2]$, $[3, 1]$, $[2, 1]$, $[7, 6]$, $[9, 8]$, $[9, 6]$ und $[8, 6]$. Man erhält systematisch alle Inversionen, indem

man die Inversionen mit dem ersten Element $\pi(1)$ bildet, wenn es welche gibt, dann die mit $\pi(2)$ als erstem Element und auf diese Weise bis zu $\pi(n-1)$ fortführt.

Jede Transposition t_i erhöht oder vermindert die Anzahl der Inversionen einer Permutation um 1, denn aus $[\pi(1), \dots, \pi(i-1), \pi(i), \pi(i+1), \pi(i+2), \dots, \pi(n)]$ wird $[\pi(1), \dots, \pi(i-1), \pi(i+1), \pi(i), \pi(i+2), \dots, \pi(n)]$. Dabei ändert sich die Anzahl der Inversionen in $[\pi(1), \dots, \pi(i-1), \pi(i+2), \dots, \pi(n)]$ nicht. Auch Inversionen mit den ersten Elementen an Positionen zwischen 1 und $i-1$ und den zweiten bei i oder $i+1$ bleiben gleich, da sich ihre Größenbeziehungen nicht ändern. Dasselbe gilt für die erste Position bei i oder $i+1$ und die zweite zwischen $i+2$ und n . Einzig die Größenbeziehung zwischen $\pi(i)$ und $\pi(i+1)$ ändert sich. War sie ursprünglich $\pi(i) < \pi(i+1)$, dann kommt eine Inversion hinzu, andernfalls verschwindet eine.

2.5.11 Satz Jede Permutation $\pi \in \mathcal{S}_n$ läßt sich als Produkt von Transpositionen darstellen. Diese Darstellung ist nicht eindeutig, aber sie besteht entweder immer aus einer geraden, oder immer aus einer ungeraden Anzahl von Transpositionen.

Beweis. Auf eine von id verschiedene Permutation $[\pi(1), \pi(2), \dots, \pi(n)]$ wenden wir die Transposition $\tau_1 := t_{\pi(k)k}$ an. Dabei ist k die größte Zahl mit $\pi(k) \neq k$, also die letzte Position, die nicht mit ihrem Eintrag übereinstimmt. Daher gilt $\pi(k) < k$, denn aus $\pi(k) = k + r$ mit einem $r > 0$ würde $k + r \neq \pi(k + r)$ folgen, im Widerspruch zur Maximalität von k . Als Ergebnis erhalten wir eine Permutation π_1 , die ab k die identische Permutation ist: $\pi_1(\ell) = \ell$ für $k \leq \ell \leq n$. Diesen Vorgang wiederholen wir für π_1 und erhalten eine Permutation π_2 , die jetzt ab einer Position $k_1 < k$ mit der identischen übereinstimmt. Nach einer endlichen Anzahl von Wiederholungen haben alle Zahlen ihre Ausgangsposition eingenommen. Wir sind daher bei der identischen Permutation angekommen: $\text{id} = \tau_\ell \tau_{\ell-1} \dots \tau_1 \pi$. Wegen $\tau_i \tau_i = \text{id}$ für $i = 1, \dots, \ell$, erhalten wir aus dieser Gleichung durch Multiplikation mit τ_ℓ , dann mit $\tau_{\ell-1}$ usw.: $\pi = \tau_1 \dots \tau_{\ell-1} \tau_\ell$. Damit ist $\pi \neq \text{id}$ ein Produkt aus Transpositionen. Für id läßt sich das natürlich ebenfalls erreichen, etwa durch $\text{id} = t_1 t_1$.

Wir nehmen an, π sei eine Permutation, für die es eine Darstellung $\pi = \tau_1 \dots \tau_{2k}$ mit einer geraden Anzahl von Transpositionen und eine Darstellung $\pi = \sigma_1 \dots \sigma_{2\ell+1}$ mit einer ungeraden Anzahl gibt. Das hieße, daß die identische Permutation ein Produkt aus einer ungeraden Anzahl von Transpositionen ist: $\text{id} = \sigma_{2\ell+1} \dots \sigma_1 \tau_1 \dots \tau_{2k}$. id wäre dann sogar das Produkt einer ungeraden Zahl von Transpositionen nächster Nachbarn, denn wir wissen, daß das für jede der beteiligten Transpositionen σ_i und τ_i gilt. Natürlich ist die Anzahl der Inversionen von id gleich Null. Andererseits erzeugt oder vernichtet jede Transposition nächster Nachbarn eine Inversion. Zählen wir also die durch die Transpositionen erzeugten Inversionen, so erhalten wir eine Summe mit einer ungeraden Anzahl Summanden, die nur 1 oder -1 sind. Als Ergebnis müßte sie 0 ergeben, was nicht möglich ist. \square

Es ist ganz instruktiv, das Verfahren aus dem Beweis einmal an einem konkreten Beispiel durchzuführen. Wir wählen $\pi := \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 3 & 7 & 2 & 5 & 1 & 8 \end{bmatrix}$ und starten mit $k = 7$ und $\pi(7) = 1$. Das führt auf $\pi_1 := t_{1,7}\pi = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 3 & 1 & 2 & 5 & 7 & 8 \end{bmatrix}$. Jetzt sind die Positionen ab $k = 7$ wieder in ihrer ursprünglichen Reihenfolge. Nun ist $k = 6$ und $\pi_1(6) = 5$.

Daher ist $\pi_2 := t_{5,6}t_{1,7}\pi = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 3 & 1 & 2 & 6 & 7 & 8 \end{bmatrix}$. Als nächste Permutation erhalten wir $\pi_3 := t_{2,5}t_{5,6}t_{1,7}\pi = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 2 & 3 & 1 & 5 & 6 & 7 & 8 \end{bmatrix}$ und schließlich $t_{1,4}t_{2,5}t_{5,6}t_{1,7}\pi = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = \text{id}$. Damit ist π eine gerade Permutation, die durch $\pi = t_{1,7}t_{5,6}t_{2,5}t_{1,4}$ aus Transpositionen aufgebaut wird.

2.5.12 Definition Eine Permutation $\pi \in \mathcal{S}_n$ heißt gerade, wenn sie das Produkt einer geraden Zahl von Transpositionen ist, sonst ungerade. Das Vorzeichen $\text{sgn}(\pi)$ einer Permutation ist 1, falls π gerade und -1 , falls π ungerade ist.

Satz 2.5.11 zeigt, daß $\text{sgn}(\pi)$ wohldefiniert ist. Darüber hinaus ist klar, daß $\text{sgn}(\pi) = (-1)^{t(\pi)}$ gilt, wenn $t(\pi)$ die Anzahl der Transpositionen einer Darstellung von π als Produkt von Transpositionen ist. Diese ist zwar nicht eindeutig, aber zwei verschiedene Versionen müssen sich immer um eine gerade Anzahl von Transpositionen unterscheiden.

Die zentrale Eigenschaft

$$\text{sgn}(\pi\sigma) = \text{sgn}(\pi) \text{sgn}(\sigma) \quad (2.16)$$

läßt sich jetzt leicht einsehen. Sind π und σ gerade, so ist auch $\pi\sigma$ gerade, denn die gerade Anzahl von Transpositionen für π und die gerade Anzahl für σ ergibt hintereinander ausgeführt eine gerade Anzahl für $\pi\sigma$. Hier stimmt (2.16) offensichtlich: $\text{sgn}(\pi) \text{sgn}(\sigma) = 1 \cdot 1 = \text{sgn}(\pi\sigma)$. Ist eine der Permutationen π oder σ gerade, die andere aber ungerade, so läßt sich $\pi\sigma$ durch eine ungerade Anzahl von Transpositionen darstellen. Es folgt $\text{sgn}(\pi) \text{sgn}(\sigma) = 1 \cdot (-1) = \text{sgn}(\pi\sigma)$. Schließlich ist $\pi\sigma$ gerade, wenn π und σ ungerade sind, so daß (2.16) auch in diesem Fall stimmt.

$\text{sgn} : \mathcal{S}_n \rightarrow \{-1, 1\}$ ist ein sogenannter *Gruppenhomomorphismus* zwischen den beiden Gruppen \mathcal{S}_n und der multiplikativen Gruppe $\{-1, 1\}$ (mit den gewohnten Rechenregeln der Multiplikation, aus denen natürlich $1^{-1} = 1$ und $(-1)^{-1} = -1$ folgt). Die Homomorphismus-Eigenschaft wird durch (2.16) wiedergegeben. Er hat die Eigenschaft $\text{sgn}(t_i) = -1$, $i = 1, \dots, n-1$, durch die er bereits eindeutig festgelegt ist. Denn jeder Homomorphismus λ mit dieser Eigenschaft erfüllt $\lambda(\pi) = \lambda(t_{i_1}t_{i_2} \cdots t_{i_k}) = \lambda(t_{i_1})\lambda(t_{i_2}) \cdots \lambda(t_{i_k}) = (-1)^k = (-1)^{t(\pi)} = \text{sgn}(\pi)$, für jede Zerlegung von π in Transpositionen t_{i_ℓ} nächster Nachbarn.

Es gibt eine Formel für $\text{sgn}(\pi)$. Dafür definieren wir

$$\begin{aligned} \Delta &:= \prod_{k=1}^{n-1} \prod_{\ell > k} (\ell - k) \\ &= (2-1)(3-1)(4-1)(5-1)(6-1) \cdots (n-1) \cdot \\ &\quad (3-2)(4-2)(5-2)(6-1) \cdots (n-2) \cdot \\ &\quad (4-3)(5-3)(6-3) \cdots (n-3) \cdot \\ &\quad \cdots \\ &\quad (n - (n-1)). \end{aligned}$$

In diesem Produkt bestehen die Faktoren aus allen möglichen positiven Differenzen der Zahlen aus \mathbb{N}_n in ihrer natürlichen Reihenfolge. Jetzt bilden wir

$$\Delta(\pi) := \prod_{k=1}^{n-1} \prod_{\ell > k} (\pi(\ell) - \pi(k)).$$

Auch hier kommen alle möglichen Differenzen vor, nur sind die Faktoren in einer anderen Reihenfolge. Außerdem sind nicht mehr alle Differenzen positiv. $\pi(\ell) - \pi(k)$ ist genau dann negativ, wenn $[\pi(\ell), \pi(k)]$ eine Inversion von π ist. Die Anzahl der negativen Faktoren in $\Delta(\pi)$ ist demnach die Anzahl $i(\pi)$ der Inversionen von π . Daher ist

$$\frac{\Delta(\pi)}{\Delta} = (-1)^{i(\pi)}.$$

Wir nennen diesen Ausdruck vorläufig $\lambda(\pi)$. Offensichtlich ist für jede Transposition t_i nächster Nachbarn $\lambda(t_i) = -1$, denn sie hat genau eine Inversion. Wenn wir jetzt noch die Homomorphismus-Eigenschaft $\lambda(\pi\sigma) = \lambda(\pi)\lambda(\sigma)$ zeigen können, wissen wir $\lambda = \text{sgn}$.

$$\Delta(\pi\sigma) = \prod_{k=1}^{n-1} \prod_{\ell > k} (\pi\sigma(\ell) - \pi\sigma(k)) = \frac{\prod_{k=1}^{n-1} \prod_{\ell > k} (\pi(\sigma(\ell)) - \pi(\sigma(k)))}{\prod_{k=1}^{n-1} \prod_{\ell > k} (\sigma(\ell) - \sigma(k))} \cdot \Delta(\sigma)$$

Der Nenner des Bruchs ist nicht Δ . Da die Differenzen $\sigma(\ell) - \sigma(k)$ im Zähler und im Nenner aber in genau derselben Form auftreten, kann man den Nenner durch Umkehrung negativer Faktoren in Δ verwandeln. Dabei ändert sich der Bruch nicht, wenn man das mit dem Argumenten von π im Zähler genauso macht:

$$\Delta(\pi\sigma) = \frac{\prod_{k=1}^{n-1} \prod_{\ell > k} (\pi(\ell) - \pi(k))}{\prod_{k=1}^{n-1} \prod_{\ell > k} (\ell - k)} \cdot \Delta(\sigma) = \frac{\Delta(\pi)}{\Delta} \cdot \Delta(\sigma) = \lambda(\pi) \cdot \Delta(\sigma).$$

Teilt man beide Seiten durch Δ , so gelangt man zu $\lambda(\pi\sigma) = \lambda(\pi)\lambda(\sigma)$. Damit haben wir

$$\text{sgn}(\pi) = \frac{\Delta(\pi)}{\Delta} = (-1)^{i(\pi)} \quad (2.17)$$

abgeleitet. Das Vorzeichen einer Permutation ist also auch durch die Anzahl der Inversionen bestimmt. Insbesondere hat eine gerade bzw. ungerade Permutation auch immer eine gerade bzw. ungerade Anzahl von Inversionen.

3 Zahlentheorie

3.1 Teilbarkeitstheorie ganzer Zahlen

3.1.1 Definition Eine Zahl $0 \neq a \in \mathbb{Z}$ heißt Teiler einer Zahl $b \in \mathbb{Z}$, wenn es eine Zahl $c \in \mathbb{Z}$ mit der Eigenschaft

$$b = c \cdot a$$

gibt. Wir schreiben dafür $a \mid b$.

Für $b = 0$ gibt es natürlich unendlich viele Teiler, denn für $c = 0$ und alle $a \in \mathbb{Z}$ gilt $0 = 0 \cdot a$. Für $b \neq 0$ gibt es bei gegebenem Teiler a nur ein Element $c \in \mathbb{Z}$, mit $b = ca$, denn für ein weiteres $c' \in \mathbb{Z}$ mit $b = c'a$ würde $ca - c'a = (c - c')a = 0$ folgen. Da $b \neq 0$ ist, muß auch $a \neq 0$ gelten, so daß nur $c = c'$ bleibt.

3.1.2 Satz (Teilen mit Rest) Für jede ganze Zahl p und jede natürliche Zahl q gibt es eindeutig bestimmte ganze Zahlen $r, t \in \mathbb{Z}$ mit der Eigenschaft

$$p = tq + r, \quad 0 \leq r < q. \quad (3.1)$$

Ein $r > 0$ heißt Rest der Teilung von p durch q .

Beweis. Wir behandeln zunächst die Eindeutigkeit. Aus $0 \leq r' < q$ und

$$p = tq + r = t'q + r'$$

folgt, wenn wir o. B. d. A. $r' \geq r$ annehmen, $(t - t')q = r' - r \geq 0$. Damit ist $t - t' \geq 0$. Wäre $t - t' > 0$, so würde $(t - t')q = r' - r \geq q$ folgen, im Widerspruch dazu, daß sowohl r' als auch r echt kleiner als q sind, was für die Differenz $r' - r$ erst recht zutrifft. Also muß $t - t' = 0$ und damit auch $r' - r = 0$ gelten.

Zur Existenz: Dazu betrachten wir die Menge

$$M := \{ \ell \in \mathbb{Z} \mid \ell q \leq p \}.$$

In M muß es ein größtes Element ℓ_0 geben, denn andernfalls würde ℓq über alle Grenzen wachsen und dabei schließlich größer als p werden. Wir setzen $r := p - \ell_0 q (\geq 0)$ und haben damit $p = \ell_0 q + r$ mit einem $r \geq 0$ erreicht. Es muß noch $r < q$ nachgewiesen werden, um die Forderungen des Satzes zu erfüllen. Wir nehmen das Gegenteil $r \geq q$ an, d. h. $r = q + d$, $d \geq 0$. Das bedeutet

$$d = p - (\ell_0 + 1)q \geq 0 \quad \text{oder} \quad (\ell_0 + 1)q \leq p.$$

Also ist auch $\ell_0 + 1 \in M$, im Widerspruch dazu, daß ℓ_0 das größte Element von M ist. $r \geq q$ ist demnach nicht möglich, es muß $r < q$ gelten. $t := \ell_0$ und r sind die Zahlen, von denen im Satz die Rede ist. \square

Beispiel: Für $p = 82$ und $q = 7$, bzw. $p = -82$ und $q = 7$, bzw. $p = 82$ und $q = 83$ gilt:

$$82 = 11 \cdot 7 + 5, \quad -82 = -12 \cdot 7 + 2, \quad 82 = 0 \cdot 83 + 82.$$

3.1.3 Definition Der größte gemeinsame Teiler ganzer Zahlen $p_1 \neq 0, p_2 \neq 0, \dots, p_n \neq 0$ ist eine natürliche Zahl t mit den Eigenschaften

$$t \mid p_1 \wedge t \mid p_2 \wedge \dots \wedge t \mid p_n, \quad (3.2)$$

$$\forall_{s \in \mathbb{N}} s \mid p_1 \wedge s \mid p_2 \wedge \dots \wedge s \mid p_n \Rightarrow s \leq t. \quad (3.3)$$

Für den größten gemeinsamen Teiler von p_1, \dots, p_n schreiben wir $\text{ggT}(p_1, \dots, p_n)$. Zahlen p_1, \dots, p_n mit größtem gemeinsamen Teiler 1 heißen teilerfremd.

Den größten gemeinsamen Teiler von 84 und 18 findet man durch wiederholtes Teilen mit Rest:

$$84 = 4 \cdot 18 + 12$$

$$18 = 1 \cdot 12 + 6$$

$$12 = 2 \cdot 6.$$

$$\text{ggT}(84, 18) = \text{ggT}(84, -18) = \text{ggT}(-84, 18) = \text{ggT}(-84, -18) = 6.$$

84											
18			18			18			18		
6	12		6	12		6	12		6	12	
6	6	6	6	6	6	6	6	6	6	6	6

3.1.4 Satz (Euklidischer Algorithmus) Für zwei natürliche Zahlen p_0 und p_1 , $p_0 \geq p_1$, ergibt sich der größte gemeinsame Teiler durch wiederholtes Teilen mit Rest:

$$p_0 = t_1 \cdot p_1 + p_2, \quad 0 \leq p_2 < p_1$$

$$p_1 = t_2 \cdot p_2 + p_3, \quad 0 \leq p_3 < p_2$$

$$p_2 = t_3 \cdot p_3 + p_4, \quad 0 \leq p_4 < p_3$$

$$\vdots$$

$$p_{n-2} = t_{n-1} \cdot p_{n-1} + p_n, \quad 0 \leq p_n < p_{n-1}$$

$$p_{n-1} = t_n \cdot p_n.$$

p_n ist der größte gemeinsame Teiler von p_0 und p_1 .

Beweis. Da die Reste p_2, p_3, \dots in jedem Schritt kleiner werden, muß das Verfahren schließlich enden. Die letzte Zeile besagt, daß p_n jedenfalls ein Teiler von p_{n-1} ist. Aus der vorletzten folgt dann, daß p_n auch p_{n-2} teilt. Auf diese Weise fortfahrend gelangt man zur zweiten Zeile und folgert, daß p_n auch p_1 teilt. $p_n \mid p_0$ folgt schließlich aus der ersten Zeile. Damit ist p_n ein gemeinsamer Teiler von p_0 und p_1 .

Sei $0 < q$ ein weiterer Teiler von p_0 und p_1 . Aus der ersten Zeile ergibt sich dann sofort, daß q den Rest p_2 teilt, aus der zweiten, daß auch p_3 von q geteilt wird usw., bis man bei der letzten Zeile angekommen ist und $q \mid p_n$ folgern kann. Das bedeutet insbesondere $q \leq p_n$. Damit ist p_n der größte gemeinsame Teiler von p_0 und p_1 . \square

3.1.5 Lemma (EUKLID) Für den größten gemeinsamen Teiler (p, q) zweier natürlicher Zahlen p und q gibt es immer eine Darstellung der Form

$$\text{ggT}(p, q) = x p + y q, \quad x, y \in \mathbb{Z}. \quad (3.4)$$

Es gilt

$$\text{ggT}(p, q) = \min \{ x p + y q \mid x, y \in \mathbb{Z} \wedge x p + y q \geq 1 \}. \quad (3.5)$$

Die Darstellung (3.4) ist keineswegs eindeutig.

Beweis. Wir zeigen (3.5). Dafür sei $r := \min \{ x p + y q \mid x, y \in \mathbb{Z}, x p + y q \geq 1 \} = x_0 p + y_0 q$ für geeignete $x_0, y_0 \in \mathbb{Z}$ und $r_0 := \text{ggT}(p, q)$. Dann ist $r_0 = r$ zu zeigen.

Als gemeinsamer Teiler von p und q teilt r_0 natürlich auch $r = x_0 p + y_0 q$. Das bedeutet $r_0 \leq r$.

Nun muß die umgekehrte Abschätzung $r \leq r_0$ gezeigt werden. Diese ist sicher erfüllt, wenn r ein gemeinsamer Teiler von p und q ist (denn r_0 ist der größte dieser Teiler). Wir gehen vom Gegenteil aus und nehmen o. B. d. A. an, daß p nicht von r geteilt wird. Nach Satz 3.1.2 gibt es eine Zahl $0 < t < r$ mit der Eigenschaft

$$p = \ell r + t,$$

für ein geeignetes $\ell \in \mathbb{Z}$. Setzen wir darin $r = x_0 p + y_0 q$,

$$p = \ell(x_0 p + y_0 q) + t$$

und lösen t nach auf:

$$0 < t = (1 - \ell x_0)p - \ell y_0 q < r.$$

Mit t haben wir ein Element der Menge $\{ x p + y q \mid x, y \in \mathbb{Z}, x p + y q > 0 \}$ gefunden, das kleiner als das Minimum r dieser Menge ist. Dieser offensichtliche Widerspruch zeigt, daß unsere Annahme falsch sein muß. Also ist r ein gemeinsamer Teiler von p und q und insbesondere $\leq r_0$. Das zeigt $r = r_0$ und damit die behauptete Darstellung $r = \text{ggT}(p, q)$.

Daß die Darstellung $\text{ggT}(p, q) = x_0 p + y_0 q$ nicht eindeutig ist, sieht man am einfachsten an einem Zahlenbeispiel:

$$\text{ggT}(5, 7) = 1 = 3 \cdot 7 - 4 \cdot 5 = -7 \cdot 7 + 10 \cdot 5. \quad \square$$

Übrigens erhält man alle Darstellungen systematisch auf folgende Weise:

$$1 = 3 \cdot 7 - 4 \cdot 5 + t \cdot 5 \cdot 7 - t \cdot 7 \cdot 5 = (3 + 5t) \cdot 7 - (4 + 7t) \cdot 5, \quad t \in \mathbb{Z}.$$

Lemma 3.1.5 gibt keine Vorschrift an, wie die Darstellung (3.5) zu finden ist. Wir stellen das Verfahren dafür an zwei ausführlichen Beispielen vor.

3.1.6 Beispiel $\text{ggT}(1314, 396) = 18$:

$$\begin{array}{ll} 1314 = 3 \cdot 396 + 126 & \Rightarrow 126 = 1314 - 3 \cdot 396 \\ 396 = 3 \cdot 126 + 18 & \Rightarrow 18 = 396 - 3 \cdot 126 \\ 126 = 7 \cdot 18. & \end{array}$$

Die Gleichungen der letzten Spalte werden nun von unten nach oben ineinander eingesetzt und zusammengefaßt:

$$\begin{aligned} 18 &= 396 - 3 \cdot (1314 - 3 \cdot 396) = 396 - 3 \cdot 1314 + 9 \cdot 396 \\ &= 10 \cdot 396 - 3 \cdot 1314. \end{aligned}$$

3.1.7 Beispiel $\text{ggT}(5610, 637) = 1$:

$$\begin{array}{ll} 5610 = 8 \cdot 637 + 514 & \Rightarrow 514 = 5610 - 8 \cdot 637 \\ 637 = 1 \cdot 514 + 123 & \Rightarrow 123 = 637 - 514 \\ 514 = 4 \cdot 123 + 22 & \Rightarrow 22 = 514 - 4 \cdot 123 \\ 123 = 5 \cdot 22 + 13 & \Rightarrow 13 = 123 - 5 \cdot 22 \\ 22 = 1 \cdot 13 + 9 & \Rightarrow 9 = 22 - 13 \\ 13 = 1 \cdot 9 + 4 & \Rightarrow 4 = 13 - 9 \\ 9 = 2 \cdot 4 + 1 & \Rightarrow 1 = 9 - 2 \cdot 4 \\ 4 = 4 \cdot 1. & \end{array}$$

Wir setzen die Gleichungen der letzten Spalte, mit der letzten beginnend, nacheinander und nach oben fortschreitend ineinander ein. Dabei werden die Zahlen 4, 9, 13, ..., 514, 637 und 5610 wie Variablen behandelt, die sukzessive durch die nachfolgenden ersetzt werden, bis zuletzt nur noch die letzten beiden, nämlich 637 und 5610, vorhanden sind.

$$\begin{aligned} 1 &= 9 - 2 \cdot (13 - 9) &= 3 \cdot 9 - 2 \cdot 13 \\ &= 3 \cdot (22 - 13) - 2 \cdot 13 &= 3 \cdot 22 - 5 \cdot 13 \\ &= 3 \cdot 22 - 5 \cdot (123 - 5 \cdot 22) &= 28 \cdot 22 - 5 \cdot 123 \\ &= 28 \cdot (514 - 4 \cdot 123) - 5 \cdot 123 &= 28 \cdot 514 - 117 \cdot 123 \\ &= 28 \cdot 514 - 117 \cdot (637 - 514) &= 145 \cdot 514 - 117 \cdot 637 \\ &= 145 \cdot (5610 - 8 \cdot 637) - 117 \cdot 637 &= 145 \cdot 5610 - 1277 \cdot 637. \end{aligned}$$

Die gesuchte Darstellung nach EUKLID ist demnach

$$1 = 145 \cdot 5610 - 1277 \cdot 637.$$

3.1.8 Definition Zusammen mit dem euklidischen Algorithmus bezeichnet man das in den Beispielen vorgestellte Verfahren als erweiterten euklidischen Algorithmus.

3.1.9 Korollar Aus $p \mid ab$ und $\text{ggT}(p, a) = 1$ folgt $p \mid b$. Andererseits ergibt sich aus $a \mid p$, $b \mid p$ und $\text{ggT}(a, b) = 1$ auch $ab \mid p$.

Beweis. $\text{ggT}(p, a) = 1$ bedeutet nach Lemma 3.1.5 $1 = xp + ya$ für geeignete $x, y \in \mathbb{Z}$. Das multiplizieren wir mit b :

$$b = bxp + yab.$$

Da ab durch p teilbar ist und bxp offensichtlich auch, ist b durch p teilbar.

Die zweite Behauptung: $a \mid p$ bedeutet $p = ta$, mit einem geeigneten $t \in \mathbb{Z}$. Wir haben daher $b \mid ta$ und $\text{ggT}(b, a) = 1$, woraus $b \mid t$ und schließlich $ab \mid p$ folgt. \square

3.1.10 Korollar Für eine Primzahl p folgt aus $p \mid ab$:

$$p \mid a \text{ oder } p \mid b.$$

Beweis. Da p eine Primzahl ist, kann ein gemeinsamer Teiler von p und a nur 1 oder p sein. Es gibt daher die folgenden beiden Fälle:

1. $\text{ggT}(p, a) = p$ bedeutet, daß p ein Teiler von a ist.
2. $\text{ggT}(p, a) = 1$ bedeutet nach Korollar 3.1.9 $p \mid b$. \square

3.1.11 Korollar Für eine Primzahl p ist \sqrt{p} niemals rational.

Beweis. Wir nehmen das Gegenteil $\sqrt{p} \in \mathbb{Q}$ an, d. h. $\sqrt{p} = \frac{a}{b}$ mit geeigneten Zahlen $a, b \in \mathbb{N}$. O. B. d. A. können wir davon ausgehen, daß der Bruch $\frac{a}{b}$ vollständig gekürzt ist, d. h., daß $\text{ggT}(a, b) = 1$ gilt. Wir quadrieren $\sqrt{p} = \frac{a}{b}$ und stellen um:

$$pb^2 = a^2.$$

Das bedeutet $p \mid a^2$, nach Korollar 3.1.10 also $p \mid a$. Daher gibt es ein $t \in \mathbb{N}$, so daß $a = tp$ gilt. Eingesetzt in obige Gleichung:

$$pb^2 = t^2p^2 \quad \Rightarrow \quad b^2 = t^2p.$$

Wie gerade eben schließen wir daraus $p \mid b$. Damit ist p ein gemeinsamer Teiler von a und b , im Widerspruch dazu, daß $\text{ggT}(a, b) = 1$ gelten sollte. \square

3.2 Rechnen Modulo p

3.2.1 Definition Für eine Zahl $p \in \mathbb{N}$ führen wir eine neue Gleichheit zweier ganzer Zahlen x und y ein. Wir sagen

$$x =_p y \quad :\Leftrightarrow \quad p \mid x - y. \quad (3.6)$$

Dafür ist die Schreibweise

$$x = y \pmod{p}, \quad (3.7)$$

der allgemeine Standard. Gesprochen: x gleich y modulo p .

Wir werden beide Notationen verwenden.

Zwei Zahlen x und y sind genau dann gleich modulo p , wenn sie sich um ein Vielfaches von p unterscheiden:

$$x =_p y \quad \Leftrightarrow \quad \exists t \in \mathbb{Z} : \quad x = y + tp.$$

3.2.2 Beispiel $5 =_2 7$, oder $5 = 7 \pmod{2}$. $-3 =_{11} 8$, oder $-3 = 8 \pmod{11}$.

Diese Gleichheit modulo p reduziert die Anzahl der verschiedenen Elemente in \mathbb{Z} drastisch, denn es werden unendlich viele ganze Zahlen als gleich angesehen (modulo p):

$$\begin{array}{ccccccc} 0 =_p & \pm p =_p & \pm 2p =_p & \pm 3p =_p & \dots \\ 1 =_p & 1 \pm p =_p & 1 \pm 2p =_p & 1 \pm 3p =_p & \dots \\ \vdots & \vdots & \vdots & \vdots & \\ p-1 =_p & p-1 \pm p =_p & p-1 \pm 2p =_p & p-1 \pm 3p =_p & \dots \end{array}$$

d. h., es gibt genau p dieser Mengen ganzer Zahlen, die modulo p nicht unterschieden werden. Offensichtlich ist jede dieser Mengen durch eine der Zahlen x zwischen 0 und $p-1$ bereits eindeutig festgelegt. Man schreibt für sie

$$[x] := \{ y \in \mathbb{Z} \mid x = y \pmod{p} \}.$$

Man kann sich leicht davon überzeugen, daß es sich bei $=_p$ um eine Äquivalenzrelation handelt und daß $[x]$ die zugehörigen Äquivalenzklassen sind. Wir werden diese Abstraktion nicht benötigen, da für unsere Zwecke die Identifikation modulo p als operativer Vorgang völlig ausreicht (wer darüber mehr wissen will, findet etwas in Abschnitt 2.3). Das bedeutet, daß wir uns vollständig innerhalb der Menge

$$\mathbb{Z}_p = \{0, 1, 2, \dots, p-1\} \quad (3.8)$$

bewegen können, wenn wir Berechnungen modulo p anstellen, auch wenn das nicht sehr praktisch ist. Tatsächlich erlauben wir es, daß in einzelnen Rechenschritten Ergebnisse durchaus außerhalb von \mathbb{Z}_p liegen dürfen, wenn wir nur das Endergebnis durch wiederholtes Teilen durch p mit Rest wieder in \mathbb{Z}_p darstellen.

Wir werden nach Satz 3.2.5 erkennen, daß sich die Multiplikation in

$$\mathbb{Z}_p^* := \{ k \in \mathbb{N} \mid k < p \wedge \text{ggT}(k, p) = 1 \} \cup \{0\} \quad (3.9)$$

besonders gut verhält, weil es zu jedem Element $0 \neq a \in \mathbb{Z}_p^*$ genau eine Element $0 \neq b \in \mathbb{Z}_p^*$ mit der Eigenschaft $ab =_p 1$ gibt. Für $p \in \mathbb{P}$ stimmt \mathbb{Z}_p mit \mathbb{Z}_p^* überein, für $p \notin \mathbb{P}$ aber nicht: $\mathbb{Z}_7^* = \{0, 1, 2, 3, 4, 5, 6\} = \mathbb{Z}_7$, aber $\mathbb{Z}_{12}^* = \{0, 1, 5, 7, 11\} \neq \mathbb{Z}_{12}$.

Bevor wir uns mit Berechnungen modulo p beschäftigen, müssen wir uns versichern, daß sie mit der neuen Gleichheit modulo p verträglich sind. Kurz gesagt vertragen sich, bis auf die Division, die ja schon in \mathbb{Z} nur eingeschränkt zur Verfügung steht, alle Grundrechenarten mit der Gleichheit modulo p. Um das einzusehen, müssen wir uns davon überzeugen, daß für modulo p gleiche Zahlen x_1 und x_2 bzw. y_1 und y_2 , das Ergebnis einer Rechnung nicht davon abhängt, ob wir sie mit x_1 und y_1 , oder mit x_2 und y_2 durchführen – vorausgesetzt, wir vergleichen diese Ergebnisse modulo p:

Aus $x_1 =_p x_2$ bzw. $y_1 =_p y_2$ folgt, daß es ganze Zahlen s und t mit $x_2 = x_1 + sp$ und $y_2 = y_1 + tp$ gibt. Dann haben wir:

$$\begin{aligned} x_2 \pm y_2 &= x_1 \pm y_1 + (s \pm t)p =_p x_1 \pm y_1, \\ x_2 y_2 &= (x_1 + sp)(y_1 + tp) = x_1 y_1 + (x_1 t + s y_2 + stp)p =_p x_1 y_1. \end{aligned}$$

Aus der letzten Gleichung ergibt sich auch $x_1^n =_p x_2^n$ für jedes $n \in \mathbb{N}$.

Als praktische Konsequenz erhalten wir: In jedem zulässigen Rechenschritt \pm und \cdot kann jedes Zwischenergebnis durch ein modulo p gleiches ersetzt werden. Das ermöglicht es, mit sehr großen Zahlen Berechnungen anzustellen, ohne tatsächlich sehr große Zahlen zu benutzen. Ein beliebtes Beispiel in diesem Zusammenhang ist folgende kleine Aufgabe:

3.2.3 Beispiel Wie lauten die letzten beiden Ziffern der Zahl 7^{1000} ?

Offensichtlich ist es nicht so ohne Weiteres möglich, diese Zahl einfach auszurechnen, um die letzten beiden Ziffern abzulesen. Erfreulicherweise ist das auch gar nicht nötig, denn die letzten beiden Ziffern von 7^{1000} können durch $7^{1000} \bmod 100$ einfach berechnet werden. Das bedeutet, daß wir nur mit Zahlen in der Größenordnung von 1000 arbeiten müssen:

$$7^{1000} = 7^{2 \cdot 500} = (7^2)^{500} = 49^{500} = (49^2)^{250} = 2401^{250} =_{100} 1^{250} = 1.$$

Also sind die letzten beiden Ziffern von 7^{1000} die 0 und die 1.

3.2.4 Beispiel (3-er, 9-er und 11-er Probe) Es gibt einen einfachen Test, mit dem eine ganze Zahl auf ihre Teilbarkeit durch 3 bzw. durch 9 geprüft werden kann: *Eine Zahl ist genau dann durch 3 bzw. 9 teilbar, wenn ihre Quersumme, also die Summe ihrer Ziffern, durch 3 bzw. 9 teilbar ist.* Das ist die sogenannte *3-er Probe* bzw. *9-er Probe*. Das liegt daran, daß $10 =_3 1$, $100 = 10^2 =_3 1^2 = 1$, ..., $10^n =_3 1^n = 1$, bzw. $10^n =_9 1$ gilt. Eine Zahl $a := a_n 10^n + a_{n-1} 10^{n-1} + \dots + a_2 10^2 + a_1 10 + a_0$ ist genau dann durch 3 teilbar, wenn $a =_3 0$ erfüllt ist. Da sich die Gleichheit modulo 3 mit der Addition und der Multiplikation verträgt, bedeutet das

$$a =_3 a_n \cdot 1 + a_{n-1} \cdot 1 + \dots + a_1 \cdot 1 + a_0 =_3 0 \quad \Leftrightarrow \quad 3 \mid a_n + a_{n-1} + \dots + a_1 + a_0.$$

Genauso ergibt sich auch die 9-er Probe. Die 11-er Probe:

Es gilt $1 \equiv_{11} 1$, $10 \equiv_{11} -1$, $100 \equiv_{11} 10^2 \equiv_{11} (-1)^2 \equiv_{11} 1$, $1000 \equiv_{11} 10^3 \equiv_{11} (-1)^3 \equiv_{11} -1$, ..., $10^n \equiv_{11} (-1)^n$. Damit folgt

$$a = \sum_{k=0}^n a_k \cdot 10^k \equiv_{11} \sum_{k=0}^n (-1)^k a_k \equiv_{11} 0 \Leftrightarrow 11 \mid \sum_{k=0}^n (-1)^k a_k.$$

Die Summe $\sum_{k=0}^n (-1)^k a_k$ wird *alternierende Quersumme* genannt.

Die Zahl 3 597 678 354 hat die Quersumme 57 und ist daher durch 3 teilbar (denn 57 hat die Quersumme 12, die offensichtlich durch 3 geteilt werden kann), nicht aber durch 9. Überprüfen wir das: $3\,597\,678\,354 = 1\,999\,226\,118 \cdot 3$ und $3\,597\,678\,354 = 399\,742\,039 \cdot 9 + 3$.

Da die alternierende Quersumme -5 ist, kann die Zahl auch nicht durch 11 geteilt werden. Tatsächlich gilt $3\,597\,678\,354 = 327\,061\,668 \cdot 11 + 6$.

53 978 173 535 hat die alternierende Quersumme $5 - 3 + 5 - 3 + 7 - 1 + 8 - 7 + 9 - 3 + 5 = 22$. Wir haben $53\,978\,173\,535 = 4\,907\,106\,685 \cdot 11$, was die 11-er Probe richtig voraussagt.

3.2.5 Satz (Inverse modulo p) Für eine natürliche Zahl $p > 1$ und für eine Zahl $0 < a \in \mathbb{Z}_p^*$, also für $0 < a < p$ und $\text{ggT}(a, p) = 1$ (vergl. (3.9)), gibt es genau eine Zahl $0 < b \in \mathbb{Z}_p^*$ mit der Eigenschaft

$$ab \equiv 1 \pmod{p}. \quad (3.10)$$

Die Zahl b wird als Inverse von a modulo p bezeichnet.

Gelegentlich schreiben wir dafür auch $a^{-1} \pmod{p}$.

Anders als in \mathbb{Z} , wo z. B. die Zahl 3 keine Inverse hat, denn wir finden keine ganze Zahl b mit der Eigenschaft $b \cdot 3 = 1$, gibt es eine solche Zahl, sagen wir modulo 7, da $5 \cdot 3 = 2 \cdot 7 + 1 \equiv_7 1$ gilt. Die Inverse modulo 7 von 3 ist daher 5.

Beweis. $0 < a \in \mathbb{Z}_p^*$ bedeutet $a < p$ und $\text{ggT}(a, p) = 1$. Nach dem Lemma von EUKLID 3.1.5 gibt es Zahlen $x, y \in \mathbb{Z}$, so daß

$$1 = xa + yp \equiv_p xa \quad (*)$$

gilt. Daher ist x ein Kandidat für b . $x \equiv_p 0$ kann nicht gelten, denn sonst müßte 1 durch p teilbar sein. Ist x ein Element von \mathbb{Z}_p^* , so ist es das gesuchte b . Sollte x nicht in \mathbb{Z}_p^* liegen, so ergibt Teilen durch p mit Rest (Satz 3.1.2):

$$x = tp + r \quad (**)$$

für ein $t \in \mathbb{Z}$ und $0 < r < p$ ($r = 0$ ist nicht möglich, denn das würde $p \mid x$, also $x \equiv_p 0$ bedeuten, s. o.). Es folgt

$$1 \equiv_p xa = atp + ra \equiv_p ra.$$

Für b wählen wir jetzt r , denn $r \in \mathbb{Z}_p^*$. Dafür ist nur noch $\text{ggT}(r, p) = 1$ nachzuweisen. Sei also $u \in \mathbb{N}$ ein gemeinsamer Teiler von r und p . Dann zeigt (**), daß u ein Teiler von x und (*), daß u auch ein Teiler von 1 ist. Das ist nur für $u = 1$ möglich. Daher sind r und p tatsächlich teilerfremd.

Die Eindeutigkeit modulo p: Für ein weiteres Element $b' \neq b$ mit den geforderten Eigenschaften $0 \neq b' \in \mathbb{Z}_p^*$ und $b'a \equiv_p 1$ folgt

$$(b - b')a \equiv_p 0,$$

d. h. $p \mid (b - b')a$. Wegen $\text{ggT}(a, p) = 1$ zeigt Korollar 3.1.9: $p \mid b - b'$. Das bedeutet insbesondere $p \leq |b - b'| < p$ – offensichtlich ein Widerspruch. \square

3.2.6 Beispiel Im Beispiel 3.1.7 haben wir mit dem erweiterten euklidischen Algorithmus $\text{ggT}(5610, 637) = 1$ und die Darstellung

$$1 = 145 \cdot 5610 - 1277 \cdot 637$$

gefunden. Also ist der Kandidat für die Inverse modulo 5610 von 637 die Zahl -1277 . Da diese nicht in \mathbb{Z}_{5610} liegt, addieren wir solange 5610, bis das Ergebnis das erste Mal in \mathbb{Z}_{5610} zu finden ist. Das ist bereits nach dem ersten Schritt der Fall: $-1277 + 5610 = 4333$ ist die Inverse modulo 5610 von 637: $637 \cdot 4333 \equiv_{5610} 1$.

3.2.7 Bemerkung Neben seiner Rolle bei den Beweisen von Korollar 3.1.9 und Satz 3.2.5 ist die Darstellung des größten gemeinsamen Teilers nach EUKLID vor allem das Hilfsmittel zur Berechnung der Inversen modulo p.

3.2.8 Bemerkung Wenn wir die Multiplikation so definieren, daß die Ergebnisse $\bmod p$ wieder in \mathbb{Z}_p liegen, dann bedeutet Satz 3.2.5, daß wir auch eine Division in \mathbb{Z}_p^* über $a \cdot b^{-1} \bmod p$ zur Verfügung haben ($b^{-1} \in \mathbb{Z}_p^*$ ist die Inverse modulo p von $b \in \mathbb{Z}_p^*$). Auf diese Weise können wir in $\mathbb{Z}_p^* \setminus \{0\}$ uneingeschränkt multiplizieren und dividieren. Eine Menge mit solchen Rechenoperationen nennt man eine *Gruppe* (diese Gruppe wird *Einheitengruppe* U_p für p genannt). Ist p sogar eine Primzahl (oder die Potenz einer solchen), dann läßt sich auch zu jedem $a \in \mathbb{Z}_p^*$ genau ein $b \in \mathbb{Z}_p^*$ mit der Eigenschaft $a + b \equiv_p 0$ finden. Jetzt ist es möglich, auf \mathbb{Z}_p^* alle Grundrechenarten durchzuführen. Eine solche Menge nennt man einen *Körper*. Bisher kennen wir nur die Körper \mathbb{Q} und \mathbb{R} , die jeweils unendlich viele Elemente enthalten. Im Gegensatz dazu ist \mathbb{Z}_p^* ein *endlicher Körper*. Ist p nicht zu groß, dann kann man sich das an der *Additions-* und *Multiplikationstabelle* veranschaulichen. Etwa für $\mathbb{Z}_7^* = \{0, 1, 2, 3, 4, 5, 6\} = \mathbb{Z}_7$:

$+/_7$	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

$\cdot/_7$	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

Dagegen $\mathbb{Z}_{12}^* = \{0, 1, 5, 7, 11\}$:

$\cdot/12$	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

Offensichtlich ist $1 + 5 = 6 \notin \mathbb{Z}_{12}^*$, so daß keine Hoffnung besteht, \mathbb{Z}_{12}^* zu einem Körper zu machen.

3.2.9 Satz (Kleiner Satz von FERMAT) Für jede Primzahl p und alle $a \in \mathbb{N}$ gilt

$$a^p \equiv_p a. \quad (3.11)$$

Ist darüber hinaus noch $\text{ggT}(a, p) = 1$ erfüllt, so folgt sogar

$$a^{p-1} \equiv_p 1. \quad (3.12)$$

Beweis. Den Beweis führen wir mit Hilfe vollständiger Induktion nach $a \in \mathbb{N}$.

Für $a = 1$ gibt es nichts zu zeigen.

$a \rightarrow a + 1$: Wir gehen von $a^p \equiv_p a$ aus und müssen $(a + 1)^p \equiv_p a + 1$ nachweisen.

$$(a + 1)^p = \sum_{k=0}^p \binom{p}{k} a^k = 1 + \sum_{k=1}^{p-1} \binom{p}{k} a^k + a^p \equiv_p 1 + a + \sum_{k=1}^{p-1} \binom{p}{k} a^k.$$

Wenn wir zeigen können, daß die letzte Summe modulo p gleich Null ist, sind wir fertig. Dazu müssen wir uns nur klar machen, daß die Binomialkoeffizienten $\binom{p}{k}$ für $1 \leq k \leq p - 1$ durch p teilbar sind: Es gilt $\binom{p}{k} \in \mathbb{N}$ und

$$\begin{aligned} k! \binom{p}{k} &= \frac{p!}{(p-k)!} = \frac{p(p-1)(p-2) \cdots (p-k+1) \cdot (p-k)!}{(p-k)!} \\ &= p(p-1)(p-2) \cdots (p-k+1). \end{aligned}$$

Das bedeutet insbesondere $p \mid k! \binom{p}{k}$, oder $p \mid k(k-1)(k-2) \cdots 3 \cdot 2 \cdot \binom{p}{k}$. Wegen $\text{ggT}(p, k) = \text{ggT}(p, k-1) = \cdots = \text{ggT}(p, 3) = \text{ggT}(p, 2) = 1$, folgt durch wiederholte Anwendung von Korollar 3.1.10: $p \mid \binom{p}{k}$.

Damit ist $a^p \equiv_p a$ gezeigt. Gilt jetzt auch noch $\text{ggT}(a, p) = 1$, so hat a nach Satz 3.2.5 eine Inverse b modulo p . Mit deren Hilfe erhalten wir

$$a^{p-1} \equiv_p a^{p-1} ab = a^p b \equiv_p ab \equiv_p 1. \quad \square$$

3.2.10 Lemma Für zwei Primzahlen $p \neq q$ folgt aus $a \equiv b \pmod{p}$ und $a \equiv b \pmod{q}$:

$$a \equiv b \pmod{pq}.$$

Beweis. Es gibt Zahlen $k, \ell \in \mathbb{Z}$, so daß $a = b + kp$ und $a = b + \ell q$ gilt. Daraus ergibt sich sofort $kp = \ell q$, d. h. $p \mid \ell q$. Wegen $\text{ggT}(p, q) = 1$ muß p die Zahl ℓ teilen, also $\ell = sp$ gelten, mit einem geeigneten $s \in \mathbb{Z}$. Das bedeutet $a = b + spq = b \pmod{pq}$. \square

3.2.11 Satz Für zwei Primzahlen $p \neq q$ und für eine Zahl $a \in \mathbb{N}$ mit $\text{ggT}(a, pq) = 1$ gilt

$$a^{(p-1)(q-1)} = 1 \pmod{pq}. \quad (3.13)$$

Beweis. Aus $\text{ggT}(a, pq) = 1$ folgt insbesondere $\text{ggT}(a, p) = 1$ und $\text{ggT}(a, q) = 1$, also die Voraussetzung für den kleinen Satz von FERMAT 3.2.9 bzgl. p und q . Demnach gilt $a^{p-1} = 1 \pmod{p}$ und $a^{q-1} = 1 \pmod{q}$. Dann folgt

$$\begin{aligned} a^{(p-1)(q-1)} &= (a^{p-1})^{q-1} \equiv_p 1^{q-1} = 1, \\ a^{(p-1)(q-1)} &= (a^{q-1})^{p-1} \equiv_q 1^{p-1} = 1. \end{aligned}$$

Nach Lemma 3.2.10 ergibt sich daraus $a^{(p-1)(q-1)} = 1 \pmod{pq}$. \square

3.3 RSA-Verschlüsselung

3.3.1 Satz (RIVEST, SHAMIR, ADLEMAN)

Das RSA-Verfahren verläuft nach folgender Vorschrift:

- i) Man wähle zwei verschiedene, große Primzahlen p und q und bilde damit $n := p \cdot q$.
- ii) Man wähle eine Zahl $1 < e \in \mathbb{N}$, die zu $\varphi := (p-1)(q-1)$ teilerfremd und kleiner als φ ist. Dann ist der öffentliche Schlüssel durch das Zahlenpaar $S_o := [e, n]$ gegeben.
- iii) $d \in \mathbb{N}$, $d < \varphi$, sei die Inverse von e modulo φ . Sie wird mit Hilfe des erweiterten euklidischen Algorithmus berechnet. Dann ist der private Schlüssel durch das Zahlenpaar $S_p := [d, n]$ gegeben.
- iv) Eine Nachricht $M \in \mathbb{N}_0$, $M < n$, wird mit S_o durch

$$C := M^e \pmod{n}$$

verschlüsselt und mit S_p durch

$$M = C^d \pmod{n}$$

entschlüsselt.

Beweis. Es ist nur zu zeigen, daß die Entschlüsselung $M \equiv_n C^d \equiv_n M^{e \cdot d}$ funktioniert. Für den Beweis machen wir eine Fallunterscheidung.

- i) $\text{ggT}(M, n) = 1$: Aus $\text{ggT}(M, pq) = 1$ folgt nach Satz 3.2.11 $M^{(p-1)(q-1)} \equiv_{pq} 1$. Wegen $ed \equiv_{\varphi} 1$ gibt es eine Zahl $\ell \in \mathbb{N}$ mit $ed = 1 + \ell\varphi = 1 + \ell(p-1)(q-1)$.

$$M^{ed} = M \cdot M^{\ell(p-1)(q-1)} = M \cdot (M^{(p-1)(q-1)})^{\ell} \equiv_{pq} M \equiv_n M.$$

- ii) $\text{ggT}(M, n) > 1$: M hat mit $n = pq$ einen gemeinsamen Teiler. Das kann nur entweder p oder q sein, denn der Teiler pq würde $M \geq n$ bedeuten, im Widerspruch zur Voraussetzung $M < n$. O.B.d.A. können wir von $\text{ggT}(M, p) = p$ und $\text{ggT}(M, q) = 1$, d.h. $M = tp$ für ein geeignetes $t \in \mathbb{N}$ ausgehen. Damit haben wir nach dem kleinen Satz von FERMAT $M^{q-1} \equiv_q 1$ und durch Potenzieren mit $\ell(p-1)$ auch $(M^{q-1})^{\ell(p-1)} \equiv_q 1$. Es gibt daher eine ganze Zahl r mit der Eigenschaft $(M^{q-1})^{\ell(p-1)} = 1 + rq$, so daß

$$M^{ed} = M \cdot (M^{q-1})^{\ell(p-1)} = M(1 + rq) = M + rMq = M + rtpq \equiv_n M. \quad \square$$

3.3.2 Beispiel $p := 101, q := 17, n = 1717, \varphi = 1600, M := 405$. Wir wählen $e := 411$.
Der erweiterte euklidische Algorithmus:

$$\begin{array}{ll} 1600 = 3 \cdot 411 + 367 & 367 = 1600 - 3 \cdot 411 \\ 411 = 1 \cdot 367 + 44 & 44 = 411 - 1 \cdot 367 \\ 367 = 8 \cdot 44 + 15 & 15 = 367 - 8 \cdot 44 \\ 44 = 2 \cdot 15 + 14 & 14 = 44 - 2 \cdot 15 \\ 15 = 1 \cdot 14 + 1 & 1 = 15 - 14 \end{array}$$

Also gilt tatsächlich $\text{ggT}(\varphi, 411) = 1$. Um die Inverse d von e modulo φ zu bestimmen, berechnen wir für den größten gemeinsamen Teiler 1 die Darstellung nach EUKLID:

$$\begin{array}{ll} 1 = 15 - 44 + 2 \cdot 15 & = 3 \cdot 15 - 44 \\ = 3 \cdot 367 - 24 \cdot 44 - 44 & = 3 \cdot 367 - 25 \cdot 44 \\ = 3 \cdot 367 - 25 \cdot 411 + 25 \cdot 367 & = 28 \cdot 367 - 25 \cdot 411 \\ = 28 \cdot 1600 - 84 \cdot 411 - 25 \cdot 411 & = 28 \cdot 1600 - 109 \cdot 411. \end{array}$$

Wir erhalten $-109 \cdot 411 = 1 - 28 \cdot 1600 \equiv_{\varphi} 1$. Die Inverse von e ist demnach -109 . Das RSA-Verfahren erfordert jedoch ein positives d . Das kann immer erreicht werden, indem ein geeignetes Vielfaches von φ zum Kandidaten für d addiert wird: $-109 + 1600 = 1491$.

Ergebnis: $S_o = [411, 1717], S_p = [1491, 1717]$.

Verschlüsselung:

$$\begin{array}{ll} C \equiv_n 405^{411} & = 405 \cdot (405^2)^{205} \\ =_n 405 \cdot 910^{205} & = 405 \cdot 910 \cdot (910^2)^{102} \\ =_n 405 \cdot 910 \cdot 506^{102} & = 405 \cdot 910 \cdot (506^2)^{51} \\ =_n 405 \cdot 910 \cdot 203^{51} & = 405 \cdot 910 \cdot 203 \cdot (203^2)^{25} \\ =_n 405 \cdot 910 \cdot 203 & =_n 809. \end{array}$$

Die verschlüsselte Nachricht ist $C = 809$.

Entschlüsselung:

$$\begin{array}{ll} C^d \equiv_n 809^{1491} & = 809 \cdot (809^2)^{745} \\ =_n 809 \cdot 304^{745} & = 809 \cdot 304 \cdot (304^2)^{372} \\ =_n 809 \cdot 304 \cdot 1415^{372} & = 809 \cdot 304 \cdot (1415^2)^{186} \end{array}$$

$$\begin{aligned}
&=_{\mathbb{N}} 809 \cdot 304 \cdot 203^{186} &= & 809 \cdot 304 \cdot (203^2)^{93} \\
&=_{\mathbb{N}} 809 \cdot 304 &=_{\mathbb{N}} & 405 = M.
\end{aligned}$$

3.4 Chinesischer Restsatz

Der folgende Satz ist für das RSA-Verfahren nicht nötig. Da wir aber gerade alle Hilfsmittel für seinen Beweis beisammen haben, nehmen wir ihn mit auf.

3.4.1 Satz (Chinesischer Restsatz) *Die Zahlen $m_1, \dots, m_n \in \mathbb{N}$ seien paarweise teilerfremd. Dann hat das Gleichungssystem*

$$\begin{aligned}
x &= a_1 \bmod m_1 \\
x &= a_2 \bmod m_2 \\
&\vdots \\
x &= a_n \bmod m_n
\end{aligned}$$

genau eine Lösung x modulo $m_1 m_2 \cdots m_n$.

Beweis. Es sei $\mu := m_1 m_2 \cdots m_n$ und $\mu_i := \prod_{k \neq i} m_k$. Dann gilt $\text{ggT}(m_i, \mu_i) = 1$, denn jeder Teiler von m_i muß laut Voraussetzung teilerfremd zu m_k sein ($k \neq i$) und nach Korollar 3.1.9 dann auch zu μ_i . Nach Satz 3.2.5 hat μ_i eine Inverse v_i modulo m_i . Also gilt $v_i \mu_i = 1 \bmod m_i$, aber $v_i \mu_i = 0 \bmod m_k$, für $k \neq i$, denn $m_k \mid \mu_i$. Wir definieren daher

$$x := \sum_{i=1}^n a_i v_i \mu_i.$$

Dann gilt $x = a_i \bmod m_i$ für $i = 1, \dots, n$, d. h., x ist eine Lösung. Für eine weitere Lösung x' dieses Gleichungssystems haben wir $m_i \mid (x - x')$ für $i = 1, \dots, n$. Da die m_i paarweise teilerfremd sind, folgt aus dem zweiten Teil von Korollar 3.1.9 schließlich $\mu \mid (x - x')$, also $x = x' \bmod \mu$. \square

3.4.2 Beispiel Ein klassisches Beispiel ist die Aufgabe, eine natürliche Zahl x zu finden, die jeweils den Rest 1 aufweist, wenn man sie durch 2, 3, 4, 5 und durch 6 teilt. Darüber hinaus soll sie auch noch durch 7 teilbar sein. Das bedeutet

$$\begin{aligned}
x &= 1 \bmod 2, & x &= 1 \bmod 3, & x &= 1 \bmod 4, \\
x &= 1 \bmod 5, & x &= 1 \bmod 6, & x &= 0 \bmod 7.
\end{aligned}$$

Der chinesische Restsatz läßt sich nicht sofort anwenden, weil die Reste 2, 4 und 6 nicht teilerfremd sind. Allerdings können wir die ersten fünf Gleichungen mit Hilfe von Lemma 3.4.5 zu $x = 1 \bmod 60$ zusammenfassen. Auf diese und die letzte Gleichung läßt sich Satz 3.4.1 anwenden: $m_1 := 60$, $m_2 := 7$, $\mu := 420$, $a_1 := 1$, $a_2 := 0$, $\mu_1 := 7$, $\mu_2 := 60$ ergibt $v_1 = 43$ als Inverse von $7 \bmod 60$ und $v_2 = 2$ als die Inverse von $60 \bmod 7$. Wir erhalten demnach $x = 1 \cdot 43 \cdot 7 + 0 \cdot 2 \cdot 60 = 301$. Alle anderen Lösungen haben die Form $301 + t \cdot 420$, für $t \in \mathbb{N}$.

Für dieses Beispiel benötigen wir noch etwas Handwerkszeug:

3.4.3 Definition Für natürliche Zahlen p_1, \dots, p_n bezeichnet $\text{kgV}(p_1, \dots, p_n)$ das kleinste gemeinsame Vielfache von p_1, \dots, p_n , also die Zahl

$$\text{kgV}(p_1, \dots, p_n) := \min \{ x \in \mathbb{N} \mid \exists_{t_1, \dots, t_n \in \mathbb{N}} x = t_1 p_1 = \dots = t_n p_n \}. \quad (3.14)$$

Natürlich ist $x = p_1 p_2 \cdots p_n$ ein gemeinsames Vielfaches der Zahlen p_1 bis p_n , so daß die Menge in (3.14) nicht leer ist. Falls p_1, \dots, p_n teilerfremd sind, ist $p_1 p_2 \cdots p_n$ auch schon das kleinste gemeinsame Vielfache dieser Zahlen.

Es gilt etwa $\text{kgV}(42, 120) = \frac{2^3 \cdot 3 \cdot 5 \cdot 2 \cdot 3 \cdot 7}{2 \cdot 3} = \frac{42 \cdot 120}{\text{ggT}(42, 120)} = 840$. Dieses Ergebnis ist kein Zufall:

3.4.4 Lemma $\text{kgV}(p, q)$ ist ein Teiler jedes gemeinsamen Vielfachen von p und q . Der Zusammenhang mit $\text{ggT}(p, q)$ ist

$$\text{kgV}(p, q) \text{ggT}(p, q) = pq. \quad (3.15)$$

Beweis. Seien $g := \text{ggT}(p, q)$ und $k := \text{kgV}(p, q)$. Dann gibt es Zahlen $s, t, m, n \in \mathbb{N}$, so daß $p = sg, q = tg$ und $k = mp = nq$ gilt.

Sei $\mu > 0$ ein gemeinsames Vielfaches von p und q und $k < \mu$. Dann gilt $\mu = dk + r$ für ein geeignetes $d \in \mathbb{N}$, und einen Rest $0 \leq r < k$. Insbesondere ist $r = \mu - dk$ ein gemeinsames Vielfaches von p und q , denn μ und k sind es. Wäre $r > 0$, dann müßte $r \geq k$ gelten, denn k ist das kleinste gemeinsame Vielfache von p und q : W! Also bleibt nur $r = 0$, d. h., $k \mid \mu$.

Natürlich gilt $\text{ggT}(s, t) =: u = 1$. Es gibt nämlich Zahlen $s', t' \in \mathbb{N}$, mit $s = u s'$ und $t = u t'$. Dann zeigt $p = s'ug$ und $q = t'ug$, daß ug ein gemeinsamer Teiler von p und q ist. Daher gilt $ug \leq g$, was nur für $u = 1$ möglich ist.

Jetzt zu (3.15). Der Ausdruck

$$\frac{pq}{g} = sq = tp$$

ist offensichtlich ein gemeinsames Vielfaches von p und q . Das bedeutet $k \mid \frac{pq}{g}$. Es gibt daher ein $\ell \in \mathbb{N}$ mit der Eigenschaft $sq = tp = \ell k = \ell m p = \ell n q$. Daraus folgt $t = \ell m$ und $s = \ell n$. Da s und t teilerfremd sind, muß $\ell = 1$ gelten. Das zeigt $\frac{pq}{g} = k$, also (3.15). \square

3.4.5 Lemma Für natürliche Zahlen p_1, p_2, \dots, p_n gilt

$$\forall_{i \in \{1, \dots, n\}} a = b \bmod p_i \quad \Rightarrow \quad a = b \bmod \text{kgV}(p_1, p_2, \dots, p_n).$$

Das ist eine Verallgemeinerung von Lemma 3.2.10.

Beweis. $a = b \bmod p_1$ und $a = b \bmod p_2$ bedeutet $a = b + tp_1 = b + sp_2$ für geeignete $t, s \in \mathbb{Z}$. Also ist $tp_1 = sp_2$ ein gemeinsames Vielfaches von p_1 und p_2 und wird nach Lemma 3.4.4 von $\text{kgV}(p_1, p_2)$ geteilt: $tp_1 = \alpha \text{kgV}(p_1, p_2)$. Das heißt $a = b + \alpha \text{kgV}(p_1, p_2)$, oder $a = b \bmod \text{kgV}(p_1, p_2)$. Zusammen mit $a = b \bmod p_3$ folgt damit $a = b \bmod \text{kgV}(\text{kgV}(p_1, p_2), p_3)$, also $a = b \bmod \text{kgV}(p_1, p_2, p_3)$, nach Übung 3.4.6, v). Das kann jetzt fortgesetzt werden. \square

3.4.6 A Für $p, q \in \mathbb{N}$ sei $g := \text{ggT}(p, q)$ und $k := \text{kgV}(p, q)$. Also gilt $p = sg$ und $q = tg$ und $k = pm = nq$ für geeignete $m, n, s, t \in \mathbb{N}$.

- i) Zeigen Sie mit Hilfe von Lemma 3.1.5, daß jeder gemeinsame Teiler von p und q ein Teiler von g ist und daß s und t teilerfremd sind.
- ii) Zeigen Sie, daß auch m und n teilerfremd sind.
- iii) Zeigen Sie $\text{ggT}(p, q, r) = \text{ggT}(\text{ggT}(p, q), r)$.
- iv) Zeigen Sie, daß es auch eine Darstellung nach EUKLID für $\text{ggT}(p, q, r)$ gibt und daß sie durch eine zu (3.5) analoge Formel gegeben ist.
- v) Zeigen Sie $\text{kgV}(p, q, r) = \text{kgV}(\text{kgV}(p, q), r)$, $\text{kgV}(p, q, r, s) = \text{kgV}(\text{kgV}(p, q, r), s)$, usw.
- vi) Zeigen Sie: Aus $\text{ggT}(p, q) = 1$, $\text{ggT}(a, b) = 1$ für $a, b \in \mathbb{N}$ und $pa = qb$ folgt $a = q$ und $b = p$.
- vii) Zeigen Sie: $ap = qb$ und $\text{ggT}(a, b) = 1$ für $a, b \in \mathbb{N}$ hat $ap = \text{kgV}(p, q)$ zur Folge.

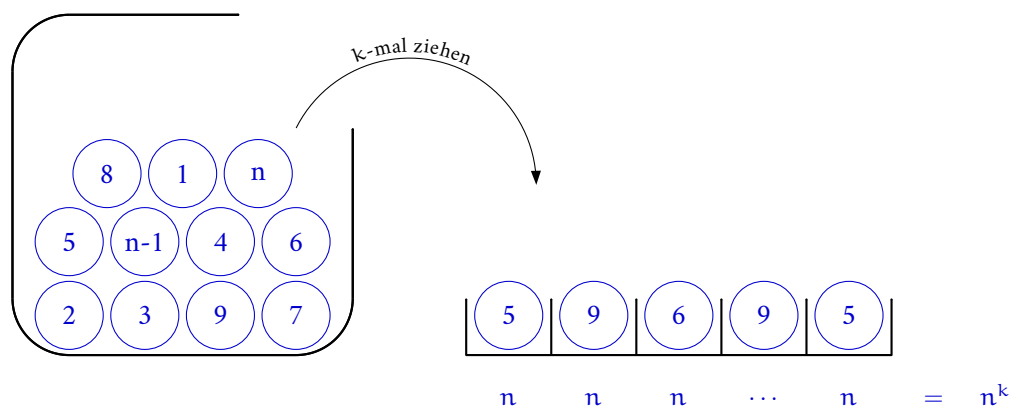
3.4.7 Beispiel Es sei $p := 30$, $q := 120$ und $r := 252$. Dann ist offensichtlich $\text{kgV}(p, q) = 120$, $\text{ggT}(120, 252) = 12$ und daher $\text{kgV}(30, 120, 252) = \frac{120 \cdot 252}{12} = 2520 = 84 \cdot p = 21 \cdot q = 10 \cdot r$. Die Zahlen 84, 21 und 10 sind teilerfremd. Damit ist nicht gemeint, daß sie paarweise teilerfremd sein müssen, sondern als Gesamtheit der drei Zahlen.

$\text{ggT}(30, 120, 252) = \text{ggT}(\text{ggT}(30, 120), 252) = \text{ggT}(30, 252) = 6$. Wegen $30 = -3 \cdot 30 + 1 \cdot 120$ und $6 = 17 \cdot 30 - 2 \cdot 252$ ist $6 = -51 \cdot 30 + 17 \cdot 120 - 2 \cdot 252$ eine mögliche Darstellung nach EUKLID für $\text{ggT}(30, 120, 252)$.

4 Kombinatorik

4.1 Die Urnenmodelle

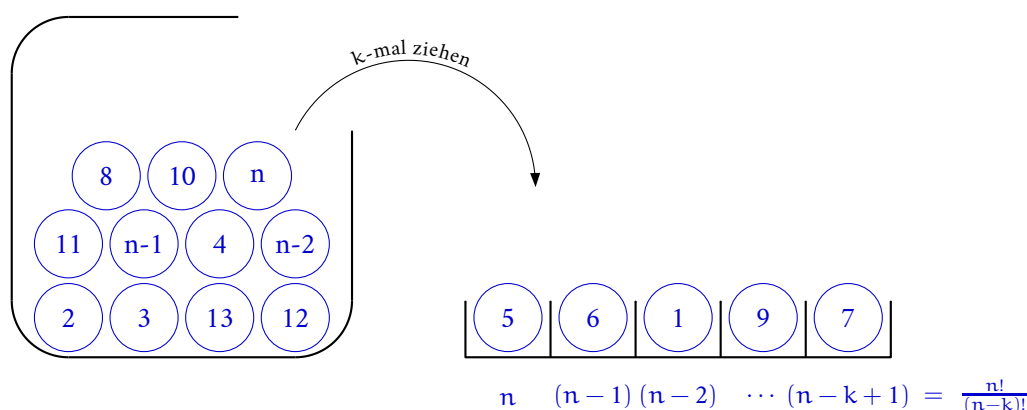
4.1.1 Ungeordnete Ziehung mit Zurücklegen Aus einer Urne mit n durchnummerierten Kugeln wird k -mal gezogen. Die erhaltene Kugel wird notiert. Anschließend wird sie wieder zurückgelegt. Die Reihenfolge der Ziehung bleibt unverändert. Daher gibt es für die erste Ziehung n mögliche Ergebnisse. Für die zweite aber auch, da in der Urne alle Kugeln wieder vorhanden sind. Für die ersten beiden Ziehungen haben wir demnach $n \cdot n = n^2$ mögliche Ziehungsergebnisse. Für die ersten drei dann n^3 , usw. Für die k Ziehungen ergeben sich so n^k Möglichkeiten.



4.1.2 Ungeordnete Ziehung ohne Zurücklegen Aus einer Urne mit n durchnummerierten Kugeln wird k -mal gezogen. Die erhaltenen Kugeln werden nicht zurückgelegt. Die Reihenfolge der Ziehung bleibt unverändert. Daher gibt es für die erste Ziehung n mögliche Ergebnisse. Für die zweite aber nur noch $n - 1$, da die Kugel der ersten Ziehung in der Urne fehlt. Für die ersten beiden Ziehungen haben wir demnach $n(n - 1)$ mögliche Ziehungsergebnisse. Für die ersten drei dann $n(n - 1)(n - 2)$, usw. Für die k Ziehungen ergeben sich auf diese Weise

$$n(n - 1) \cdots (n - k + 1) = \frac{n(n - 1) \cdots (n - k + 1) \cdot (n - k) \cdots 3 \cdot 2}{(n - k)!} = \frac{n!}{(n - k)!}$$

Möglichkeiten.

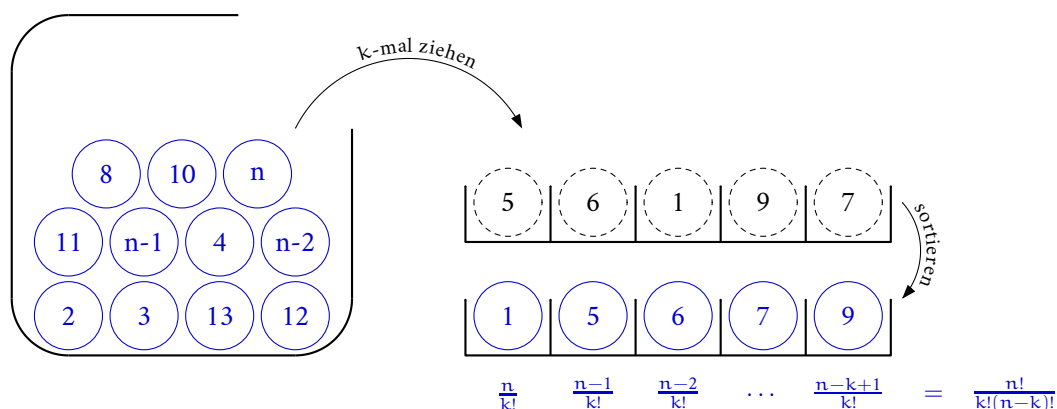


Einen wichtigen Spezialfall erhalten wir für $k = n$, also für den Fall, bei dem die Urne vollständig geleert wird. Es ergeben sich $\frac{n!}{(n-n)!} = n!$ Ziehungsergebnisse (vergl. (1.30)). Da die Ziehung ungeordnet ist, entstehen dabei alle möglichen Anordnungen der Zahlen 1 bis n , also alle *Permutationen* dieser Zahlen (vergl. Beispiel 4.1.9, 6), oder 2.5.10).

4.1.3 Satz Mit den Elementen einer n -elementigen Menge lassen sich $n!$ Permutationen bilden. D. h., es gibt $n!$ Möglichkeiten n unterscheidbare Objekte anzuordnen.

Es gibt $\frac{n!}{(n-k)!}$ Möglichkeiten für ungeordnete Ziehungen ohne Zurücklegen von k Kugeln aus einer Urne mit n Kugeln.

4.1.4 Geordnete Ziehung ohne Zurücklegen Aus einer Urne mit n durchnummerierten Kugeln wird k -mal gezogen. Die erhaltenen Kugeln werden nicht zurückgelegt. Die Reihenfolge der Ziehung ist unwichtig. Nach der Ziehung ordnet man daher die Kugeln nach aufsteigenden Nummern an. Dabei geht die Reihenfolge der ursprünglichen Ziehung natürlich verloren. Um die Anzahl der möglichen Ziehungsergebnissen zu erfahren, stellen wir uns zunächst alle möglichen Ausgänge vor, bevor die Kugeln sortiert werden. Das sind $\frac{n!}{(n-k)!}$ k -Tupel aus verschiedenen Zahlen, (wie etwa 56197 für $k = 5$) denn jetzt handelt es sich noch um eine ungeordnete Ziehung ohne Zurücklegen.



Wir gruppieren sie in Blöcke, die jeweils nur Tupel mit den gleichen k Zahlen, jedoch in unterschiedlicher Reihenfolge, aufweisen (also etwa 15679, 15697, 15967, ..., 97651). Jeder Block

enthält genau $k!$ Elemente, denn er enthält alle Permutationen seines ersten Tupels. Durch das Sortieren fallen für jeden Block diese $k!$ -Möglichkeiten zu einer einzigen zusammen (vergl. Beispiel 4.1.9, 2) und 3)). Die Zahl $\frac{n!}{(n-k)!}$ der ungeordneten Tupel reduziert sich daher um den Faktor $k!$ auf $\frac{n!}{k!(n-k)!} = \binom{n}{k}$.

Es gibt weitere Interpretationen dieses Ergebnisses. Die vielleicht wichtigste betrifft die Anzahl der Möglichkeiten, aus einer Menge mit n Elementen, k -elementige Teilmengen zu bilden. Das liegt daran, daß das Ziehungsergebnis ungeordnet ist. Das Ergebnis 15679 ist vollständig durch die Menge $\{1, 5, 6, 7, 9\}$ beschrieben, da die Reihenfolge ihrer Elemente definitionsgemäß nicht von Bedeutung ist. Jedes Ziehungsergebnis bestimmt eine k -elementige Teilmenge von $\{1, 2, \dots, n-1, n\}$ und jede solche Teilmenge einen möglichen Ausgang der geordneten Ziehung ohne Zurücklegen.

4.1.5 Satz $\binom{n}{k}$ ist die Anzahl

- i) der möglichen k -elementigen Teilmengen einer n -elementigen Menge,
- ii) der möglichen Anordnungen von k gleichen Objekten auf n Plätze,
- iii) der möglichen geordneten Ziehungen von k Kugeln aus einer Urne mit n Kugeln.

Eine Verallgemeinerung des Binomialkoeffizienten $\binom{n}{k}$ ist der Multinomialkoeffizient

$$\binom{n}{k_1, k_2, \dots, k_r} := \frac{n!}{k_1! k_2! \dots k_r!}, \quad (4.1)$$

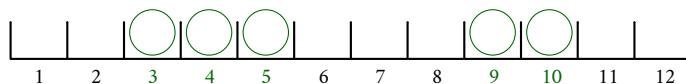
mit $k_1 + k_2 + \dots + k_r = n$, $k_i \in \mathbb{N}_0$. Er beschreibt die Anzahl der möglichen Anordnungen von k_1, \dots, k_r jeweils gleichen Objekten auf n Plätze. Außerdem übernimmt er die Rolle des Binomialkoeffizienten in der Verallgemeinerung des binomischen Lehrsatzes 1.2.16

$$(a_1 + a_2 + \dots + a_r)^n = \sum_{k_1 + \dots + k_r = n} \binom{n}{k_1, k_2, \dots, k_r} a_1^{k_1} a_2^{k_2} \dots a_r^{k_r}. \quad (4.2)$$

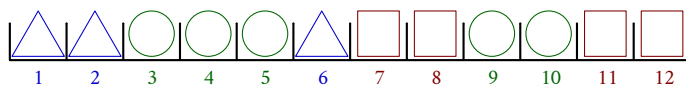
Bemerkung: Die Notation des Multinomialkoeffizienten geht für $r = 2$ nicht konform mit dem Binomialkoeffizienten. Die Übersetzung in diesem Fall ist $\binom{n}{k} = \binom{n}{k, n-k}$.

Beweis. i) Den Beweis haben wir im Wesentlichen schon erbracht. Wir müssen uns noch mit dem Sonderfall $\binom{n}{0} = 1$ beschäftigen. Auch hier stimmt die Behauptung, denn die einzige 0-elementige Menge ist die leere Menge, die Teilmenge jeder Menge ist.

ii) Die Anzahl möglicher Anordnungen machen wir uns an einem Beispiel klar: Eine Anordnung von 5 Kreisen auf 12 Plätze ist durch die Menge der Platznummern, im folgenden Beispiel $\{3, 4, 5, 9, 10\}$, eindeutig bestimmt. Daher gibt es genau so viele Anordnungen, wie 5-elementige Teilmengen der 12-elementigen Menge $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Das sind $\binom{12}{5}$.



Die Interpretation der Multinomialkoeffizienten $\binom{n}{k_1, k_2, \dots, k_r}$ demonstrieren wir am Beispiel der $\binom{12}{3, 4, 5}$ möglichen Anordnungen von 3 Dreiecken, 4 Quadraten und 5 Kreisen auf 12 Plätze.



Zunächst verteilen wir die 5 Kreise auf den 12 verfügbaren Plätzen. Dafür gibt es $\binom{12}{5}$ Möglichkeiten. Bei jeder sind 7 Plätze nicht durch Kreise besetzt, die wir jetzt mit den 3 Dreiecken und 4 Quadraten belegen können. Das geht jeweils auf $\binom{7}{3}$ verschiedene Weisen. Insgesamt gibt es daher

$$\binom{12}{5} \binom{7}{3} = \binom{12}{7} \binom{7}{3} = \frac{12!}{5! 7!} \cdot \frac{7!}{3! 4!} = \frac{12!}{3! 4! 5!} = \binom{12}{3, 4, 5}$$

Möglichkeiten, die Dreiecke, die Quadrate und die Kreise zu verteilen.

Die Verallgemeinerung des binomischen Lehrsatzes zeigen wir nur für den Fall $(a_1 + a_2 + a_3)^n$. Hier läßt sich das Wesentliche verstehen.

$$\begin{aligned} (a_1 + a_2 + a_3)^n &= (a_1 + (a_2 + a_3))^n \stackrel{(1.32)}{=} \sum_{k=0}^n \binom{n}{k} a_1^{n-k} (a_2 + a_3)^k \\ &= \sum_{k=0}^n \sum_{\ell=0}^k \binom{n}{k} \binom{k}{\ell} a_1^{n-k} a_2^{k-\ell} a_3^{\ell} = \sum_{k=0}^n \sum_{\ell=0}^k \frac{n!}{(n-k)! k!} \frac{k!}{(k-\ell)! \ell!} a_1^{n-k} a_2^{k-\ell} a_3^{\ell} \\ &= \sum_{k=0}^n \sum_{\ell=0}^k \binom{n}{n-k, k-\ell, \ell} a_1^{n-k} a_2^{k-\ell} a_3^{\ell} = \sum_{k_1+k_2+k_3=n} \binom{n}{k_1, k_2, k_3} a_1^{k_1} a_2^{k_2} a_3^{k_3}. \quad \square \end{aligned}$$

4.1.6 Korollar Die Anzahl aller Teilmengen einer n -elementigen Menge ist 2^n .

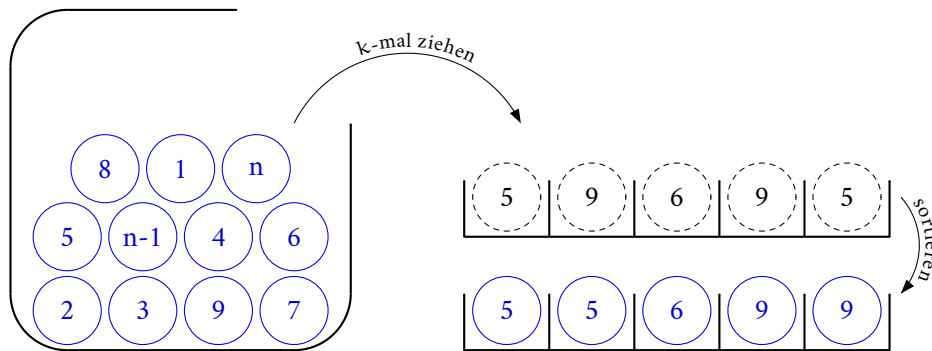
Beweis. Die Anzahl aller Teilmengen setzt sich aus der Anzahl der 0-, der 1-, der 2-elementigen usw. zusammen. Eine Anwendung des binomischen Lehrsatzes 1.2.16 ergibt dafür

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = \sum_{k=0}^n \binom{n}{k} 1^k \cdot 1^{n-k} = (1+1)^n = 2^n. \quad \square$$

4.1.7 Geordnete Ziehung mit Zurücklegen Durch geeignete Kodierung der Ziehungsergebnisse führen wir diesen Fall auf die Anordnung von Objekten zurück. Das Ergebnis 55699 etwa ist durch Angabe der Häufigkeiten k_i für das Auftreten der Zahl i eindeutig charakterisiert. Für 55699 heißt das $k_1 = k_2 = k_3 = k_4 = 0$, $k_5 = 2$, $k_6 = 1$, $k_7 = k_8 = 0$, $k_9 = 2$, ..., $k_n = 0$. Da wir $k = 5$ mal ziehen, muß $k_1 + k_2 + \dots + k_n = 5$ gelten. Jede mögliche Ziehung wird durch eine *Zerlegung*

$$k_1 + k_2 + \dots + k_n = k$$

der Anzahl k der Ziehungen in n Summanden $k_i \in \mathbb{N}_0$ eindeutig festgelegt. Wir müssen also nur noch bestimmen, wie viele Zerlegungen dieser Art eine natürliche Zahl k haben kann.



An Beispielen, wie $0+0+0+0+2+1+0+0+2+0 = 5$ für $k = 5$ und $n = 10$, demonstrieren wir die Kodierung, die es gestattet, das Ziehungsergebnis durch Anordnungen von Objekten auf einer festen Anzahl von Plätzen zu beschreiben. Wir ersetzen einfach die Häufigkeiten k_i für das Ziehungsergebnis i durch die entsprechende Anzahl von Punkten \bullet . Das bedeutet

$$\begin{aligned}
 0+0+0+0+2+1+0+0+2+0 &\simeq +++++\bullet\bullet+ \bullet + + + \bullet\bullet+ \\
 3+0+0+1+0+0+0+0+0+1 &\simeq \bullet\bullet\bullet + + + \bullet + + + + + \bullet \\
 1+0+1+1+0+0+1+0+1+0 &\simeq \bullet + + \bullet + \bullet + + + \bullet + + \bullet + \\
 &\dots \qquad \qquad \qquad \dots
 \end{aligned}$$

Offensichtlich handelt es sich jetzt nur noch darum, 5 Punkte auf $10 - 1 + 5 = 14$ Plätze zu verteilen. Dafür gibt es, wie wir bereits wissen, $\binom{14}{5} = 2002$ Möglichkeiten. Im allgemeinen Fall haben wir $n - 1 +$ -Zeichen, um mit den n Häufigkeiten k_1, \dots, k_n die Summe $k_1 + k_2 + \dots + k_n = k$ zu bilden. In der Kodierung bedeutet das, k Punkte auf $n - 1 + k$ Plätze zu verteilen, wofür es $\binom{n+k-1}{k}$ Möglichkeiten gibt.

4.1.8 Satz $\binom{n+k-1}{k}$ ist die Anzahl der Möglichkeiten

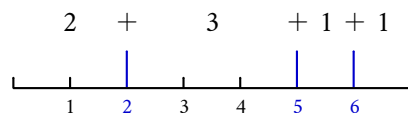
- i) einer geordneten Ziehung mit Zurücklegen von k Kugeln aus einer Urne mit n Kugeln,
- ii) die Zahl k in n nicht negative ganze Zahlen zu zerlegen.

$\binom{k-1}{n-1}$ ist die Anzahl der Möglichkeiten, die Zahl k in n natürliche Zahlen zu zerlegen.

Beweis. Es ist nur noch die Anzahl der Zerlegungen von k durch n natürliche Zahlen zu bestimmen. Offensichtlich muß jetzt $k \geq n$ gelten. Die Kodierung machen wir uns am Beispiel $k = 7$ und $n = 4$ klar:

$$\begin{aligned}
 7 &= 4+1+1+1 = 1+4+1+1 = 1+1+4+1 = 1+1+1+4 \\
 &= 3+2+1+1 = 3+1+2+1 = 3+1+1+2 = 1+3+1+2 = 1+1+3+2 = 1+3+2+1 \\
 &= 2+3+1+1 = 2+1+3+1 = 2+1+1+3 = 1+2+1+3 = 1+1+2+3 = 1+2+3+1 \\
 &= 2+2+2+1 = 2+2+1+2 = 2+1+2+2 = 1+2+2+2.
 \end{aligned}$$

Die Zerlegung $7 = 2 + 3 + 1 + 1$ etwa kodieren wir durch aufeinanderfolgende Strecken mit 2 Einheiten, 3 Einheiten und zwei mit der Länge einer Einheit:



Diese Anordnung ist eindeutig durch die Positionen 2, 5 und 6 bestimmt, bei denen die Strecken aneinander stoßen. Bei einer Streckenlänge k kommen $k-1$ Positionen für die $n-1$ möglichen Trennstellen der n aufeinander folgenden Strecken in Frage. Es geht also um die Anzahl der Möglichkeiten, $n-1$ Trennstellen aus $k-1$ auszuwählen. Das sind $\binom{k-1}{n-1}$, wie wir inzwischen wissen. \square

Die vier Urnenmodelle lassen sich übersichtlich in folgender Tabelle zusammenfassen (hier ist u bzw. g die Abkürzung für *ungeordnet* bzw. *geordnet* und $m Z$ bzw. $o Z$ für *mit Zurücklegen* bzw. *ohne Zurücklegen*):

	$m Z$	$o Z$
u	n^k	$\frac{n!}{(n-k)!}$
g	$\binom{n+k-1}{k}$	$\binom{n}{k}$

(4.3)

4.1.9 Beispiel Wir wählen $n = 5$ und $k = 3$, um die verschiedenen Ziehungsmodelle an einem konkreten Fall vorzustellen.

1) Ungeordnete Ziehung mit Zurücklegen. Es gibt $5^3 = 125$ Ergebnisse:

111	112	113	114	115	311	312	313	314	315	511	512	513	514	515
121	122	123	124	125	321	322	323	324	325	521	522	523	524	525
131	132	133	134	135	331	332	333	334	335	531	532	533	534	535
141	142	143	144	145	341	342	343	344	345	541	542	543	544	545
151	152	153	154	155	351	352	353	354	355	551	552	553	554	555
211	212	213	214	215	411	412	413	414	415					
221	222	223	224	225	421	422	423	424	425					
231	232	233	234	235	431	432	433	434	435					
241	242	243	244	245	441	442	443	444	445					
251	252	253	254	255	451	452	453	454	455					

2) Ungeordnete Ziehung ohne Zurücklegen. Es gibt $\frac{5!}{(5-3)!} = 5 \cdot 4 \cdot 3 = 60$ Möglichkeiten:

123	124	125	134	135	145	234	235	245	345
132	142	152	143	153	154	242	253	254	354
312	412	512	413	513	514	423	523	524	534
213	214	215	314	315	415	324	325	425	435
231	241	251	341	351	451	342	352	452	453
321	421	521	431	531	541	432	532	542	543

3) Geordnete Ziehung ohne Zurücklegen. Es gibt $\binom{5}{3} = \frac{5 \cdot 4 \cdot 3!}{3! \cdot 2!} = \frac{5 \cdot 4}{2} = 10$ Ergebnisse:

123	124	125	134	135	145	234	235	245	345
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

4) Geordnete Ziehung mit Zurücklegen. Es gibt $\binom{5+3-1}{3} = 35$ Möglichkeiten:

111	112	113	114	115	333	334	335
	122	123	124	125		344	345
		133	134	135			355
			144	145		444	445
				155			455
	222	223	224	225			555
		233	234	235			
			244	245			
				255			

5) Es gibt $\binom{5-1}{3-1} = 6$ Möglichkeiten, die Zahl 5 in 3 Summanden aus \mathbb{N} zu zerlegen, nämlich $5 = 3 + 1 + 1 = 1 + 3 + 1 = 1 + 1 + 3 = 2 + 2 + 1 = 2 + 1 + 2 = 1 + 2 + 2$, aber $\binom{3+5-1}{5} = 21$ mit Summanden aus \mathbb{N}_0 ($n = 3, k = 5$):

$$\begin{aligned}
 5 &= 5 + 0 + 0 &= 0 + 5 + 0 &= 0 + 0 + 5 \\
 &= 4 + 1 + 0 &= 4 + 0 + 1 &= 0 + 4 + 1 &= 1 + 4 + 0 &= 1 + 0 + 4 &= 0 + 1 + 4 \\
 &= 3 + 2 + 0 &= 3 + 0 + 2 &= 0 + 3 + 2 &= 2 + 3 + 0 &= 2 + 0 + 3 &= 0 + 2 + 3 \\
 &= 3 + 1 + 1 &= 1 + 3 + 1 &= 1 + 1 + 3 \\
 &= 2 + 2 + 1 &= 2 + 1 + 2 &= 1 + 2 + 2
 \end{aligned}$$

6) Es gilt $(a_1 + a_2 + a_3)^4 =$

$$\begin{aligned}
 &a_1^4 + a_2^4 + a_3^4 \\
 &+ 4a_1^3a_2 + 4a_1^3a_3 + 4a_2^3a_3 \\
 &+ 4a_1^2a_2^2 + 4a_1^2a_3^2 + 4a_2^2a_3^2 \\
 &+ 6a_1^2a_2a_3 + 6a_1a_2^2a_3 + 6a_1a_2a_3^2 \\
 &+ 12a_1^2a_2a_3 + 12a_1a_2^2a_3 + 12a_1a_2a_3^2
 \end{aligned}$$

7) Es gibt $4! = 24$ Möglichkeiten, die Zahlen 1, 2, 3 und 4 anzuordnen:

$$\begin{array}{cccccc}
 1, 2, 3, 4 & 1, 3, 2, 4 & 3, 1, 2, 4 & 2, 1, 3, 4 & 2, 3, 1, 4 & 3, 2, 1, 4 \\
 1, 2, 4, 3 & 1, 3, 4, 2 & 3, 1, 4, 2 & 2, 1, 4, 3 & 2, 3, 4, 1 & 3, 2, 4, 1 \\
 1, 4, 2, 3 & 1, 4, 3, 2 & 3, 4, 1, 2 & 2, 4, 1, 3 & 2, 4, 3, 1 & 3, 4, 2, 1 \\
 4, 1, 2, 3 & 4, 1, 3, 2 & 4, 3, 1, 2 & 4, 2, 1, 3 & 4, 2, 3, 1 & 4, 3, 2, 1
 \end{array}$$

Die $5! = 120$ Möglichkeiten, die Zahlen 1, 2, 3, 4, 5 anzuordnen, kann man aus den Permutationen von 1, 2, 3, 4 systematisch dadurch gewinnen, daß man die Zahl 5 an jede anhängt, sie dann jeweils an die vierte Position setzt, an die dritte, die zweite und die erste:

$$\begin{array}{cccccc}
 1, 2, 3, 4, 5 & 1, 3, 2, 4, 5 & 3, 1, 2, 4, 5 & 2, 1, 3, 4, 5 & 2, 3, 1, 4, 5 & 3, 2, 1, 4, 5 \\
 1, 2, 4, 3, 5 & 1, 3, 4, 2, 5 & 3, 1, 4, 2, 5 & 2, 1, 4, 3, 5 & 2, 3, 4, 1, 5 & 3, 2, 4, 1, 5 \\
 1, 4, 2, 3, 5 & 1, 4, 3, 2, 5 & 3, 4, 1, 2, 5 & 2, 4, 1, 3, 5 & 2, 4, 3, 1, 5 & 3, 4, 2, 1, 5 \\
 4, 1, 2, 3, 5 & 4, 1, 3, 2, 5 & 4, 3, 1, 2, 5 & 4, 2, 1, 3, 5 & 4, 2, 3, 1, 5 & 4, 3, 2, 1, 5 \\
 1, 2, 3, 5, 4 & 1, 3, 2, 5, 4 & 3, 1, 2, 5, 4 & 2, 1, 3, 5, 4 & 2, 3, 1, 5, 4 & 3, 2, 1, 5, 4 \\
 1, 2, 4, 5, 3 & 1, 3, 4, 5, 2 & 3, 1, 4, 5, 2 & 2, 1, 4, 5, 3 & 2, 3, 4, 5, 1 & 3, 2, 4, 5, 1 \\
 1, 4, 2, 5, 3 & 1, 4, 3, 5, 2 & 3, 4, 1, 5, 2 & 2, 4, 1, 5, 3 & 2, 4, 3, 5, 1 & 3, 4, 2, 5, 1 \\
 4, 1, 2, 5, 3 & 4, 1, 3, 5, 2 & 4, 3, 1, 5, 2 & 4, 2, 1, 5, 3 & 4, 2, 3, 5, 1 & 4, 3, 2, 5, 1 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 5, 1, 2, 3, 4 & 5, 1, 3, 2, 4 & 5, 3, 1, 2, 4 & 5, 2, 1, 3, 4 & 5, 2, 3, 1, 4 & 5, 3, 2, 1, 4 \\
 5, 1, 2, 4, 3 & 5, 1, 3, 4, 2 & 5, 3, 1, 4, 2 & 5, 2, 1, 4, 3 & 5, 2, 3, 4, 1 & 5, 3, 2, 4, 1 \\
 5, 1, 4, 2, 3 & 5, 1, 4, 3, 2 & 5, 3, 4, 1, 2 & 5, 2, 4, 1, 3 & 5, 2, 4, 3, 1 & 5, 3, 4, 2, 1 \\
 5, 4, 1, 2, 3 & 5, 4, 1, 3, 2 & 5, 4, 3, 1, 2 & 5, 4, 2, 1, 3 & 5, 4, 2, 3, 1 & 5, 4, 3, 2, 1
 \end{array}$$

4.1.10 Die VANDERMONDSche Identität Für $a, b \in \mathbb{N}$ und $b \leq a$ gilt:

$$\begin{aligned}
 (1+x)^a(1+x)^b &= (1+x)^{a+b} \\
 &= \sum_{n=0}^{a+b} \binom{a+b}{n} x^n = \sum_{k=0}^a \sum_{\ell=0}^b \binom{b}{k} \binom{a}{\ell} x^{k+\ell} \\
 &= \sum_{n=0}^b \sum_{k=0}^n \binom{b}{k} \binom{a}{n-k} x^n \\
 &\quad + \sum_{n=b+1}^a \sum_{k=0}^b \binom{b}{k} \binom{a}{n-k} x^n + \sum_{n=a+1}^{a+b} \sum_{k=n-a}^b \binom{b}{k} \binom{a}{n-k} x^n.
 \end{aligned}$$

b	b0	b1	b2	b3	b4	b5	b6
↑	30	31	32	33	34	35	36
k	20	21	22	23	24	25	26
1	10	11	12	13	14	15	16
0	00	01	02	03	04	05	06

$0 \quad 1 \quad \ell \quad \rightarrow \quad b \quad \rightarrow \quad a$

Die Doppelsumme wird entlang der Diagonalen konstanter $k + \ell$ ausgeführt. Ein Koeffizientenvergleich ergibt

$$\begin{aligned}
 \sum_{k=0}^n \binom{b}{k} \binom{a}{n-k} &= \binom{a+b}{n}, \quad n \leq b, & \sum_{k=0}^b \binom{b}{k} \binom{a}{n-k} &= \binom{a+b}{n}, \quad b \leq n \leq a, \\
 \sum_{k=n-a}^b \binom{b}{k} \binom{a}{n-k} &= \binom{a+b}{n}, \quad a \leq n \leq a+b.
 \end{aligned}$$

Treffen wir die Vereinbarung, daß $\binom{p}{q} := 0$ für $q < 0$, oder $q > p$ zu setzen ist, dann lassen sich diese drei Fälle für $a, b, n \in \mathbb{N}$, $n \leq a+b$, zur VANDERMONDSchen Identität zusammenfassen:

$$\sum_{k=0}^n \binom{b}{k} \binom{a}{n-k} = \binom{a+b}{n}. \quad (4.4)$$

Wir haben diese Identität rein algebraisch erhalten. Es gibt aber auch eine anschauliche Deutung. Dazu stellen wir uns eine Urne mit $a+b$ Kugeln vor, aus der n -mal gezogen wird. Dafür gibt es $\binom{a+b}{n}$ Möglichkeiten. Wir können uns aber auch vorstellen, daß wir die $a+b$ Kugeln auf zwei Urnen aufteilen, eine mit b und eine mit a Kugeln. Dann ziehen wir k mal aus der ersten und $n-k$ -mal aus der zweiten Urne, was im Ergebnis der Ziehung von n Kugeln aus der Urne mit $a+b$ Kugeln entspricht. Es gibt $\binom{b}{k}$ Ergebnisse für die Ziehung aus der ersten Urne, und für jede dieser Ziehungen jeweils $\binom{a}{n-k}$ Möglichkeiten für die zweite Urne. Das sind also $\binom{b}{k} \binom{a}{n-k}$ Möglichkeiten. Um alle Möglichkeiten zu erhalten, muß jetzt noch über k summiert werden, womit wir auch bei (4.4) angekommen sind.

Für $n = b$ erhalten wir, wegen $\sum_{k=0}^n \binom{n}{k} \binom{a}{n-k} = \sum_{k=0}^n \binom{n}{n-k} \binom{a}{n-k} = \sum_{k=0}^n \binom{n}{k} \binom{a}{k}$:

$$\sum_{k=0}^n \binom{n}{k} \binom{a}{k} = \binom{a+n}{n}. \quad (4.5)$$

Das ergibt für $a = n$:

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}. \quad (4.6)$$

Schreiben wir (4.4) systematisch in der Form

$$\sum_{k_1+k_2=n} \binom{a_1}{k_1} \binom{a_2}{k_2} = \binom{a_1+a_2}{n},$$

so erkennt man schnell, wie diese Beziehung verallgemeinert werden kann. Setzen wir $a_2 = a_3 + a_4$ und verwenden (4.4):

$$\begin{aligned} \binom{a_1+a_3+a_4}{n} &= \sum_{k_1+k_2=n} \binom{a_1}{k_1} \binom{a_3+a_4}{k_2} = \sum_{k_1+k_2=n} \sum_{k_3+k_4=k_2} \binom{a_1}{k_1} \binom{a_3}{k_3} \binom{a_4}{k_4} \\ &= \sum_{k_1+k_3+k_4=n} \binom{a_1}{k_1} \binom{a_3}{k_3} \binom{a_4}{k_4}. \end{aligned}$$

Diese Überlegung lässt sich wiederholen und ergibt schließlich die verallgemeinerte VANDERMONDSche Identität

$$\sum_{k_1+k_2+\dots+k_p=n} \binom{a_1}{k_1} \binom{a_2}{k_2} \dots \binom{a_r}{k_r} = \binom{a_1+a_2+\dots+a_r}{n}. \quad (4.7)$$

4.1.11 Beispiel (Lotto) Wir überlegen uns die Gewinnchancen der einzelnen Gewinnränge beim Lotto. Bekanntlich werden für ein Spiel auf einem quadratischen Feld mit den Zahlen 1 bis 49 sechs angekreuzt, sagen wir $\{1, 2, 3, 4, 5, 6\}$. Die Ziehung der Lottozahlen ist geordnet, ohne Zurücklegen: Nachdem sechs Kugeln aus einer Trommel mit 49 durchnummerierten Kugeln gezogen sind, werden die gezogenen Nummern in aufsteigender Reihenfolge als Ziehungsergebnis präsentiert. Da beim Ankreuzen der sechs Zahlen auf dem Tippzettel keine Reihenfolge festgelegt werden kann, darf sie für das Ziehungsergebnis auch keine Rolle spielen. Die Anzahl möglicher Ziehungen beträgt $\binom{49}{6} = 13983816$, also knapp 14 Millionen. Wenn man auch noch die Zusatzzahl spielt (eine Zahl zwischen 0 und 9, die auf dem Tippschein steht), multipliziert sich die Anzahl der Möglichkeiten noch einmal mit 10 und ergibt rund 140 Millionen Möglichkeiten. Da sechs Richtige nur auf eine Weise erzielt werden können, ist die Chance dafür $1 : \binom{49}{6} \approx 7.15 \cdot 10^{-8}$, und bei sechs Richtigen mit Zusatzzahl $\approx 7.15 \cdot 10^{-9}$.

Es gibt weitere Gewinnränge. 5 Richtige können auf $6 \cdot 43 = 258$ Arten getippt werden. Bei unserem fiktiven Beispiel $\{1, 2, 3, 4, 5, 6\}$ kommt man in diesen Gewinnrang, wenn genau eine der sechs Zahlen durch eine der 43 nicht gezogenen ersetzt wird, wie etwa $\{7, 2, 3, 4, 5, 6\}$, $\{1, 12, 3, 4, 5, 6\}$, ... $\{1, 2, 3, 4, 5, 49\}$. Genau 5 Richtige zu tippen hat demnach die Chance $258 : \binom{49}{6} \approx 1.84 \cdot 10^{-5} \approx 1 : 50\,000$.

Bei 4 Richtigen sind zwei der korrekten Zahlen durch nicht gezogene zu ersetzen. Für jede der $\binom{6}{2}$ Möglichkeiten, zwei Kugeln des tatsächlichen Ziehungsergebnisses auszuwählen, gibt es $\binom{43}{2}$ Weisen, sie durch zwei der 43 nicht gezogenen zu ersetzen. Das sind $\binom{6}{2} \cdot \binom{43}{2} = 13545$ Möglichkeiten. Die Chance dafür beträgt daher $\binom{6}{2} \cdot \binom{43}{2} : \binom{49}{6} \approx 9.69 \cdot 10^{-4} \approx 1 : 1000$.

3 Richtige sind auf $\binom{6}{3} \cdot \binom{43}{3} = 246820$ Weisen zu erhalten, mit einer Chance von $\binom{6}{3} \cdot \binom{43}{3} : \binom{49}{6} \approx 0.018 \approx 1 : 55$.

Obwohl sie nicht mehr zu den Gewinnrängen zählen, bestimmen wir noch die Möglichkeiten, genau zwei und genau eine Richtige zu tippen, nämlich $\binom{6}{4} \cdot \binom{43}{4} = 1\,851\,150$, bzw. $\binom{6}{5} \cdot \binom{43}{5} =$

5 775 588. Wie nicht anders zu erwarten, entfällt der Löwenanteil $\binom{6}{6} \cdot \binom{43}{6} = 6\,096\,454$ der Möglichkeiten darauf, alle Zahlen falsch zu tippen. Die Chance, beim Lotto nichts zu gewinnen, beläuft sich somit auf $(1\,851\,150 + 5\,775\,588 + 6\,096\,454) : \binom{49}{6} \approx 0.98 = 98 : 100$.

Wenn unsere Analyse der einzelnen Gewinnränge zutreffend ist, sollten sich die Anzahl ihrer Realisierungen zur Gesamtzahl $\binom{49}{6}$ aller Möglichkeiten aufsummieren. Das kann man anhand der konkreten Zahlen leicht überprüfen – oder wir vertrauen auf die Folgerung (4.5) aus der VANDERMONDSchen Identität:

$$\begin{aligned} & \binom{6}{0}\binom{43}{0} + \binom{6}{1}\binom{43}{1} + \binom{6}{2}\binom{43}{2} + \binom{6}{3}\binom{43}{3} + \binom{6}{4}\binom{43}{4} + \binom{6}{5}\binom{43}{5} + \binom{6}{6}\binom{43}{6} \\ &= \sum_{k=0}^6 \binom{6}{k}\binom{43}{k} \stackrel{(4.5)}{=} \binom{6+43}{6} = \binom{49}{6}. \end{aligned}$$

5 Vektorräume

5.1 \mathbb{R}^2 und \mathbb{R}^3 als Vektorraum

In Definition 1.2.4 haben wir die Produktmenge von zwei oder mehr Mengen eingeführt und als Beispiel dafür in 1.2.5 den \mathbb{R}^n durch

$$\mathbb{R}^n = \{ [x_1, x_2, \dots, x_n] \mid x_1, x_2, \dots, x_n \in \mathbb{R} \}$$

definiert. Die Elemente $\mathbf{x} = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$ bezeichnen wir als *Vektoren*. Vorläufig werden wir die Schreibweise dieser Vektoren den Gegebenheiten anpassen, d. h., benötigen wir die Koordinaten eines Vektors $\mathbf{x} \in \mathbb{R}^n$ im fließenden Text, so schreiben wir sie als Zeilen $[x_1, \dots, x_n]$, in abgesetzten Formeln dagegen als Spalten

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Wir unterscheiden Vektoren \mathbf{x} durch die fette Schreibweise von gewöhnlichen Zahlen x . Auf \mathbb{R}^n führen wir eine Addition für Vektoren und die sogenannte *Skalarmultiplikation* von Vektoren mit einer Zahl (Skalar) aus \mathbb{R} ein.

5.1.1 Definition Sind $\mathbf{x} = [x_1, \dots, x_n]$ und $\mathbf{y} = [y_1, \dots, y_n]$ Vektoren aus \mathbb{R}^n und $t \in \mathbb{R}$, so wird durch

$$\mathbf{x} + \mathbf{y} := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad (5.1)$$

$$t \cdot \mathbf{x} := \begin{bmatrix} t \cdot x_1 \\ \vdots \\ t \cdot x_n \end{bmatrix} \quad (5.2)$$

die Summe $\mathbf{x} + \mathbf{y}$ von \mathbf{x} und \mathbf{y} bzw. die Skalarmultiplikation $t \cdot \mathbf{x}$ von \mathbf{x} mit t definiert.

Natürlich werden wir meist $t\mathbf{x}$ statt $t \cdot \mathbf{x}$ schreiben. Außerdem ist es üblich, wie beim Rechnen mit gewöhnlichen Zahlen, das Symbol $-\mathbf{x}$ für $(-1) \cdot \mathbf{x}$ zu vereinbaren und dann $\mathbf{x} - \mathbf{y}$ für $\mathbf{x} + (-1) \cdot \mathbf{y}$ zu schreiben. Der Vektor $\mathbf{0} := [0, \dots, 0]$ wird als *Nullvektor* bezeichnet.

Eine geometrische Interpretation der Vektoren und ihrer Rechengesetze geben wir zunächst für den Fall \mathbb{R}^2 und später auch für \mathbb{R}^3 an. Wir stellen uns einen Vektor $\mathbf{x} = [x_1, x_2]$ als eine

gerichtete Strecke in einem zweidimensionalen Koordinatensystem vor. Ausgehend von einem beliebigen Punkt lautet die Vorschrift: Gehe x_1 Einheiten in Richtung der x_1 -Achse und x_2 Einheiten in Richtung der x_2 -Achse. Das Ende dieser Strecke markiert man durch eine Pfeilspitze, um die Richtung, in die der Vektor zeigt, kenntlich zu machen. Machen wir uns diese Interpretation von Vektoren zu eigen, so ergibt sich zwanglos, daß die Parallelverschiebung einer solchen gerichteten Strecke denselben Vektor darstellt, denn der Startpunkt in der oben angeführten Vorschrift ist beliebig. Wählen wir allerdings den Ursprung O des Koordinatensystems als Startpunkt, dann zeigen die Vektoren \mathbf{x} und \mathbf{y} auf Punkte $X := [x_1, x_2]$ bzw. $Y := [y_1, y_2]$, mit denselben Koordinaten x_1 und x_2 , wie die zugehörigen Vektoren.

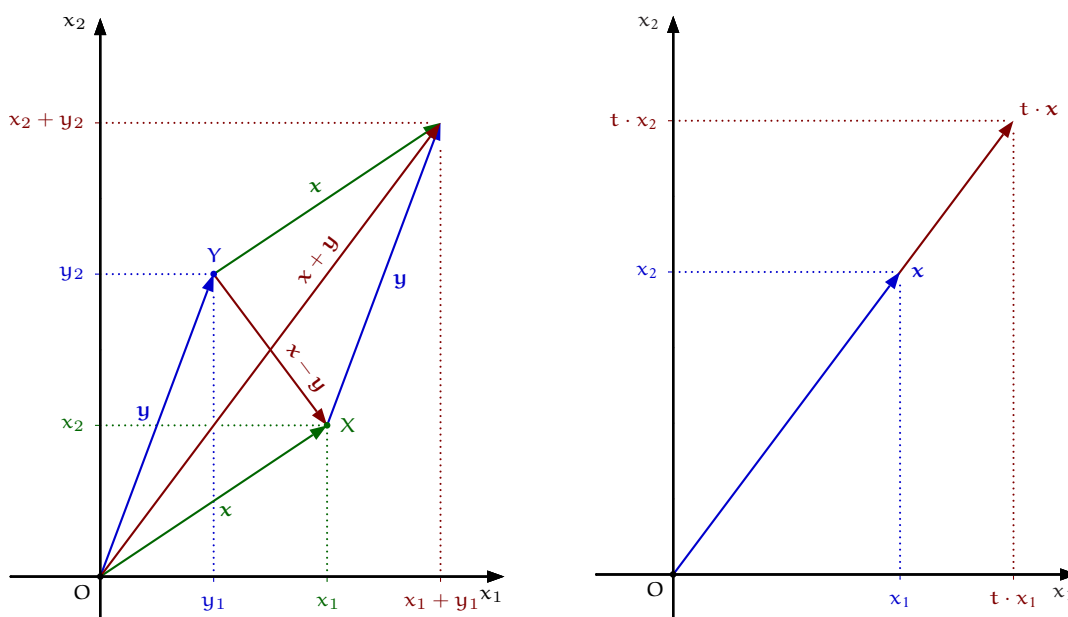
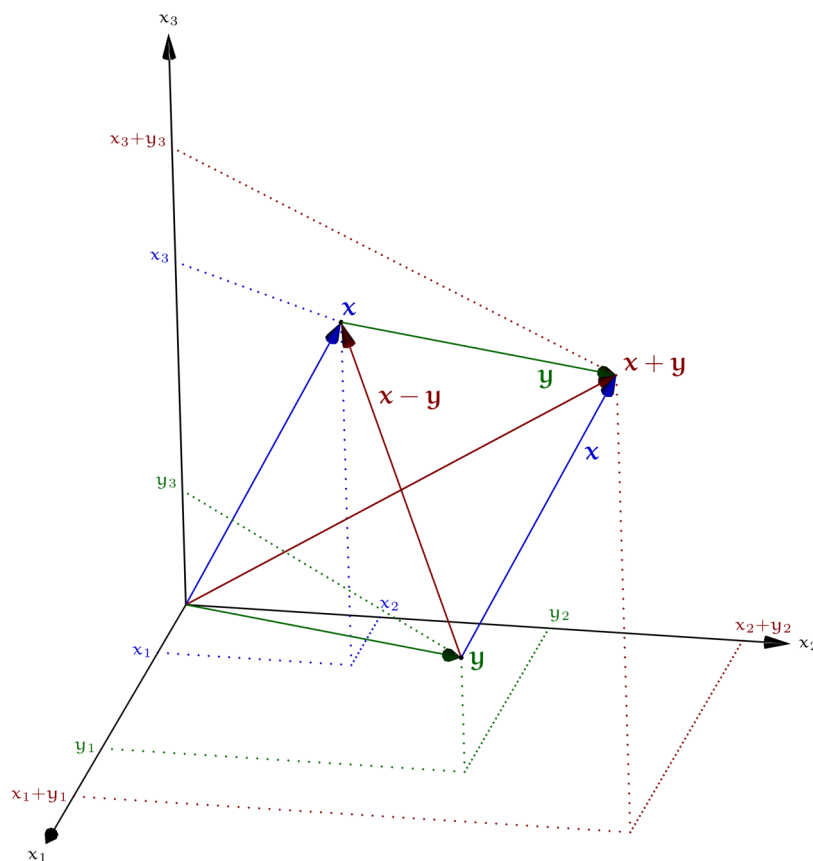


Abb. 5.1 Addition, Subtraktion und Skalarmultiplikation im \mathbb{R}^2

Die Addition $\mathbf{x} + \mathbf{y}$ von \mathbf{x} und \mathbf{y} ist die Hintereinanderausführung der Vorschriften \mathbf{x} und \mathbf{y} . Das heißt, sich x_1 Einheiten für \mathbf{x} und y_1 für \mathbf{y} in Richtung der x_1 -Achse zu bewegen, insgesamt also um $x_1 + y_1$ Einheiten. Dasselbe gilt für die x_2 -Richtung. Es ergibt sich so die Rechenvorschrift aus (5.1). Graphisch bedeutet das, die gerichtete Strecke \mathbf{y} an das Ende der gerichteten Strecke \mathbf{x} zu fügen. Abb. 5.1 zeigt, daß dieser Vorgang nicht von der Reihenfolge abhängt, daß also $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ gilt. Den Vektor, der bei Y startet und bei X endet bezeichnen wir mit \overrightarrow{YX} . Für ihn gilt $\mathbf{y} + \overrightarrow{YX} = \mathbf{x}$, also $\overrightarrow{YX} = \mathbf{x} - \mathbf{y}$ und insbesondere $\overrightarrow{OX} = \mathbf{x}$. Da X und \overrightarrow{OX} dieselben Koordinaten haben, werden wir meist X mit dem Vektor $\mathbf{x} = \overrightarrow{OX}$ identifizieren und von dem Punkt \mathbf{x} sprechen. Die Skalarmultiplikation $t\mathbf{x}$ eines Vektors \mathbf{x} mit dem Faktor t ist eine Streckung von \mathbf{x} um das t -Fache.

Die Interpretation der Vektoroperationen Addition, Subtraktion und Skalarmultiplikation bleibt in höheren Dimensionen, aber vornehmlich in der dritten, erhalten. Im \mathbb{R}^3 hat man gegenüber \mathbb{R}^2 lediglich eine weitere Richtung zur Verfügung, in die man sich bewegen kann, beschrieben durch die dritte Koordinate x_3 eines Vektors $\mathbf{x} = [x_1, x_2, x_3] \in \mathbb{R}^3$.

Abb. 5.2 Addition und Subtraktion im \mathbb{R}^3

Die allgemeinste Operation, die mit Vektoren durchzuführen ist, besteht in einer Kombination aus Streckung (Skalarmultiplikation) und Addition. D. h., wir werden keinen allgemeineren Konstruktionen, als den sogenannten *Linearkombinationen*

$$t_1 \mathbf{x}_1 + t_2 \mathbf{x}_2 + \cdots + t_m \mathbf{x}_m \quad (5.3)$$

aus Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_m$ und Skalaren t_1, \dots, t_m begegnen. In den folgenden Abschnitten stellen wir die grundlegenden Objekte *Gerade*, *Ebene* und *Determinante* im Raum \mathbb{R}^3 vor, wo sie noch sehr anschaulich sind, bevor wir sie später auf höhere Dimensionen verallgemeinern.

5.1.2 Geraden und Ebenen Eine Gerade g durch zwei verschiedene Punkte \mathbf{a} und \mathbf{b} wird durch diese eindeutig festgelegt. Sie läßt sich durch zwei Vektoren charakterisieren, nämlich

durch den sog. *Stützvektor* $\mathbf{q} := \mathbf{a}$ der Geraden und ihren *Richtungsvektor* $\mathbf{u} := \mathbf{b} - \mathbf{a}$:

$$g := \{ \mathbf{q} + t\mathbf{u} \mid t \in \mathbb{R} \}. \quad (5.4)$$

Man beachte dabei, daß die Gerade, d. h. die Menge g , durch A und B eindeutig bestimmt ist, nicht aber die Bestimmungsstücke \mathbf{q} und \mathbf{u} . Für \mathbf{q} hätte man durchaus auch \mathbf{b} wählen können und statt \mathbf{u} z. B. $2\mathbf{u}$ oder $-\mathbf{u}$ (vergl. Abb. 5.3).

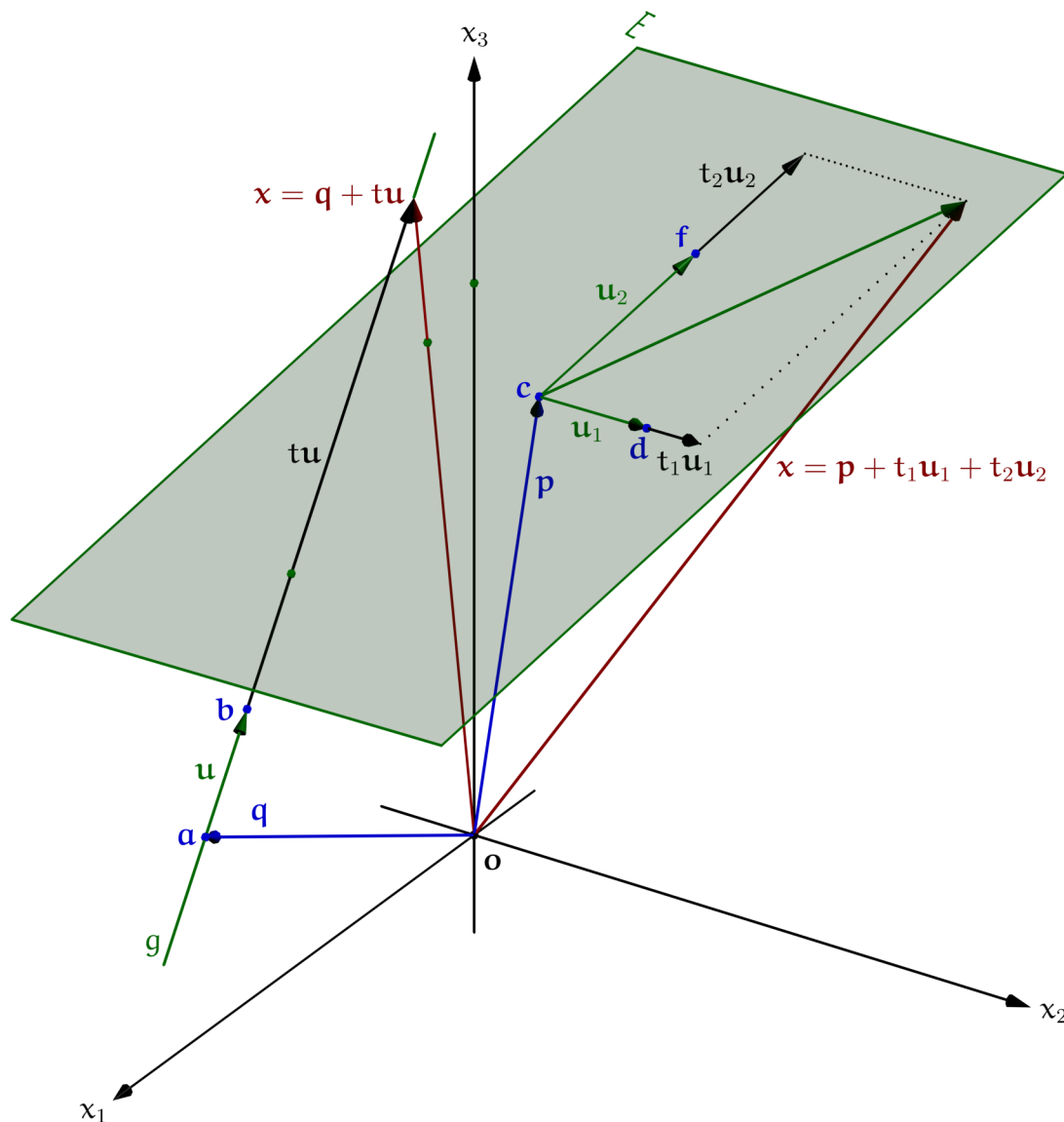


Abb. 5.3 Gerade und Ebene im \mathbb{R}^3

Eine Ebene E wird durch drei verschiedene Punkte c , d und f , die nicht auf einer Geraden liegen, eindeutig bestimmt. Wie für eine Gerade wählt man etwa $\mathbf{p} := c$ als Stützvektor und die beiden

Richtungsvektoren $\mathbf{u}_1 := \mathbf{d} - \mathbf{c}$, $\mathbf{u}_2 := \mathbf{f} - \mathbf{c}$. Dann ist die sogenannte *Parameterdarstellung* einer Ebene E folgendermaßen definiert:

$$E := \{ \mathbf{p} + t_1 \mathbf{u}_1 + t_2 \mathbf{u}_2 \mid t_1, t_2 \in \mathbb{R} \}. \quad (5.5)$$

5.1.3 Die Norm Unter der *euklidischen Norm* $\|\mathbf{x}\|$ eines Vektors \mathbf{x} (im Folgenden aber einfach als *Norm* bezeichnet) aus \mathbb{R}^2 oder \mathbb{R}^3 wollen wir zunächst die Länge der gerichteten Strecke \mathbf{x} verstehen. Für $\mathbf{x} = [x_1, x_2] \in \mathbb{R}^2$ wird sie also einfach über den Satz des PYTHAGORAS bestimmt:

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\| := \sqrt{x_1^2 + x_2^2}. \quad (5.6)$$

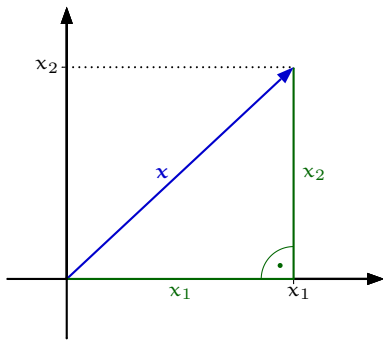
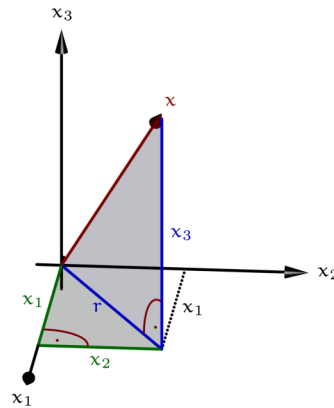


Abb. 5.4 Zur Norm im \mathbb{R}^2 und \mathbb{R}^3



Um die Norm eines Vektors $\mathbf{x} = [x_1, x_2, x_3] \in \mathbb{R}^3$ zu finden, bemühen wir den Satz des PYTHAGORAS zweimal. Für das rechtwinklige Dreieck mit den Katheten r und x_3 (vergl. Abb. 5.4) mit der Hypotenuse $\|\mathbf{x}\|$ erhalten wir $\|\mathbf{x}\|^2 = r^2 + x_3^2$. Das Dreieck mit den Katheten x_1, x_2 und der Hypotenuse r liefert uns den fehlenden Ausdruck $r^2 = x_1^2 + x_2^2$. Eingesetzt ergibt sich die Norm von \mathbf{x} zu

$$\left\| \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right\| := \sqrt{x_1^2 + x_2^2 + x_3^2}. \quad (5.7)$$

Den Abstand zweier Punkte \mathbf{x} und \mathbf{y} bezeichnen wir mit $d(\mathbf{x}, \mathbf{y})$. Er ist offensichtlich die Länge des Verbindungsvektors $\mathbf{x} - \mathbf{y}$ (vergl. Abb. 5.2):

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}. \quad (5.8)$$

5.1.4 Skalarprodukt Der Winkel α zwischen Vektoren $\mathbf{x} = [x_1, x_2, x_3]$ und $\mathbf{y} = [y_1, y_2, y_3]$ lässt sich mit Hilfe des *Skalarprodukts*

$$\langle \mathbf{x} | \mathbf{y} \rangle := \|\mathbf{x}\| \|\mathbf{y}\| \cos(\alpha) \quad (5.9)$$

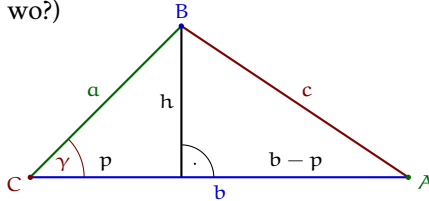
dieser beiden Vektoren bestimmen, wenn wir einen Weg finden, die rechte Seite nur unter Verwendung der Koordinaten x_i und y_i von \mathbf{x} bzw. \mathbf{y} auszurechnen. Dann lässt sich (5.9) nämlich nach $\cos(\alpha)$ auflösen und anschließend nach α . Der Weg zur Formel für (5.9) verläuft über den Kosinussatz.

5.1.5 Satz (Kosinussatz) Für ein ebenes Dreieck mit den Seiten a, b, c und dem Winkel γ zwischen a und b gilt

$$c^2 = a^2 + b^2 - 2ab \cos(\gamma). \quad (5.10)$$

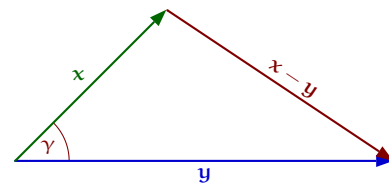
Für $\gamma = \frac{\pi}{2}$ ergibt sich der Satz des PYTHAGORAS.

Beweis. Der Skizze entnimmt man $p = a \cos(\gamma)$ und $c^2 = h^2 + (b - p)^2 = a^2 - p^2 + b^2 + p^2 - 2bp = a^2 + b^2 - 2ab \cos(\gamma)$. (Es wurde zweimal der Satz des PYTHAGORAS verwendet, wo?)



Wegen $\cos\left(\frac{\pi}{2}\right) = 0$ ergibt sich der Satz des PYTHAGORAS als Sonderfall. Das ist natürlich nicht weiter überraschend, da wir ihn für den Beweis ja vorausgesetzt haben. \square

Jetzt können wir uns daran machen, eine Rechenmethode für das Skalarprodukt zu finden. Wir wenden den Kosinussatz auf die Situation in nebenstehender Skizze an. Dabei ist $a = \|\mathbf{x}\|$, $b = \|\mathbf{y}\|$ und $c = \|\mathbf{x} - \mathbf{y}\|$:



$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos(\gamma) \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x} | \mathbf{y} \rangle, \end{aligned}$$

also

$$\begin{aligned} 2\langle \mathbf{x} | \mathbf{y} \rangle &= x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2 \\ &= x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2 \\ &\quad - x_1^2 - 2x_1y_1 - y_1^2 - x_2^2 - 2x_2y_2 - y_2^2 - x_3^2 - 2x_3y_3 - y_3^2 \\ &= 2x_1y_1 + 2x_2y_2 + 2x_3y_3. \end{aligned}$$

Wir erhalten die einfache Formel

$$\langle \mathbf{x} | \mathbf{y} \rangle = x_1y_1 + x_2y_2 + x_3y_3. \quad (5.11)$$

Natürlich entfällt der dritte Summand x_3y_3 , wenn die Vektoren aus \mathbb{R}^2 sind. Wählen wir für \mathbf{y} den Vektor \mathbf{x} , dann zeigt (5.7) den engen Zusammenhang des Skalarprodukts mit der Norm:

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}. \quad (5.12)$$

Mit Hilfe der Gleichung (5.11) ist es nun eine leichte Übungsaufgabe, die Eigenschaften

$$\text{i) } \langle \mathbf{x} | \mathbf{x} \rangle \geq 0, = 0 \iff \mathbf{x} = \mathbf{0}, \quad (\text{Definitheit}) \quad (5.13)$$

$$\text{ii) } \langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle, \quad (\text{Symmetrie}) \quad (5.14)$$

$$\text{iii) } \langle t\mathbf{x} + s\mathbf{y} | \mathbf{z} \rangle = t\langle \mathbf{x} | \mathbf{z} \rangle + s\langle \mathbf{y} | \mathbf{z} \rangle \quad (\text{Linearität}) \quad (5.15)$$

nachzurechnen, die den Namen *Skalarprodukt* rechtfertigen. Diese Rechenregeln sind in der ursprünglichen Definition (5.9) kaum zu erkennen.

5.1.6 Beispiel Es sei $\mathbf{x} := [8, -4, 1]$, $\mathbf{y} := [2, 5, 14]$ und $\mathbf{z} := [4, 7, -4]$. Dann ist gemäß (5.7) $\|\mathbf{x}\| = \sqrt{64 + 16 + 1} = \sqrt{81} = 9$, $\|\mathbf{y}\| = \sqrt{4 + 25 + 196} = \sqrt{225} = 15$ und $\|\mathbf{z}\| = 9$. Das Skalarprodukt ergibt sich nach (5.11) zu $\langle \mathbf{x} | \mathbf{y} \rangle = 16 - 20 + 14 = 10$. Nun lösen wir (5.9) nach $\cos(\alpha)$ auf:

$$\cos(\alpha) = \frac{\langle \mathbf{x} | \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{10}{9 \cdot 15} = \frac{2}{27}.$$

Wir erhalten $\alpha = \cos^{-1}(\frac{2}{27}) \approx 1.497$, oder $\alpha \approx 1.497 \cdot \frac{180^\circ}{\pi} \approx 85.75^\circ$, wenn wir das Ergebnis statt in Bogenmaß lieber in Grad angeben wollen.

Für den Winkel β zwischen \mathbf{x} und \mathbf{z} bestimmen wir

$$\langle \mathbf{x} | \mathbf{z} \rangle = 32 - 28 - 4 = 0.$$

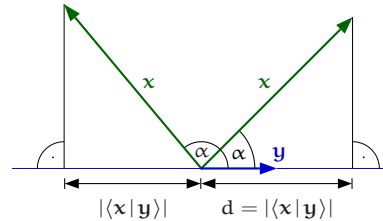
In diesem Fall brauchen wir die Formel für $\cos(\beta)$ gar nicht mehr bemühen, denn der Kosinus wird in dem uns ausschließlich interessierenden Bereich $\beta \in [0, \pi]$ nur an der Stelle $\beta = \frac{\pi}{2}$ Null. Die beiden Vektoren \mathbf{x} und \mathbf{z} schließen also einen rechten Winkel ein, d. h. sie stehen senkrecht aufeinander. Man sagt auch \mathbf{x} und \mathbf{z} sind *orthogonal* und schreibt dafür mitunter $\mathbf{x} \perp \mathbf{z}$:

$$\mathbf{x} \perp \mathbf{z} : \Longleftrightarrow \langle \mathbf{x} | \mathbf{z} \rangle = 0. \quad (5.16)$$

5.1.7 Geometrische Interpretation des Skalarprodukts Ist einer der beiden Vektoren \mathbf{x} oder \mathbf{y} , sagen wir \mathbf{y} , *normiert*, d. h. gilt $\|\mathbf{y}\| = 1$, dann läßt sich der Betrag $|\langle \mathbf{x} | \mathbf{y} \rangle|$ des Skalarprodukts als Länge der Projektion von \mathbf{x} auf die Richtung \mathbf{y} interpretieren.

An der Skizze sehen wir $|\cos(\alpha)| = \frac{d}{\|\mathbf{x}\|}$. Dabei ist d die Länge der Projektion von \mathbf{x} auf die Richtung \mathbf{y} . Wegen $\|\mathbf{y}\| = 1$ folgt zusammen mit (5.9)

$$d = \|\mathbf{x}\| |\cos(\alpha)| = \|\mathbf{x}\| \|\mathbf{y}\| \cos(\alpha) = |\langle \mathbf{x} | \mathbf{y} \rangle|.$$



Wenn wir den Betrag in $|\langle \mathbf{x} | \mathbf{y} \rangle|$ weglassen, dann sagt uns das Vorzeichen von $\langle \mathbf{x} | \mathbf{y} \rangle$ auch noch die Richtung der Projektion relativ zur Richtung von \mathbf{y} . Ist es positiv, dann liegt sie in der Richtung in die \mathbf{y} zeigt. Ist es negativ, dann liegt sie in der entgegengesetzten Richtung, denn $\cos(\alpha)$ ist für $\frac{\pi}{2} < \alpha < \pi$ negativ.

5.1.8 Satz (CAUCHY-SCHWARZ-Ungleichung) Für alle Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ gilt

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|. \quad (5.17)$$

Das Gleichheitszeichen gilt dabei genau dann, wenn \mathbf{x} und \mathbf{y} parallel sind.

Beweis. Die Ungleichung (5.17) ist eine einfache Konsequenz aus der Tatsache $|\cos(\alpha)| \leq 1$: $|\langle \mathbf{x} | \mathbf{y} \rangle| = \|\mathbf{x}\| \|\mathbf{y}\| |\cos(\alpha)| \leq \|\mathbf{x}\| \|\mathbf{y}\|$. Falls \mathbf{x} und \mathbf{y} parallel sind, ist $\alpha = 0$ oder $\alpha = \pi$ und $\cos(\alpha) = 1$ bzw. $\cos(\alpha) = -1$, also $|\cos(\alpha)| = 1$ und damit $|\langle \mathbf{x} | \mathbf{y} \rangle| = \|\mathbf{x}\| \|\mathbf{y}\|$. Gehen wir andererseits von dieser Gleichheit aus, dann muß $|\cos(\alpha)| = 1$ gelten, was nur bei $\alpha = 0$ oder $\alpha = \pi$ vorkommt. Für diese Winkel sind \mathbf{x} und \mathbf{y} parallel. \square

5.1.9 Das Kreuzprodukt Anders als das Skalarprodukt in \mathbb{R}^3 , das aus zwei Vektoren \mathbf{x} und \mathbf{y} eine Zahl $\langle \mathbf{x} | \mathbf{y} \rangle$ aus \mathbb{R} (einen *Skalar*) macht, gibt es im \mathbb{R}^3 ein weiteres Produkt, daß als Ergebnis einen Vektor aus \mathbb{R}^3 liefert. Es wird als *Kreuzprodukt* $\mathbf{x} \times \mathbf{y}$ bezeichnet und hat die schöne Eigenschaft (neben weiteren), einen Vektor zu erzeugen, der senkrecht auf den beiden Ausgangsvektoren steht. Ein Vektor \mathbf{n} , der senkrecht auf $\mathbf{x} = [x_1, x_2, x_3]$ und senkrecht auf $\mathbf{y} = [y_1, y_2, y_3]$ steht, muß die beiden Gleichungen $\langle \mathbf{x} | \mathbf{n} \rangle = x_1 n_1 + x_2 n_2 + x_3 n_3 = 0$ und $\langle \mathbf{y} | \mathbf{n} \rangle = y_1 n_1 + y_2 n_2 + y_3 n_3 = 0$ erfüllen. Da wir das GAUSS-Verfahren zur Lösung eines solchen *linearen Gleichungssystems* noch nicht zur Verfügung haben, (siehe 6.1, Seite 99) geben wir das Ergebnis einfach an und rechnen dafür nach, daß es die gewünschten Eigenschaften aufweist:

$$\mathbf{x} \times \mathbf{y} := \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}. \quad (5.18)$$

5.1.10 Satz Das Kreuzprodukt $\mathbf{x} \times \mathbf{y}$ hat folgende Eigenschaften:

$$(\mathbf{t}\mathbf{x} + \mathbf{s}\mathbf{y}) \times \mathbf{z} = \mathbf{t}\mathbf{x} \times \mathbf{z} + \mathbf{s}\mathbf{y} \times \mathbf{z} \quad (5.19)$$

$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}, \quad (5.20)$$

$$\mathbf{x} \times \mathbf{y} = \mathbf{0} \iff \mathbf{x} \parallel \mathbf{y} \quad (5.21)$$

$$\mathbf{x} \times \mathbf{y} \perp \mathbf{x}, \quad \mathbf{x} \times \mathbf{y} \perp \mathbf{y}, \quad (5.22)$$

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin(\alpha), \quad (5.23)$$

α ist der Winkel, den \mathbf{x} und \mathbf{y} einschließen. $\|\mathbf{x} \times \mathbf{y}\|$ ist der Flächeninhalt des Parallelogramms, das von \mathbf{x} und \mathbf{y} aufgespannt wird.

Beweis: (5.19) und (5.20) sind simple Übungen. (5.21): $\mathbf{x} \parallel \mathbf{y}$ bedeutet, daß \mathbf{x} ein Vielfaches von \mathbf{y} ist: $\mathbf{x} = \mathbf{t}\mathbf{y}$. Es folgt $\mathbf{x} \times \mathbf{y} = \mathbf{t}\mathbf{x} \times \mathbf{y} = -\mathbf{t}\mathbf{x} \times \mathbf{x} = -\mathbf{x} \times \mathbf{y}$, also $\mathbf{x} \times \mathbf{y} = \mathbf{0}$. Dabei haben wir (5.20) auf $\mathbf{x} \times \mathbf{x}$ angewandt. Für die Umkehrung können wir o. B. d. A. von $\mathbf{x} \neq \mathbf{0}$ und $\mathbf{y} \neq \mathbf{0}$ ausgehen. Dann ist wenigstens eine der Koordinaten von \mathbf{x} , sagen wir x_1 ungleich 0. Aus $\mathbf{x} \times \mathbf{y} = \mathbf{0}$ folgt dann mit (5.18):

$$x_1 \mathbf{y} = \begin{bmatrix} x_1 y_1 \\ x_1 y_2 \\ x_1 y_3 \end{bmatrix} = \begin{bmatrix} x_1 y_1 \\ x_2 y_1 \\ x_3 y_1 \end{bmatrix} = y_1 \mathbf{x}, \quad \text{also} \quad \mathbf{x} \parallel \mathbf{y}.$$

(5.22) ist eine einfache Rechenaufgabe:

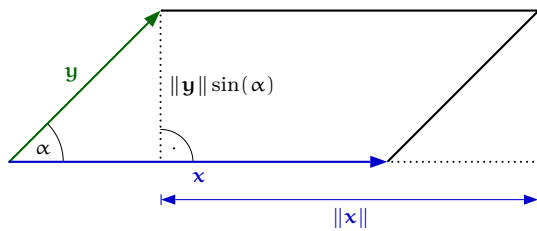
$$\begin{aligned} \langle \mathbf{x} \times \mathbf{y} | \mathbf{x} \rangle &= (x_2 y_3 - x_3 y_2)x_1 + (x_3 y_1 - x_1 y_3)x_2 + (x_1 y_2 - x_2 y_1)x_3 \\ &= x_2 y_3 x_1 - x_3 y_2 x_1 + x_3 y_1 x_2 - x_1 y_3 x_2 + x_1 y_2 x_3 - x_2 y_1 x_3 = 0. \end{aligned}$$

$\langle \mathbf{x} \times \mathbf{y} | \mathbf{y} \rangle = -\langle \mathbf{y} \times \mathbf{x} | \mathbf{y} \rangle = 0$. Das zeigt (5.22).

(5.23) ist aufwendiger (vergl. Übung 7.4.6):

$$\|\mathbf{x} \times \mathbf{y}\|^2 = (x_2 y_3 - x_3 y_2)^2 + (x_3 y_1 - x_1 y_3)^2 + (x_1 y_2 - x_2 y_1)^2$$

$$\begin{aligned}
&= x_2^2 y_3^2 + x_3^2 y_2^2 - 2x_2 x_3 y_2 y_3 + x_3^2 y_1^2 + x_1^2 y_3^2 - 2x_1 x_3 y_1 y_3 \\
&\quad + x_1^2 y_2^2 + x_2^2 y_1^2 - 2x_1 x_2 y_1 y_2 + x_1^2 y_1^2 + x_2^2 y_2^2 + x_3^2 y_3^2 - x_1^2 y_1^2 - x_2^2 y_2^2 - x_3^2 y_3^2 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1 y_1 + x_2 y_2 + x_3 y_3)^2 \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - \langle \mathbf{x} | \mathbf{y} \rangle^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 (1 - \cos^2(\alpha)) = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2(\alpha).
\end{aligned}$$



Da $\sin(\alpha)$ in dem uns interessierenden Bereich $\alpha \in [0, \pi]$ nicht negativ ist, können wir aus obiger Gleichung auf beiden Seiten die Wurzel ziehen und erhalten (5.23). Die Interpretation von $\|\mathbf{x}\| \|\mathbf{y}\| \sin(\alpha)$ als Flächeninhalt des Parallelogramms, das von \mathbf{x} und \mathbf{y} erzeugt wird, sieht man an der Skizze. \square

Die etwas sperrige Formel (5.18) läßt sich durch das folgende grafische Verfahren leicht merken: Die Koordinaten von \mathbf{x} und die von \mathbf{y} werden jeweils zweimal untereinander geschrieben und dann wie folgt über Kreuz multipliziert:

$$\begin{array}{cc}
\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_1 \\ x_2 \\ x_3 \end{array} & \begin{array}{c} y_1 \\ y_2 \\ y_3 \\ y_1 \\ y_2 \\ y_3 \end{array}
\end{array}
: \begin{bmatrix} x_2 y_3 - y_2 x_3 \\ x_3 y_1 - y_3 x_1 \\ x_1 y_2 - y_1 x_2 \end{bmatrix} = \mathbf{x} \times \mathbf{y}. \quad (5.24)$$

5.1.11 Die Normalenform der Ebene Eine Ebene $E \subset \mathbb{R}^3$ läßt sich durch einen Punkt $\mathbf{p} \in E$ und einen Vektor \mathbf{n} , der in eine zu E senkrechten Richtung zeigt, charakterisieren. Genauer: \mathbf{n} steht für alle $\mathbf{x}, \mathbf{y} \in E$ senkrecht auf $\mathbf{x} - \mathbf{y}$. \mathbf{n} ist der sogenannte *Normalenvektor* von E . Er ist bis auf ein Vielfaches durch die Ebene festgelegt. Verlangen wir, daß \mathbf{n} die Länge 1 hat, so machen wir das gelegentlich durch die Schreibweise \mathbf{n}_0 deutlich. Für einen weiteren Vektor $\mathbf{x} \in E$ gilt also

$$\langle \mathbf{x} - \mathbf{p} | \mathbf{n}_0 \rangle = \langle \mathbf{x} | \mathbf{n}_0 \rangle - \langle \mathbf{p} | \mathbf{n}_0 \rangle = 0.$$

Für alle Vektoren $\mathbf{x} \in E$ ist demnach das Skalarprodukt mit dem Normalenvektor \mathbf{n}_0 die feste Zahl $d_0 := \langle \mathbf{p} | \mathbf{n}_0 \rangle$. Gemäß 5.1.7 handelt es sich, bis auf ein eventuelles Vorzeichen, um die Länge der Projektion von \mathbf{x} auf die Richtung \mathbf{n}_0 . Laut Abb. 5.5 ist das aber gerade der Abstand der Ebene E vom Ursprung O (durch Änderung des Vorzeichens von \mathbf{n}_0 kann immer $d_0 \geq 0$ erreicht werden). Also liegen alle Punkte von E in der Menge $\{\mathbf{x} \in \mathbb{R}^3 \mid \langle \mathbf{x} | \mathbf{n}_0 \rangle = d_0\}$. Abb. 5.5 zeigt aber auch, daß alle Vektoren \mathbf{x} mit derselben Projektion d_0 in Richtung \mathbf{n}_0 zu E gehören (wir wollen es hier mit der Anschauung bewenden lassen, da ein strenger Beweis eine genauere Kenntnis der Begriffe *Dimension* und *Basis* erfordert, die uns im Augenblick noch nicht zur Verfügung stehen). Damit haben wir die Normalenform einer Ebene gefunden, die einen vorgegebenen Punkt \mathbf{p} enthält und den Normalenvektor \mathbf{n}_0 der Länge 1 hat:

$$E = \{ \mathbf{x} \in \mathbb{R}^3 \mid \langle \mathbf{x} | \mathbf{n}_0 \rangle = d_0 \}, \quad d_0 = \langle \mathbf{p} | \mathbf{n}_0 \rangle, \quad \|\mathbf{n}_0\| = 1. \quad (5.25)$$

Man kann in (5.25) auf die Forderung $\|\mathbf{n}_0\| = 1$ verzichten, wenn man die Interpretation von d_0 als Abstand der Ebene zum Ursprung nicht benötigt. Aus Abb. 5.5 läßt sich auch eine Formel für den Abstand $d(\mathbf{y}, E)$ eines Punktes \mathbf{y} von der Ebene E ablesen. $\langle \mathbf{y} | \mathbf{n}_0 \rangle$ ist die Projektion von \mathbf{y} auf die Richtung \mathbf{n}_0 . Ziehen wir davon den Abstand d_0 ab, so erhalten wir, bis auf ein mögliches Vorzeichen (warum?) den Abstand $d(\mathbf{y}, E)$. Wir haben damit die sogenannte *HESSE-Form* für den Abstand eines Punktes \mathbf{y} von der Ebene E gefunden. Für praktische Rechnungen möchte man sich gern von der unbequemen Forderung $\|\mathbf{n}_0\| = 1$ befreien können. Die Ebenengleichung lautet dann für einen beliebigen Normalenvektor \mathbf{n} : $\langle \mathbf{y} | \mathbf{n} \rangle = d$, mit einer festen Zahl d , die i. Allg. nicht mehr als Abstand interpretierbar ist. Um den Anschluß an (5.25) zu finden, muß \mathbf{n} normiert werden, was einfach durch $\mathbf{n}_0 = \frac{1}{\|\mathbf{n}\|} \mathbf{n}$ erreicht wird. Die zugehörige Ebenengleichung lautet jetzt $\langle \mathbf{y} | \mathbf{n}_0 \rangle = \frac{d}{\|\mathbf{n}\|} = d_0$. Setzen wir das in

$$d(\mathbf{y}, E) = |\langle \mathbf{y} | \mathbf{n}_0 \rangle - d_0| \quad (5.26)$$

ein: $d(\mathbf{y}, E) = \left| \left\langle \mathbf{y} \left| \frac{\mathbf{n}}{\|\mathbf{n}\|} \right\rangle - \frac{d}{\|\mathbf{n}\|} \right|$, also

$$d(\mathbf{y}, E) = \frac{1}{\|\mathbf{n}\|} |\langle \mathbf{y} | \mathbf{n} \rangle - d|. \quad (5.27)$$

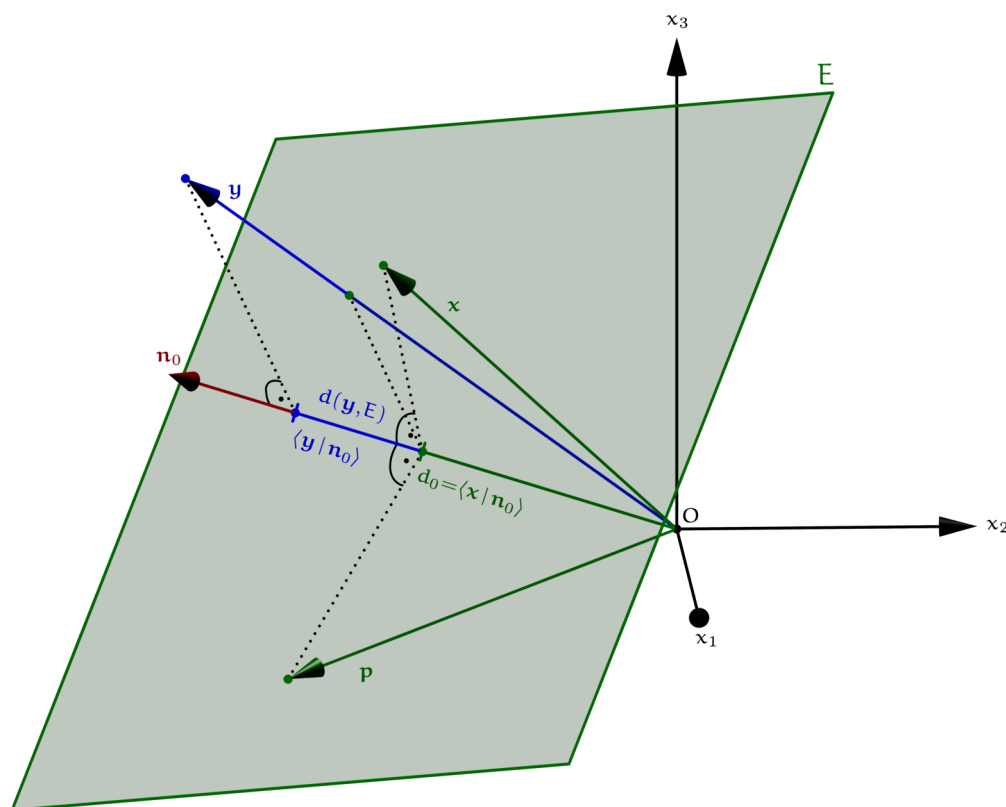


Abb. 5.5 Zur Normalenform einer Ebene

5.1.12 Beispiel Wir gehen von drei Punkten \mathbf{p} , \mathbf{q} und \mathbf{r} aus, bestimmen die Normalenform der Ebene E , die diese Punkte enthält und berechnen anschließend den Abstand eines weiteren Punktes \mathbf{s} von E .

$$\mathbf{p} := \frac{1}{4} [5, -8, -2], \mathbf{q} := \frac{1}{4} [12, 1, 2], \mathbf{r} := \frac{1}{4} [3, -6, 6], \mathbf{s} := \frac{1}{2} [2, -5, 4].$$

Wir wählen \mathbf{p} als Stützvektor, $\mathbf{u}_1 := 4(\mathbf{q} - \mathbf{p}) = [7, 9, 4]$ und $\mathbf{u}_2 := 2(\mathbf{r} - \mathbf{p}) = [-1, 1, 4]$ als Richtungsvektoren. Die Parameterdarstellung von E lautet damit

$$E = \left\{ \frac{1}{4} \begin{bmatrix} 5 \\ -8 \\ -2 \end{bmatrix} + t_1 \begin{bmatrix} 7 \\ 9 \\ 4 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix} \mid t_1, t_2 \in \mathbb{R} \right\}.$$

Der Normalenvektor ist ein geeignetes Vielfaches des Kreuzprodukts $\mathbf{u}_1 \times \mathbf{u}_2$:

$$\begin{bmatrix} 7 & -1 \\ 9 & 1 \\ 4 & 4 \\ 7 & -1 \\ 9 & 1 \\ 4 & 4 \end{bmatrix} : \begin{bmatrix} 36 & - & 4 \\ -4 & - & 28 \\ 7 & + & 9 \end{bmatrix} = \begin{bmatrix} 32 \\ -32 \\ 16 \end{bmatrix}. \text{ Wir wählen also } \mathbf{n} := \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Die Ebenengleichung lautet damit $\langle \mathbf{x} | \mathbf{n} \rangle = \langle \mathbf{q} | \mathbf{n} \rangle = 6 - 2 + 2 = 6$, oder ausgeschrieben in der sogenannten *Koordinatenform*: $2x_1 - 2x_2 + x_3 = 6$. Das ist die Version, mit der üblicherweise gerechnet wird. Wir haben jetzt $E = \{\mathbf{x} \in \mathbb{R}^3 \mid 2x_1 - 2x_2 + x_3 = 6\}$. Die HESSE-Form ergibt sich aus $|\langle \mathbf{y} | \mathbf{n} \rangle - 6|$ einfach durch Teilen mit $\|\mathbf{n}\| = 3$:

$$d(\mathbf{y}, E) = \frac{1}{3} \cdot |2y_1 - 2y_2 + y_3 - 6|.$$

$$\text{Also ist } d(\mathbf{s}, E) = \frac{1}{3} \cdot |2 + 5 + 2 - 6| = 1.$$

5.1.13 Spatprodukt und Determinante Wir haben für das Skalarprodukt und das Kreuzprodukt geometrische Interpretationen gefunden. Daher sollte auch eine für die Kombination $\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle$ dieser beiden Produkte möglich sein. Wir werden gleich sehen, daß es sich dabei im Wesentlichen um das Volumen des *Spats* handelt, der von den Vektoren \mathbf{x} , \mathbf{y} und \mathbf{z} aufgespannt wird (Abb. 5.6). Daher rührt der Name *Spatprodukt* für $\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle$.

Wir wissen, daß $\|\mathbf{x} \times \mathbf{y}\|$ den Inhalt der Grundfläche des Spats angibt (5.1.7).

Die Projektion von \mathbf{z} auf $\mathbf{n}_0 := \frac{\mathbf{x} \times \mathbf{y}}{\|\mathbf{x} \times \mathbf{y}\|}$ hat die Länge $h := \|\mathbf{h}\| = |\langle \mathbf{n}_0 | \mathbf{z} \rangle| = \frac{1}{\|\mathbf{x} \times \mathbf{y}\|} |\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle|$. Um also $\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle$ interpretieren zu können, müssen wir zunächst $h \cdot \|\mathbf{x} \times \mathbf{y}\|$ verstehen. Das ist allerdings leicht, denn es handelt sich einfach um das Volumen des senkrechten Prismas über der Grundfläche des Spats (die gestrichelten Linien in Abb. 5.6). Worüber man etwas nachdenken muß, ist, daß es auch das Volumen des Spats ist.

Die Idee besteht darin, den Spat über volumengleiche Körper in das Prisma zu verwandeln. Im ersten Schritt wird vom Spat entlang der Fläche \mathbf{xabc} ein Keil abgetrennt, der an der gegenüberliegenden Seite wieder angesetzt wird (die gepunkteten Linien in Abb. 5.6).

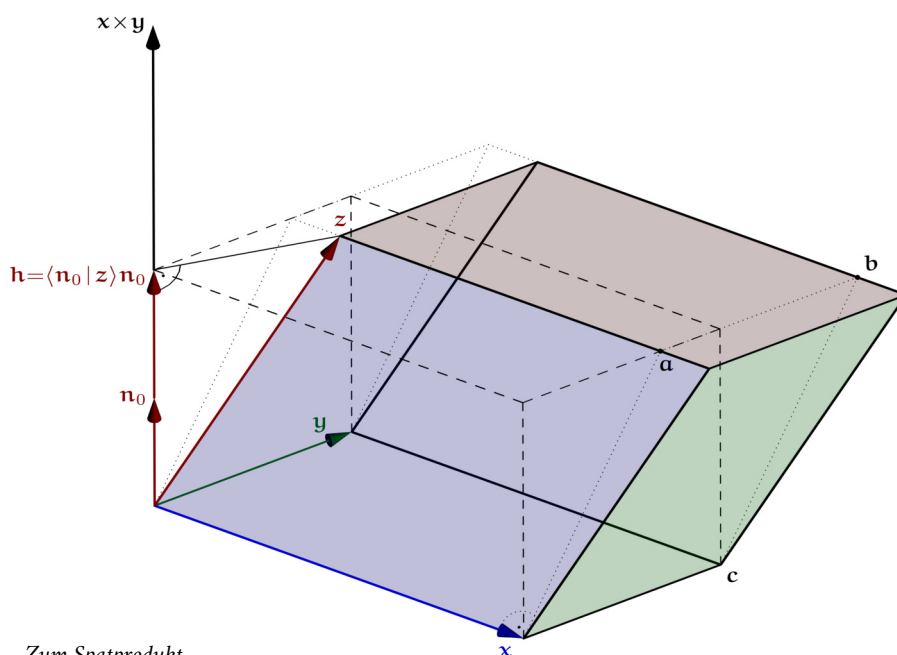


Abb. 5.6 Zum Spatprodukt

Im zweiten Schritt wiederholt man diesen Vorgang entlang der y -Richtung und gelangt so zum Prisma. Man überlege sich die nötigen Zwischenschritte, falls der Spat sehr viel schiefer ist, als in Abb. 5.6.

Als Zwischenergebnis haben wir, daß $|\langle x \times y | z \rangle|$ das Volumen des Spats wiedergibt, der von den Vektoren x , y und z aufgespannt wird. Da das Skalarprodukt auch negativ werden kann, muß $\langle x \times y | z \rangle$ noch eine Information enthalten, die über das Volumen hinaus geht.

In Abb. 5.6 ist das Spatprodukt positiv. Ersetzen wir z durch $-z$, so erhalten wir einen Spat gleichen Volumens. Der Ausdruck $\langle x \times y | -z \rangle$ wird jetzt aber negativ. Vergleicht man die Lage der sogenannten *Dreibeine* x , y , z und x , y , $-z$, so erkennt man, daß sie durch keine Bewegung im Raum zur Deckung gebracht werden können (ausprobieren). Ordnet man den Daumen, Zeige-

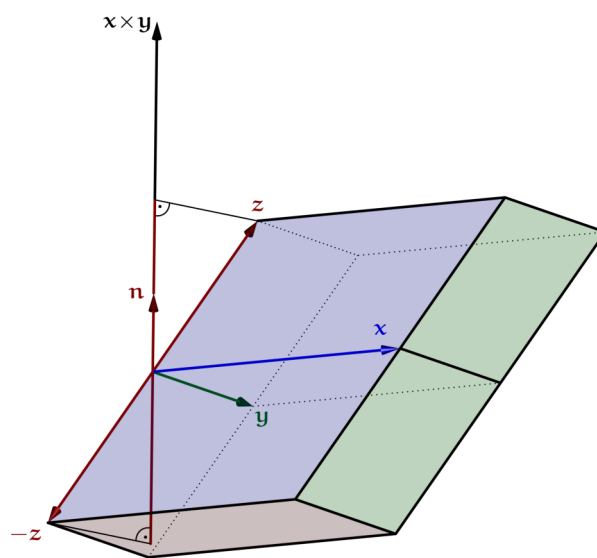


Abb. 5.7 Orientierung

und Ringfinger der rechten Hand so zueinander an, daß sie nicht in einer Ebene liegen, dann läßt sie sich so drehen, daß x in Richtung des Daumens, y in Richtung des Zeigefingers und z

in Richtung des Mittelfingers zeigt. Für $\mathbf{x}, \mathbf{y}, -\mathbf{z}$ ist das nicht möglich. Man sagt, das Dreibein $\mathbf{x}, \mathbf{y}, \mathbf{z}$ gehorcht der *Rechte-Hand-Regel*. Das Spatprodukt kann offensichtlich zwischen diesen beiden Situationen unterscheiden. Man vereinbart daher, daß drei Vektoren (*mathematisch*) *positiv orientiert* heißen, wenn ihr Spatprodukt positiv ist und andernfalls (*mathematisch*) *negativ orientiert*. Die Verallgemeinerung dieser Überlegungen auf höhere Dimensionen führt auf den Begriff der *Determinante*. Daher werden wir für das Spatprodukt auch die Schreibweise $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ verwenden und von der Determinante der Vektoren $\mathbf{x}, \mathbf{y}, \mathbf{z}$ sprechen.

Vertauschen wir die Reihenfolge der Vektoren \mathbf{x}, \mathbf{y} und \mathbf{z} , so ändert sich dabei offensichtlich nicht das Volumen des Spats, da dieses ja nicht von der Abfolge abhängen kann, in der wir die Vektoren aufzählen. Allerdings kann sich die Orientierung ändern, was am Vorzeichen der Determinante erkennbar ist. So ist etwa $\det(\mathbf{y}, \mathbf{x}, \mathbf{z}) = \langle \mathbf{y} \times \mathbf{x} | \mathbf{z} \rangle = -\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle = -\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$, denn das Kreuzprodukt ändert sein Vorzeichen, wenn die Reihenfolge der Vektoren vertauscht wird (5.20). $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \det(\mathbf{z}, \mathbf{x}, \mathbf{y})$ wird durch unsere Überlegungen zum Volumen nahegelegt, muß aber streng genommen nachgerechnet werden, denn diese Gleichung stellt auch die Behauptung über ein gleiches Vorzeichen auf. Wir haben oben zwar die Rechte-Hand-Regel eingeführt und $\mathbf{x}, \mathbf{y}, \mathbf{z}$ sowie $\mathbf{z}, \mathbf{x}, \mathbf{y}$ gehorchen derselben Regel (ausprobieren), wir müssen uns aber immer noch davon überzeugen, daß sich die Ausdrücke $\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle$ und $\langle \mathbf{z} \times \mathbf{x} | \mathbf{y} \rangle$ auch tatsächlich daran halten.

Abb. 5.6 zeigt einen Spat mit positiver Orientierung. Die Vektoren $\mathbf{x}, \mathbf{y}, \mathbf{z}$ gehorchen der Rechte-Hand-Regel und das Skalarprodukt $\langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle$ ist positiv. Nach 5.9 muß also $\mathbf{x} \times \mathbf{y}$ und \mathbf{z} in dieselbe Richtung zeigen. Das bedeutet, daß \mathbf{x}, \mathbf{y} und $\mathbf{x} \times \mathbf{y}$ ebenfalls der Rechte-Hand-Regel genügen müssen.

5.1.14 Satz Das Spatprodukt $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \langle \mathbf{x} \times \mathbf{y} | \mathbf{z} \rangle$ ist, bis auf ein mögliches Vorzeichen, das Volumen des Spats, der von \mathbf{x}, \mathbf{y} und \mathbf{z} aufgespannt wird. Das Vorzeichen gibt die Orientierung der drei Vektoren an. Es gilt für alle $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w} \in \mathbb{R}^3$ und $t, s \in \mathbb{R}$:

$$\det \begin{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \end{bmatrix} = \begin{matrix} x_1 y_2 z_3 + x_3 y_1 z_2 + x_2 y_3 z_1 \\ - x_3 y_2 z_1 - x_1 y_3 z_2 - x_2 y_1 z_3, \end{matrix} \quad (5.28)$$

$$\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \det(\mathbf{z}, \mathbf{x}, \mathbf{y}) = \det(\mathbf{y}, \mathbf{z}, \mathbf{x}) = -\det(\mathbf{x}, \mathbf{z}, \mathbf{y}) = -\det(\mathbf{z}, \mathbf{y}, \mathbf{x}) = -\det(\mathbf{y}, \mathbf{x}, \mathbf{z}), \quad (5.29)$$

$$\det(\mathbf{x}, \mathbf{y}, t\mathbf{v} + s\mathbf{w}) = t \det(\mathbf{x}, \mathbf{y}, \mathbf{v}) + s \det(\mathbf{x}, \mathbf{y}, \mathbf{w}), \quad (5.30)$$

$$\det(\mathbf{x}, \mathbf{y}, \mathbf{z} + t\mathbf{x} + s\mathbf{y}) = \det(\mathbf{x}, \mathbf{y}, \mathbf{z}). \quad (5.31)$$

$\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$ gilt genau dann, wenn \mathbf{x}, \mathbf{y} und \mathbf{z} in einer Ebene durch den Ursprung liegen.

Beweis. Gleichung (5.28) ist simples Ausrechnen des Spatprodukts. Mit dieser Gleichung kann man sich dann leicht von (5.29) überzeugen. Da alle Faktoren x_i, y_j und z_k in den Produkten $x_i y_j z_k$ linear vorkommen (d. h. nicht mit höheren Potenzen als 1), ist (5.30) nur eine leichte Fleißaufgabe. Man sagt, die Determinante ist in der letzten Komponente linear. Für (5.31) genügt es daher, wenn wir uns mit $\det(\mathbf{x}, \mathbf{y}, \mathbf{z} + \mathbf{x})$ beschäftigen. Wir erhalten $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) +$

$\det(\mathbf{x}, \mathbf{y}, \mathbf{x})$. Für den zweiten Summanden gilt aber nach (5.29) $\det(\mathbf{x}, \mathbf{y}, \mathbf{x}) = -\det(\mathbf{x}, \mathbf{y}, \mathbf{x})$, woraus natürlich $\det(\mathbf{x}, \mathbf{y}, \mathbf{x}) = 0$ folgt. Also haben wir $\det(\mathbf{x}, \mathbf{y}, \mathbf{z} + \mathbf{x}) = \det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ gezeigt. Die letzte Behauptung ist im Lichte der Volumeninterpretation der Determinante klar. \square

Ist $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$, so nennt man $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *linear abhängig*, andernfalls *linear unabhängig*. Natürlich gelten Gleichungen wie (5.30) und (5.31) auch für die erste und die zweite Variable von $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$. Das ist eine direkte Konsequenz von (5.29).

Niemand berechnet eine Determinante gemäß (5.28). Genau wie beim Kreuzprodukt gibt es eine einfache grafische Merkmregel:

$$\begin{bmatrix} x_1 & y_1 & z_1 & x_1 & y_1 \\ x_2 & y_2 & z_2 & x_2 & y_2 \\ x_3 & y_3 & z_3 & x_3 & y_3 \end{bmatrix} : \quad \begin{aligned} & x_1 y_2 z_3 + y_1 z_2 x_3 + z_1 x_2 y_3 \\ & - x_3 y_2 z_1 - y_3 z_2 x_1 - z_3 x_2 y_1 = \det(\mathbf{x}, \mathbf{y}, \mathbf{z}). \end{aligned} \quad (5.32)$$

5.1.15 Die Determinante für den \mathbb{R}^2

Hier soll sie, wie im \mathbb{R}^3 , einen Inhalt messen, nur daß es sich jetzt um den Flächeninhalt eines Parallelogramms handelt, das von zwei Vektoren $\mathbf{x} = [x_1, x_2]$ und $\mathbf{y} = [y_1, y_2]$ aufgespannt wird. Mit einem kleinen Trick profitieren wir dabei von der Arbeit, die wir uns für den Fall \mathbb{R}^3 gemacht haben. Wir betten die Vektoren in den \mathbb{R}^3 ein, indem wir ihnen eine weitere Koordinate 0 hinzufügen: $\mathbf{x} = [x_1, x_2, 0]$ und $\mathbf{y} = [y_1, y_2, 0]$. Als dritten Vektor wählen wir $\mathbf{z} := [0, 0, 1]$ und erhalten durch den Aufspann von \mathbf{x} , \mathbf{y} und \mathbf{z} ein Prisma der Höhe 1 und dem Parallelogramm als Grundfläche. Sein Volumen $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ hat daher dieselbe Maßzahl wie der Flächeninhalt des Parallelogramms (Abb. 5.8). In (5.28) eingesetzt bleiben nur die Summanden übrig, die keinen Faktor z_1 oder z_2 aufweisen, also $x_1 y_2 - x_2 y_1$. Das ist die richtige Definition der Determinante im \mathbb{R}^2 :

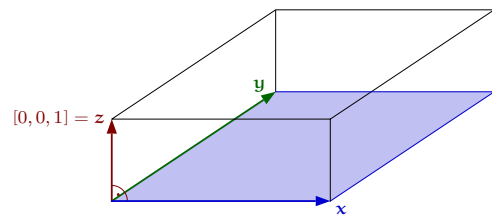


Abb. 5.8 Zur Determinante im \mathbb{R}^2

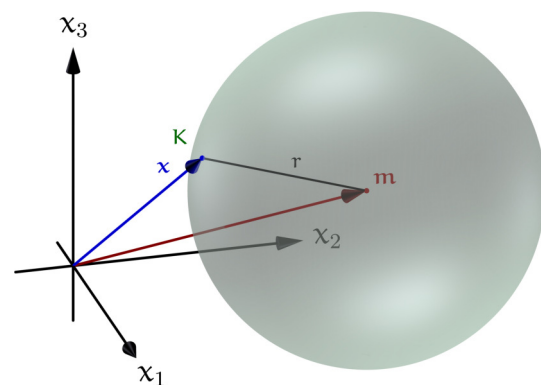
$$\det \left[\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right] := x_1 y_2 - x_2 y_1. \quad (5.33)$$

5.1.16 Die Kugel Eine Kugeloberfläche K ist der geometrische Ort aller Punkte \mathbf{x} , die von einem gegebenen Punkt \mathbf{m} – dem *Mittelpunkt* – einen festen Abstand r haben. r ist der *Radius* von K . Das bedeutet also

$$K = \{ \mathbf{x} \in \mathbb{R}^3 \mid \|\mathbf{x} - \mathbf{m}\| = r \}. \quad (5.34)$$

In der sogenannten *Koordinatenform* der Kugelgleichung verwendet man üblicherweise die quadrierte Version von $\|\mathbf{x} - \mathbf{m}\| = r$:

$$(x_1 - m_1)^2 + (x_2 - m_2)^2 + (x_3 - m_3)^2 = r^2. \quad (5.35)$$



5.1.17 A Zeigen Sie, daß die Tangentialebene T an die Kugel K mit Mittelpunkt \mathbf{m} und Radius r in einem Punkt $\mathbf{p} \in K$ durch die Menge

$$T = \{ \mathbf{x} \in \mathbb{R}^3 \mid \langle \mathbf{x} - \mathbf{m} \mid \mathbf{p} - \mathbf{m} \rangle = r^2 \} . \quad (5.36)$$

gegeben ist. Finden Sie eine Darstellung der Kugel (also nicht nur der Sphäre). Zeigen Sie für $\mathbf{m} := [3, 2, 2]$, $r := 1.5$ und $\mathbf{p} := [4, 1.5, 3]$, daß $\mathbf{x} := [-9, 213.5, 122]$ zu T gehört.

5.1.18 A $g := \{ \mathbf{p} + t \mathbf{u} \mid t \in \mathbb{R} \}$ sei eine Gerade mit einem Richtungsvektor \mathbf{u} der Länge 1. Machen Sie sich anhand der Abbildung 5.9 klar, daß der Abstand $d(\mathbf{q}, g)$ eines Punktes \mathbf{q} von g durch

$$d(\mathbf{q}, g) = \|\mathbf{u} \times (\mathbf{q} - \mathbf{p})\| \quad (5.37)$$

bestimmt ist. Berechnen Sie damit den Abstand von $\mathbf{q} := [15, 6, 5]$ zu der Geraden

$$g := \left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix} \mid t \in \mathbb{R} \right\} .$$

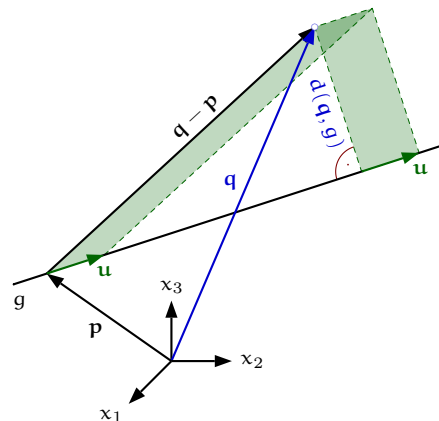


Abb. 5.9 Abstand Punkt–Gerade

5.1.19 Euklidisches Skalarprodukt und euklidische Norm im \mathbb{R}^n Für den euklidischen Raum \mathbb{R}^3 stehen uns zwei wichtige Werkzeuge zur Verfügung: Das Skalarprodukt zur Winkelmessung und die Norm zur Längen- und Abstandsmessung. Gleichung (5.12) zeigt, daß das erste das zweite schon bereitstellt. Daher müssen wir nur das Skalarprodukt auf den \mathbb{R}^n verallgemeinern und erhalten eine Norm gleich mitgeliefert. Im Hinblick auf (5.11) ist klar, wie das sinnvoll geschehen kann: Für $\mathbf{x} = [x_1, \dots, x_n] \in \mathbb{R}^n$ und $\mathbf{y} = [y_1, \dots, y_n] \in \mathbb{R}^n$ definieren wir das *euklidische Skalarprodukt* durch

$$\langle \mathbf{x} \mid \mathbf{y} \rangle := x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{k=1}^n x_k y_k . \quad (5.38)$$

Die zugehörige *euklidische Norm* ist dann

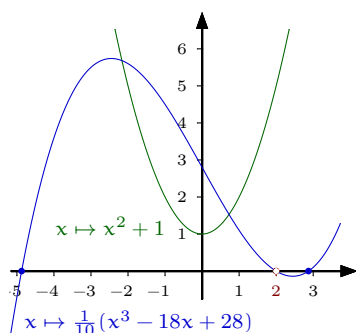
$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x} \mid \mathbf{x} \rangle} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\sum_{k=1}^n x_k^2} . \quad (5.39)$$

Das Augenmerk liegt im \mathbb{R}^n nicht mehr auf der Winkelmessung, sondern darauf, mit dem Skalarprodukt den Orthogonalitätsbegriff zur Verfügung zu haben:

$$\mathbf{x} \perp \mathbf{y} \iff \langle \mathbf{x} \mid \mathbf{y} \rangle = 0 . \quad (5.40)$$

Wenn nicht ausdrücklich etwas anderes gesagt wird, verstehen wir unter dem Skalarprodukt und der zugehörigen Norm auf \mathbb{R}^n immer das euklidische Skalarprodukt und die euklidische Norm.

5.2 Die komplexen Zahlen



Es gibt einfache algebraische Gleichungen, die in \mathbb{R} nicht lösbar sind. Die wichtigste Gleichung dieser Art, und wie wir sehen werden, auch die einzige, deren Lösung wir zu \mathbb{R} hinzufügen müssen, ist

$$x^2 = -1. \quad (5.41)$$

Da alle Quadrate in \mathbb{R} nicht negativ sind, kann es hier keine reelle Lösung geben. Der Graph der Funktion $x \mapsto x^2 + 1$ macht das ebenfalls deutlich. Eine Lösung von (5.41) wäre eine Nullstelle, die diese Funktion offensichtlich nicht hat.

Bei quadratischen Gleichungen $ax^2 + bx + c = 0$, mit reellen Koeffizienten a , b und c , für die $4ac > b^2$ gilt, versagt die Lösungsformel, die sogenannte *Mitternachtsformel* $x_{1/2} = \frac{1}{2a}(-b \pm \sqrt{b^2 - 4ac})$ regelmäßig, weil der Ausdruck $b^2 - 4ac$ unter der Wurzel negativ ist. Allerdings bedeutet das einfach, daß die Funktion $x \mapsto ax^2 + bx + c$ eben keine reelle Nullstelle hat. Anders sieht das bei kubischen Gleichungen, wie etwa $x^3 + px + q = 0$, aus. Auch für solche Gleichungen sind Lösungsformeln bekannt, wie die *Cardanischen Formeln* (siehe 5.2.11)

$$x = u - v, \quad \text{mit} \quad u := \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}}, \quad v := \sqrt[3]{\sqrt{\Delta} + \frac{q}{2}} = \frac{p}{3u}. \quad (5.42)$$

Dabei ist $\Delta := \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$ die sogenannte *Diskriminante*. Etwa für die Gleichung

$$x^3 - 18x + 28 = 0$$

ist $\Delta = -20$. Daher kann man (5.42) nicht anwenden, obwohl $x^3 - 18x + 28 = (x^2 + 2x - 14)(x - 2)$ zeigt, daß es eine glatte Lösung, nämlich $x_1 = 2$, geben muß (tatsächlich sind $x_{2/3} = -1 \pm \sqrt{15}$ zwei weitere). Die Formel (5.42) würde die unmöglichen Zwischenergebnisse $u = \sqrt[3]{2\sqrt{-5} - 14}$ und $v = \sqrt[3]{2\sqrt{-5} + 14}$ ergeben. Will man sich damit nicht abfinden, dann kann man ja mal sehen, was passiert, wenn man sich über das Denkverbot $\sqrt{-20} = 2\sqrt{-5}$ hinwegsetzt. Bei diesem Beispiel ist das besonders einfach, weil man weiß, daß $u = 1 - \sqrt{-5}$ die formale dritte Wurzel ist:

$$(1 - \sqrt{-5})^3 = 1 - 3\sqrt{-5} + 3\sqrt{-5}^2 - \sqrt{-5}^3 = 1 - 3\sqrt{-5} - 15 + 5\sqrt{-5} = 2\sqrt{-5} - 14.$$

Genauso erhält man $v = -1 - \sqrt{-5}$. Diese formalen Manipulationen liefern überraschenderweise tatsächlich ein richtiges Endergebnis $x = 1 - \sqrt{-5} + 1 + \sqrt{-5} = 2$. Solche Rechnungen kann man als einen der Gründe vermuten, warum man angefangen hat darüber nachzudenken, ob man nicht durch Hinzunahme fiktiver Lösungen von Gleichungen des Typs (5.41) zu \mathbb{R} einen Zahlenkörper erhalten kann, in dem Probleme, wie im vorgestellten Beispiel, nicht mehr auftreten. Dabei hat sich herausgestellt, daß die einzige *neue Zahl*, die benötigt wird, die Lösung von (5.41) ist. Diese Zahl wird mit i bezeichnet, was auf ihren imaginären Charakter hinweisen soll, wenn man sie vom Standpunkt reeller Zahlen aus betrachtet. Die Wurzeln u und v wären dann als $1 - \sqrt{5}i$ und $-1 - \sqrt{5}i$ zu interpretieren. Bevor wir die komplexen Zahlen formal einführen, überlegen wir uns, was wir benötigen, um konsistent mit ihnen rechnen

zu können. Dafür bilden wir zunächst einmal ganz naiv den allgemeinsten Ausdruck $z_1 + iz_2$, $z_1, z_2 \in \mathbb{R}$, den wir aus reellen Zahlen und i bilden können. Die scheinbar allgemeinere Form $z_0 + iz_1 + i^2 z_2 + \dots + i^n z_n$ läßt sich immer in $\tilde{z}_1 + i\tilde{z}_2$ verwandeln, wenn wir verwenden, was die Zahl i einzig ausmacht:

$$i^2 = -1. \quad (5.43)$$

Überlegen wir uns, welche Rechenregeln wir zu definieren haben. Für Addition, Subtraktion und Multiplikation sind sie offensichtlich, einzig bei der Division müssen wir zunächst das dritte Binom einsetzen, um den Bruch in die Standardform $u_1 + iu_2$ zu bringen:

$$\begin{aligned} (z_1 + iz_2) \pm (w_1 + iw_2) &= z_1 \pm w_1 + i(z_2 \pm w_2), \\ (z_1 + iz_2)(w_1 + iw_2) &= z_1 w_1 + i(z_1 w_2 + z_2 w_1) + i^2 z_2 w_2 \\ &= z_1 w_1 - z_2 w_2 + i(z_1 w_2 + z_2 w_1), \\ \frac{z_1 + iz_2}{w_1 + iw_2} &= \frac{(z_1 + iz_2)(w_1 - iw_2)}{(w_1 + iw_2)(w_1 - iw_2)} \\ &= \frac{z_1 w_1 - i^2 z_2 w_2 + i(z_2 w_1 - z_1 w_2)}{w_1^2 - i^2 w_2^2} \\ &= \frac{z_1 w_1 + z_2 w_2}{w_1^2 + w_2^2} + i \frac{z_2 w_1 - z_1 w_2}{w_1^2 + w_2^2}. \end{aligned}$$

5.2.1 Die Gaußsche Zahlenebene Um die komplexen Zahlen auf mathematisch solidem Grund einzuführen, überlegen wir uns, was in obiger Rechnung an Struktur wirklich benutzt wurde. Die erste Abstraktion folgt aus der Beobachtung, daß zur Festlegung von $z_1 + iz_2$ tatsächlich nur das Zahlenpaar $[z_1, z_2] \in \mathbb{R}^2$ nötig ist. Daher werden wir die komplexen Zahlen als Punktmenge mit der Ebene \mathbb{R}^2 (der sogenannten *Gaußschen Zahlenebene*) identifizieren und auf dieser die Rechenregeln einführen, die von unseren Vorüberlegungen nahegelegt werden.

5.2.2 Definition Die komplexen Zahlen \mathbb{C} sind die Menge \mathbb{R}^2 , versehen mit folgenden Rechenregeln

$$\begin{aligned} [z_1, z_2] \pm [w_1, w_2] &:= [z_1 \pm w_1, z_2 \pm w_2], \\ [z_1, z_2] \cdot [w_1, w_2] &:= [z_1 w_1 - z_2 w_2, z_1 w_2 + z_2 w_1]. \end{aligned}$$

Die Zahlen $[x, 0]$ werden mit den reellen Zahlen $x \in \mathbb{R}$ identifiziert. $[0, 1]$ wird mit dem Symbol i bezeichnet. Damit hat jede komplexe Zahl $z \in \mathbb{C}$ die Gestalt $z = z_1 + iz_2$, $z_1, z_2 \in \mathbb{R}$. Dabei heißen $\operatorname{Re}(z) := z_1$ und $\operatorname{Im}(z) := z_2$ Realteil bzw. Imaginärteil der Zahl z .

Die Addition und die Subtraktion in \mathbb{C} sind einfach die entsprechenden Vektoroperationen des \mathbb{R}^2 . Die Multiplikation ist eine Operation, die über die Vektorraumstruktur des \mathbb{R}^2 hinaus geht. Die Division haben wir bisher nicht definiert. Da sie die Umkehrung der Multiplikation darstellt, muß sie sich aus dieser Rechenregel gewinnen lassen.

5.2.3 A Zeigen Sie, daß die Multiplikation aus der Definition die üblichen Rechenregeln respektiert, d. h., daß $z \cdot w = w \cdot z$, $z \cdot (u + w) = z \cdot u + z \cdot w$, $z \cdot (u \cdot w) = (z \cdot u) \cdot w$ gilt.

5.2.4 Definition Für jede komplexe Zahl $z = z_1 + iz_2$ ist

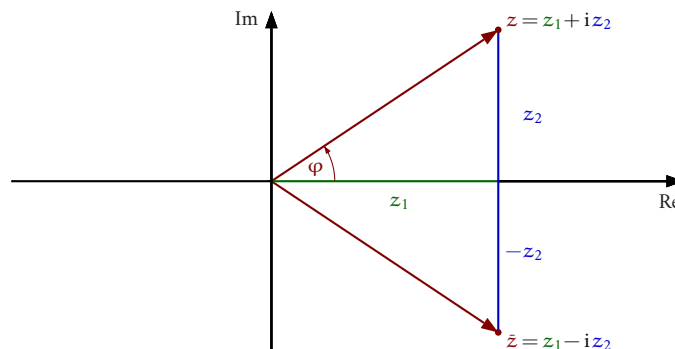
$$|z| := \sqrt{z_1^2 + z_2^2} \quad (5.44)$$

ihr Betrag und

$$\bar{z} := z_1 - iz_2 \quad (5.45)$$

die zu z konjugiert komplexe Zahl.

Der Betrag $|z|$ von z ist gerade die euklidische Norm des Vektors $[z_1, z_2] \in \mathbb{R}^2$, der zu z korrespondiert. Der Identifizierung von \mathbb{C} mit \mathbb{R}^2 , was die Vektorraumstruktur angeht, trägt auch die Darstellung komplexer Zahlen als Vektoren in der GAUSSschen Zahlenebene Rechnung:



5.2.5 A

- i) Zeigen Sie $z \cdot \bar{z} = |z|^2$ und $\overline{z \pm w} = \bar{z} \pm \bar{w}$, sowie $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$. Verwenden Sie das, um nachzuweisen, daß \mathbb{C} nullteilerfrei ist, d. h., daß aus $z \neq 0$ und $z \cdot w = 0$ immer $w = 0$ folgt.
- ii) Folgern Sie aus i), daß die Gleichung $z \cdot w = u$ für jedes $z \neq 0$ eine eindeutig bestimmte Lösung w hat und daß diese durch $w = \frac{1}{|z|^2} \bar{z} \cdot u$ gegeben ist. Folgern sie schließlich die Kürzungsregel für \mathbb{C} , indem Sie die Gleichung $z \cdot w = u$ mit einer Zahl $v \neq 0$ multiplizieren.

Die Lösung $w = \frac{1}{|z|^2} \bar{z}$ von $z \cdot w = 1$ bezeichnen wir mit $\frac{1}{z}$, $1/z$ oder z^{-1} .

- iii) Zeigen Sie: $\overline{z/w} = \bar{z}/\bar{w}$, $|\bar{z}| = |z|$,
 $\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$, $\operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z})$,
 $|z \cdot w| = |z| \cdot |w|$, $\left| \frac{z}{w} \right| = \frac{|z|}{|w|}$, $|z + w| \leq |z| + |w|$.

Verwenden Sie für die letzte Ungleichung, der sog. *Dreiecksungleichung*, daß aus $0 \leq a \leq b$ auch $0 \leq \sqrt{a} \leq \sqrt{b}$ folgt (warum eigentlich?).

Zeigen Sie die *umgekehrte Dreiecksungleichung* $|z - w| \geq ||z| - |w||$.

- iv) Zeigen Sie den *Zwei-Quadrate-Satz*: Das Produkt aus der Summe von jeweils zwei Quadraten ganzer Zahlen ist wieder eine solche Summe.
 $(a^2 + b^2)(c^2 + d^2) = e^2 + f^2$, $a, b, c, d, e, f \in \mathbb{Z}$.

Bemerkung: Ab jetzt schreiben wir, wie allgemein üblich, meist zw statt $z \cdot w$.

Aus der Skizze der GAUSSschen Zahlenebene auf Seite 78 entnimmt man leicht die sog. *Polar-darstellung* einer komplexen Zahl:

$$z = |z| (\cos(\varphi) + i \sin(\varphi)). \quad (5.46)$$

Der Winkel $\varphi \in (-\pi, \pi]$ ist eindeutig durch

$$\varphi = \begin{cases} \arccos\left(\frac{\operatorname{Re}(z)}{|z|}\right), & \operatorname{Im}(z) \geq 0, \\ -\arccos\left(\frac{\operatorname{Re}(z)}{|z|}\right), & \operatorname{Im}(z) < 0. \end{cases} \quad (5.47)$$

bestimmt. Dabei ist \arccos , oder \cos^{-1} im Moment einfach ein Name für die Möglichkeit, vom Wert des Kosinus auf den zugehörigen Winkel zu schließen (vergl. 11.4.7).

5.2.6 Die EULER-Formel Der Polardarstellung (5.46) ist nicht ohne Weiteres anzusehen, daß sie den eigentlichen Grund dafür liefert, daß das Rechnen mit komplexen Zahlen so leistungsfähig ist. Dazu fehlt noch eine berühmte Formel, nämlich die *EULER-Formel*, die wir hier im Vorgriff auf die erst später zu behandelnden Reihenentwicklung von Funktionen angeben wollen. Sie lautet (vergl. Beispiel 11.6.11)

$$e^{i\varphi} = \cos(\varphi) + i \sin(\varphi). \quad (5.48)$$

Mit dieser Formel ist eine Verbindung zwischen den trigonometrischen Funktionen und der Exponentialfunktion geschaffen, die doch scheinbar nichts miteinander zu tun haben, wenn man von den Funktionsgraphen ausgeht. Eine vorläufige Rechtfertigung dafür, den Ausdruck $\cos(\varphi) + i \sin(\varphi)$ mit der Exponentialfunktion in Verbindung zu bringen, liefern die *Additionssätze der Trigonometrie* (siehe auch S. 138, 276)

$$\cos(\varphi + \psi) = \cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi), \quad (5.49)$$

$$\sin(\varphi + \psi) = \sin(\varphi) \cos(\psi) + \cos(\varphi) \sin(\psi), \quad (5.50)$$

Damit erhalten wir nämlich (das zunächst noch formale) Potenzrechengesetz

$$\begin{aligned} e^{i\varphi} e^{i\psi} &= (\cos(\varphi) + i \sin(\varphi))(\cos(\psi) + i \sin(\psi)) \\ &= \cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi) + i(\sin(\varphi) \cos(\psi) + \cos(\varphi) \sin(\psi)) \\ &= \cos(\varphi + \psi) + i \sin(\varphi + \psi) = e^{i(\varphi + \psi)}. \end{aligned}$$

Die Polardarstellung einer komplexen Zahl lautet jetzt

$$z = |z| e^{i\varphi}. \quad (5.51)$$

Der Fortschritt dieser Formel gegenüber der normalen Darstellung $z = z_1 + i z_2$ liegt darin, daß die etwas umständlich auszuführenden Rechenoperationen der Multiplikation, Division und

insbesondere der Potenzierung auf die einfachen Potenzrechengesetze zurückgeführt werden. Für $z = |z| e^{i\varphi}$ und $w = |w| e^{i\psi}$ gilt

$$zw = |z||w| e^{i(\varphi+\psi)}, \quad (5.52)$$

$$\frac{z}{w} = \frac{|z|}{|w|} e^{i(\varphi-\psi)}, \quad (5.53)$$

$$z^n = |z|^n e^{in\varphi}. \quad (5.54)$$

An diesen Formeln lassen sich einfache geometrische Interpretationen der Multiplikation und der Division ablesen (für die Addition haben wir ja bereits die Vektoraddition). Besonders einfach werden sie am *Einheitskreis* $\mathbb{C}_1 := \{z \in \mathbb{C} \mid |z| = 1\}$. Hier bedeutet Multiplikation zw und Division $\frac{z}{w}$ einfach die Addition bzw. Subtraktion der Winkel φ und ψ der beteiligten Faktoren z und w . Außerhalb des Einheitskreises ist noch die Multiplikation bzw. Division der Beträge zu berücksichtigen.

5.2.7 Beispiel (HERONS Formel) Eine überraschende Anwendung des Rechnens mit komplexen Zahlen ist eine elegante Herleitung von HERONS Formel

$$F = \sqrt{s(s-a)(s-b)(s-c)} \quad (5.55)$$

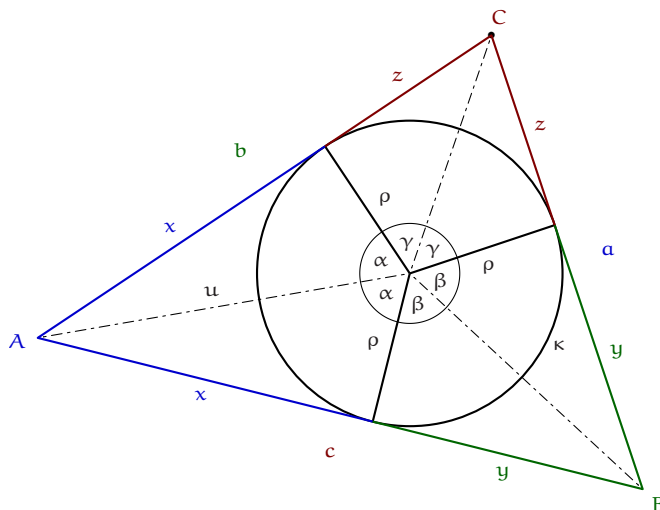


Abb. 5.10 Zu HERONS Formel

für den Flächeninhalt eines Dreiecks, bei gegebenen Seitenlängen a, b und c . Dabei ist $s := \frac{1}{2}(a+b+c) = x+y+z$ der halbe Umfang des Dreiecks (siehe Abb. 5.10).

ρ ist der Radius des *Inkreises* κ , also des größten Kreises, der ganz im Dreieck enthalten ist. Die Zahl $\rho + ix$ muß laut Skizze die Polardarstellung $u e^{i\alpha}$ haben, mit einem geeigneten $u > 0$.

Genauso erhalten wir zwei weitere Polardarstellungen:

$$\rho + iy = v e^{i\beta}, \quad \rho + iz = w e^{i\gamma}.$$

Das Produkt $uvw e^{i(\alpha+\beta+\gamma)}$ dieser drei Zahlen ist $-uvw$, also reell, denn $\alpha + \beta + \gamma = \pi$ und $e^{i\pi} = -1$. Sein Imaginärteil muß daher verschwinden:

$$\text{Im}(\rho + ix)(\rho + iy)(\rho + iz) = \rho^2(x+y+z) - xyz = \rho^2 s - xyz = 0.$$

Daraus läßt sich der *Inkreisradius* ρ bestimmen. Laut Skizze ist $y+z = a$, also $x+y+z = s = x+a$ und damit $x = s-a$. Genauso erhalten wir $y = s-b$ und $z = s-c$. Das ergibt

$$\rho = \sqrt{\frac{1}{s}(s-a)(s-b)(s-c)}. \quad (5.56)$$

Da sich das Dreieck in drei Rechtecke mit den Flächeninhalten ρx , ρy und ρz umwandeln läßt, gilt $F = \rho(x + y + z) = \rho s$. Daraus ist (5.55) leicht zu folgern.

Zur Bestimmung des Inkreismittelpunkts siehe Aufgabe 6.4.13.

Dieser Beweis stammt von Miles Dillon Edwards (Lassiter High School, Marietta, GA 30066).

Für ein Dreieck mit den Seiten $a = 2$, $b = 3$ und $c = 4$ ist $\rho = \frac{1}{6} \sqrt{15}$ und $F = \frac{3}{4} \sqrt{15}$.

5.2.8 A Zeigen Sie:

$$\begin{aligned} & e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = -1, \quad e^{2\pi i} = 1, \quad |e^{i\varphi}| = 1, \quad \overline{e^{i\varphi}} = e^{-i\varphi}. \\ \text{i)} \quad & e^{i\frac{\pi}{4}} = \frac{1}{\sqrt{2}}(1 + i), \quad e^{i\frac{\pi}{3}} = \frac{1}{2}(1 + \sqrt{3}i). \end{aligned}$$

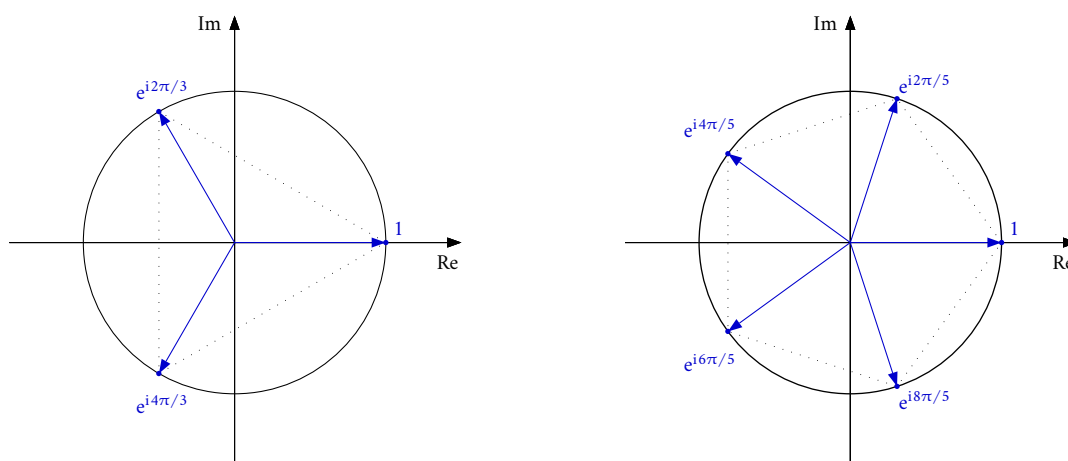
- ii) Multiplikation einer Zahl z mit $\pm i$ bedeutet eine Rotation von z um $\pm \frac{\pi}{2}$.
Multiplikation einer Zahl z mit $e^{i\psi}$ bedeutet eine Rotation von z um ψ .

5.2.9 Die n -ten Wurzeln Die Bildung von w^n für $w = |w|e^{i\psi}$ bedeutet, den Betrag $|w|$ zur n -ten Potenz zu erheben und den Winkel ψ mit n zu multiplizieren. Eine n -te Wurzel von z ist eine Zahl w , deren n -te Potenz w^n gerade $z = |z|e^{i\varphi}$ ergibt: $w^n = |w|^n e^{in\psi} = |z|e^{i\varphi} = z$. Der Betrag von w sollte also die n -te Wurzel von $|z|$ sein. Für ψ führt sicher die Wahl $\frac{\varphi}{n}$ zum Ziel:

$$w^n = \left(\sqrt[n]{|z|} e^{i\frac{\varphi}{n}} \right)^n = \sqrt[n]{|z|}^n e^{i\frac{\varphi}{n} \cdot n} = |z|e^{i\varphi} = z.$$

Da es aber schon in \mathbb{R} , z. B. für die Quadratwurzel, zwei Lösungen gibt, ist zu erwarten, daß weitere Lösungen der Gleichung $w^n = z$ existieren. Wir werden gleich sehen, daß es sogar immer genau n verschiedene Lösungen gibt. Der Grund dafür ist die Gleichung $e^{p \cdot 2\pi i} = 1$, $p \in \mathbb{Z}$, die einfach aus der Tatsache folgt, daß $\cos(p \cdot 2\pi) = 1$ und $\sin(p \cdot 2\pi) = 0$ gilt. Bevor wir die allgemeine Lösung der Gleichung $z^n = w$ angeben, beschäftigen wir uns mit dem wichtigen Spezialfall $u^n = 1$, der den Schlüssel zur allgemeinen Lösungstheorie liefert. Eine Lösung ist natürlich $u_0 := 1$. Eine weitere ist $u_1 := e^{i\frac{2\pi}{n}}$. Allerdings sind auch $u_p := e^{ip \cdot \frac{2\pi}{n}}$ Lösungen, denn $u_p^n = e^{ip \cdot \frac{2\pi}{n} \cdot n} = e^{ip \cdot 2\pi} = 1$. Nur für $p = 1, 2, \dots, n-1$ entstehen auf diese Weise neue Lösungen. u_0, u_1, \dots, u_{n-1} sind die sogenannten *n -ten Einheitswurzeln*.

In der folgenden Skizze sind die dritten und die fünften Einheitswurzeln zu sehen.



5.2.10 Satz Für $z = |z| e^{i\varphi}$ gibt es immer genau n verschiedene Lösungen w_0, w_1, \dots, w_{n-1} der Gleichung $w^n = z$, die aus der Grundlösung $w_0 := \sqrt[n]{|z|} e^{i\frac{\varphi}{n}}$ durch Multiplikation mit den n -ten Einheitswurzeln u_0, u_1, \dots, u_{n-1} entstehen ($p = 0, 1, \dots, n-1$):

$$w_p = \sqrt[n]{|z|} e^{i\frac{1}{n}(\varphi + p \cdot 2\pi)} = w_0 u_p, \quad (5.57)$$

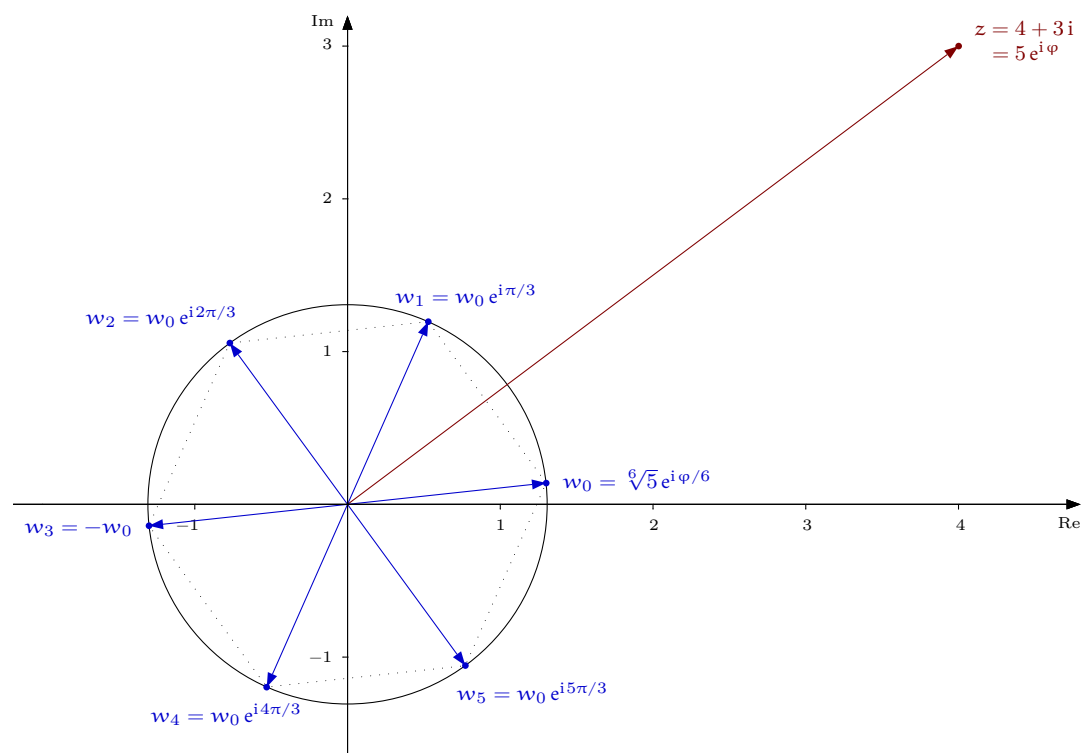
$$u_p = e^{ip \cdot \frac{2\pi}{n}}. \quad (5.58)$$

Beweis. Wir müssen nachrechnen, daß tatsächlich $w_p^n = z$ gilt. Das folgt aber leicht aus $w_0^n = z$ und $u_p^n = 1$.

Jetzt müssen wir nur noch zeigen, daß auch alle Lösungen gefunden wurden. Dazu überlegen wir zunächst, daß für ψ in der Gleichung $e^{i\psi} = 1$ nur die Zahlen $p \cdot 2\pi$, $p \in \mathbb{Z}$ möglich sind (eine einfache Übung). Für eine Lösung $w = |w| e^{i\psi}$ von $w^n = z$ muß $|w|^n = |z|$, also $|w| = \sqrt[n]{|z|}$ gelten. $w^n = z$ führt dann auf $e^{in\psi} = e^{i\varphi}$ und nach Multiplikation mit $e^{-i\varphi}$ auf $e^{i(n\psi - \varphi)} = 1$. Das bedeutet $n\psi - \varphi = p \cdot 2\pi$, oder $\psi = \frac{1}{n}(\varphi + p \cdot 2\pi)$, $p \in \mathbb{Z}$. Einzig die Einschränkung von p auf den Bereich von 0 bis $n-1$ muß noch begründet werden. Würden wir $p > n-1$ zulassen, also $p = r \cdot n + q$ mit einer Zahl $r \in \mathbb{N}$ und $q \in \{0, \dots, n-1\}$, so erhielten wir keine neue Lösung, sondern einfach wieder w_q aus $\{w_0, w_1, \dots, w_{n-1}\}$:

$$\sqrt[n]{|z|} e^{i(\varphi + (r \cdot n + q) \cdot 2\pi)/n} = \sqrt[n]{|z|} e^{i(\varphi + q \cdot 2\pi)/n} \cdot e^{ir \cdot n \cdot 2\pi/n} = w_q \cdot e^{ir \cdot 2\pi} = w_q.$$

Genauso überlegt man sich, daß auch $p < 0$ zu keiner neuen Lösung führt. □



Bestimmen wir die sechsten Wurzeln von $z = 4 + 3i = 5 e^{i\varphi}$ mit $\varphi = \arccos(0.8) \approx 0.6435$: $w_0 = \sqrt[6]{5} e^{i\frac{\varphi}{6}}$ und

$$w_1 = w_0 e^{i\frac{\pi}{3}}, \quad w_2 = w_0 e^{i\frac{2\pi}{3}}, \quad w_3 = -w_0, \quad w_4 = w_0 e^{i\frac{4\pi}{3}}, \quad w_5 = w_0 e^{i\frac{5\pi}{3}}.$$

Wie man sieht entstehen auch im Falle allgemeiner Wurzeln regelmäßige n -Ecke, nur daß diese jetzt um den Winkel von w_0 gedreht sind.

5.2.11 Kubische Gleichungen Eine kubische Gleichung $x^3 + bx^2 + cx + d = 0$ kann durch die Substitution $y := x - e$ in die Form

$$y^3 + py + q = 0 \tag{5.59}$$

gebracht werden, mit $p := c - \frac{1}{3}b^2$ und $q := \frac{2}{27}b^3 - \frac{1}{3}bc + d$. Für $e := -\frac{b}{3}$ erhält man nämlich

$$\begin{aligned} 0 &= x^3 + bx^2 + cx + d \\ &= y^3 + 3ey^2 + 3e^2y + e^3 + b(y^2 + 2ey + e^2) + c(y + e) + d \\ &= y^3 + (3e + b)y^2 + (3e^2 + 2eb + c)y + e^3 + e^2b + ce + d \\ &= y^3 + \left(\frac{b^2}{3} - \frac{2b^2}{3} + c\right)y - \frac{b^3}{27} + \frac{b^3}{9} - \frac{bc}{3} + d \\ &= y^3 + \left(c - \frac{1}{3}b^2\right)y + \frac{2}{27}b^3 - \frac{1}{3}bc + d. \end{aligned}$$

Der Lösungsansatz für (5.59) ist $y = u - v$, mit den beiden Unbekannten u und v :

$$\begin{aligned} 0 &= u^3 - 3u^2v + 3uv^2 - v^3 + (u - v)p + q \\ &= u^3 - v^3 + 3uv(v - u) + (u - v)p + q = u^3 - v^3 + (p - 3uv)(u - v) + q \\ &= u^3 - v^3 + q, \quad \text{falls } p = 3uv \text{ gilt.} \end{aligned}$$

Wenn also u und v so bestimmt werden können, daß die Gleichungen

$$0 = u^3 - v^3 + q, \tag{5.60}$$

$$p = 3uv \tag{5.61}$$

gelten, dann ist $y = u - v$ eine Lösung von (5.59). Multiplizieren wir (5.60) mit u^3 , so erhalten wir eine quadratische Gleichung für u^3 :

$$0 = u^6 + qu^3 - \left(\frac{p}{3}\right)^3,$$

mit den Lösungen

$$u^3 = -\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3},$$

die, wie immer bei den Lösungen quadratischer Gleichungen, richtig zu interpretieren sind, falls die sogenannte *Diskriminante* $\Delta := \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$ negativ, oder gar komplex sein sollte.

Wir wählen

$$u := \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} \quad \text{und} \quad v := \sqrt[3]{\sqrt{\Delta} + \frac{q}{2}}.$$

Dann gilt offensichtlich $u^3 - v^3 = -q$ und $uv = \sqrt[3]{\Delta - \frac{q^2}{4}} = \sqrt[3]{\frac{p^3}{27}} = \frac{p}{3}$ – letzteres allerdings nur, wenn die Wurzeln reell sind. Die Gleichungen (5.60) und (5.61) sind erfüllt. Damit ist durch

$$y_1 := \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \sqrt[3]{\sqrt{\Delta} + \frac{q}{2}} = \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \frac{p}{3\sqrt[3]{\sqrt{\Delta} - \frac{q}{2}}} \quad (5.62)$$

eine Formel für eine Lösung der kubischen Gleichung gefunden. Auch hier muß gesagt werden, daß es sich im Falle komplexer Wurzel $\sqrt{\Delta}$ weniger um eine Lösungsformel handelt, als vielmehr um eine Anweisung, aus welcher komplexen Zahl eine dritte Wurzel gezogen werden soll. Deshalb ist der zweite Teil von (5.62) dem ersten vorzuziehen, wenn in den Wurzeln komplexe Zahlen vorkommen, da sonst nicht klar ist, welche dritte Wurzel von $\sqrt{\Delta} + \frac{q}{2}$ gemeint ist. Es genügt, eine Lösung in der Form $y = u - v$ zu finden, da die beiden anderen systematisch mit Hilfe der dritten Einheitswurzeln (5.58) $u_k := \exp(k \frac{2\pi i}{3})$, $k = 0, 1, 2$, konstruiert werden können:

$$y_2 = u_1 u - u_2 v, \quad y_3 = u_2 u - u_1 v. \quad (5.63)$$

Für y_2 etwa gilt $3u_1 u \cdot u_2 v = 3uv = p$ und $(u_1 u)^3 - (u_2 v)^3 = u^3 - v^3 = -q$.

5.2.12 Satz Eine kubische Gleichung $x^3 + bx^2 + cx + d = 0$ kann mit der Substitution $x = y - \frac{b}{3}$ auf die Standardform

$$y^3 + py + q = 0 \quad (5.64)$$

gebracht werden. Dabei ist $p := c - \frac{1}{3}b^2$ und $q := \frac{2}{27}b^3 - \frac{1}{3}bc + d$. Deren Lösungen werden (für $p, q \neq 0$) durch die sogenannten Cardanischen Formeln gegeben:

$$y_1 = \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \frac{p}{3\sqrt[3]{\sqrt{\Delta} - \frac{q}{2}}} = u - v, \quad (5.65)$$

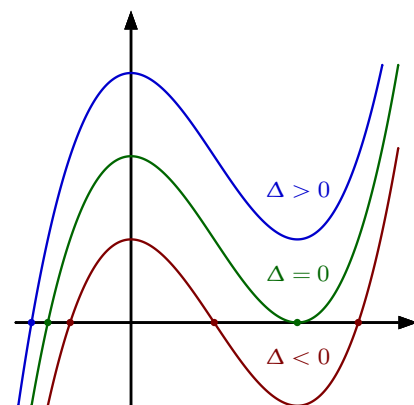
$$y_2 = u_1 u - u_2 v, \quad y_3 = u_2 u - u_1 v. \quad (5.66)$$

$u_0 = 1$, $u_{1/2} = \frac{1}{2}(-1 \pm i\sqrt{3})$ sind die dritten Einheitswurzeln. $\Delta := (\frac{q}{2})^2 + (\frac{p}{3})^3$ ist die sogenannte Diskriminante. Für reelle p und q bedeutet

$$\Delta > 0: \quad y_1 = \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \sqrt[3]{\sqrt{\Delta} + \frac{q}{2}} \in \mathbb{R}, \\ y_3 = y_2 \notin \mathbb{R} \text{ und } y_2 \neq y_3,$$

$$\Delta = 0: \quad y_1 = \frac{3q}{p} \in \mathbb{R} \text{ und } y_2 = y_3 = -\frac{3q}{2p} \in \mathbb{R},$$

$$\Delta < 0: \quad y_{1/2/3} = 2 \operatorname{Re}\left(u_{0/1/2} \sqrt[3]{-\frac{q}{2} + i\sqrt{|\Delta|}}\right), \text{ oder} \\ y_1 = 2\sqrt{\frac{|p|}{3}} \cos(\varphi_0), \\ y_{2/3} = -2\sqrt{\frac{|p|}{3}} \cos(\varphi_0 \mp \frac{\pi}{3}), \\ \varphi_0 := \frac{1}{3} \arccos\left(-\frac{q}{2}\sqrt{\frac{27}{|p|^3}}\right).$$



Beweis. $p \neq 0$ bedeutet, daß $\sqrt{\Delta} \neq \frac{q}{2}$ gilt, so daß (5.65) sinnvoll ist. Wir müssen zeigen, daß $y_1 = u - v$, mit $u = \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}}$ und $v = \frac{p}{3\sqrt[3]{\sqrt{\Delta} - \frac{q}{2}}}$ die beiden Gleichungen (5.60) und (5.61) erfüllen. Letztere ist offensichtlich. Bleibt (5.60):

$$u^3 - v^3 = \sqrt{\Delta} - \frac{q}{2} - \frac{\left(\frac{p}{3}\right)^3}{\sqrt{\Delta} - \frac{q}{2}} = \frac{\Delta + \left(\frac{q}{2}\right)^2 - q\sqrt{\Delta} - \left(\frac{p}{3}\right)^3}{\sqrt{\Delta} - \frac{q}{2}} = \frac{\frac{q^2}{2} - q\sqrt{\Delta}}{\sqrt{\Delta} - \frac{q}{2}} = -q.$$

Dabei haben wir nur die definierende Eigenschaft der beteiligten Wurzeln verwendet, also $u^3 = \sqrt{\Delta} - \frac{q}{2}$ und $\sqrt{\Delta}^2 = \Delta$. Wir können daher eine beliebige Quadratwurzel aus der (möglicherweise komplexen) Zahl Δ und eine beliebige dritte Wurzel aus $\sqrt{\Delta} - \frac{q}{2}$ wählen, um die erste Lösung y_1 zu bilden. Daß es sich dann bei y_2 und y_3 ebenfalls um Lösungen handelt, haben wir oben schon gezeigt.

Jetzt noch der Fall reeller Koeffizienten $p \neq 0$ und $q \neq 0$ und die Rolle der Diskriminante. Für $\Delta \geq 0$ haben wir es nur mit der dritten Wurzel aus reellen Zahlen zu tun, so daß für y_1 die reelle Wurzel gewählt werden kann, für die die gewohnten Rechenregeln gelten:

$$\begin{aligned} y_1 &= \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \frac{p}{3\sqrt[3]{\sqrt{\Delta} - \frac{q}{2}}} = \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \frac{p\sqrt[3]{\sqrt{\Delta} + \frac{q}{2}}}{3\sqrt[3]{\left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2 - \left(\frac{q}{2}\right)^2}} \\ &= \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \sqrt[3]{\sqrt{\Delta} + \frac{q}{2}} = u - v. \end{aligned}$$

Offensichtlich sind u und v verschieden und daher $y_1 \neq 0$. Wäre $y_2 = u_1 u - u_2 v = u_1 u - \overline{u_1} v$ reell, so müsste $u_1 u - \overline{u_1} v = \overline{u_1} u - u_1 v$ gelten, woraus sofort $2i \operatorname{Im}(u_1)u = -2i \operatorname{Im}(u_1)v$, also $u = -v$ folgen würde. Das wiederum wäre zu $\Delta = 0$ äquivalent. Nur in diesem Fall sind also die weiteren Lösungen y_2 und y_3 reell und darüber hinaus gleich: $v = -u$ hat $y_1 = -2\sqrt[3]{\frac{q}{2}}$ und $y_2 = (u_1 + \overline{u_1})u = 2 \operatorname{Re}(u_1)u = \sqrt[3]{\frac{q}{2}} = y_3$ zur Folge. $\Delta = 0$ bedeutet $\left(\frac{p}{3}\right)^3 = -\left(\frac{q}{2}\right)^2$, also $\left(\frac{p}{3}\right)^3 \frac{q}{2} = -\left(\frac{q}{2}\right)^3$, oder $-\frac{q}{2} = \left(\frac{3q}{2p}\right)^3$, woraus $y_1 = \frac{3q}{p}$ und $y_2 = -\frac{3q}{2p}$ unmittelbar folgt.

Für $\Delta > 0$ dagegen ist $y_2 \notin \mathbb{R}$ und daher $y_2 \neq \overline{y_2} = y_3$. Damit sind drei verschiedene Lösungen vorhanden, von denen genau eine reell ist.

Für $\Delta < 0$ wählen wir $\sqrt{\Delta} := i\sqrt{|\Delta|}$ und erhalten mit $u = \sqrt[3]{-\frac{q}{2} + i\sqrt{|\Delta|}}$ die erste Lösung $y_1 = u - \frac{p}{3u}$. Die formen wir um (vergl. Aufgabe 5.2.15):

$$\begin{aligned} y_1 &= u - \frac{p\overline{u}}{3|u|^2} = u - \frac{p\overline{u}}{3\sqrt[3]{\left|-\frac{q}{2} + i\sqrt{|\Delta|}\right|^2}} = u - \frac{p\overline{u}}{3\sqrt[3]{\left|-\frac{q}{2} + i\sqrt{|\Delta|}\right|^2}} \\ &= u - \frac{p\overline{u}}{3\sqrt[3]{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3 - \left(\frac{q}{2}\right)^2}} = u + \overline{u} = 2 \operatorname{Re} \sqrt[3]{-\frac{q}{2} + i\sqrt{|\Delta|}}. \end{aligned}$$

Die Formeln für y_2 und y_3 ergeben sich jetzt mit $v = -\overline{u}$ aus einer einfachen Rechnung. u , $u_1 u$ und $u_2 u$ sind alle dritten Wurzeln von $-\frac{q}{2} + i\sqrt{|\Delta|}$. Sie bilden also die Eckpunkte eines

gleichseitigen Dreiecks. Wegen $u \notin \mathbb{R}$ liegt es nicht symmetrisch zur reellen Achse, weshalb die Realteile der Eckpunkte y_1 , y_2 und y_3 paarweise verschieden sein müssen.

Die zweite Form der Lösungen ergibt sich mit Hilfe des Ansatzes $y = t \cos(\varphi)$ und der Gleichung $\cos(3\varphi) = 4 \cos^3(\varphi) - 3 \cos(\varphi)$ (vergl. Aufgabe 5.2.19). Setzen wir das in (5.64) ein, um φ und t zu bestimmen:

$$0 = t^3 \cos^3(\varphi) + p t \cos(\varphi) + q = \frac{1}{4} t^3 \cos(3\varphi) + t \left(\frac{3}{4} t^2 + p \right) \cos(\varphi) + q.$$

Wenn wir $t = \sqrt{-\frac{4}{3}p}$ wählen, verschwindet der Term mit $\cos(\varphi)$. Das ist möglich, denn $\Delta = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2 < 0$ hat $p < 0$ zur Folge. Damit führt $t = 2\sqrt{\frac{|p|}{3}}$ auf $\cos(3\varphi) = -\frac{q}{2} \sqrt{\left(\frac{3}{|p|}\right)^3}$. Auch diese Gleichung ist lösbar, denn $\Delta < 0$ bedeutet auch $\left|\left(\frac{p}{3}\right)^3\right| = \left(\frac{|p|}{3}\right)^3 > \left(\frac{q}{2}\right)^2$, also $\frac{|q|}{2} < \sqrt{\left(\frac{|p|}{3}\right)^3}$ und $\frac{|q|}{2} \sqrt{\left(\frac{3}{|p|}\right)^3} < 1$. Die Lösungen sind $\varphi_k := \frac{1}{3} \arccos\left(-\frac{q}{2} \sqrt{\left(\frac{3}{|p|}\right)^3}\right) + k \frac{2\pi}{3}$, $k = 0, 1, 2$. Für $k = 0$ erhalten wir $y_1 := 2\sqrt{\frac{|p|}{3}} \cos(\varphi_0)$. Für $k = 1$ ergibt sich, wegen $\cos(\varphi_0 + \frac{2\pi}{3}) = \cos(\varphi_0 - \frac{\pi}{3} + \pi) = -\cos(\varphi_0 - \frac{\pi}{3})$ und für $k = 2$, wegen $\cos(\varphi_0 + \frac{4\pi}{3}) = \cos(\varphi_0 + \frac{\pi}{3} + \pi) = -\cos(\varphi_0 + \frac{\pi}{3})$ die behaupteten Ausdrücke für y_2 und y_3 . \square

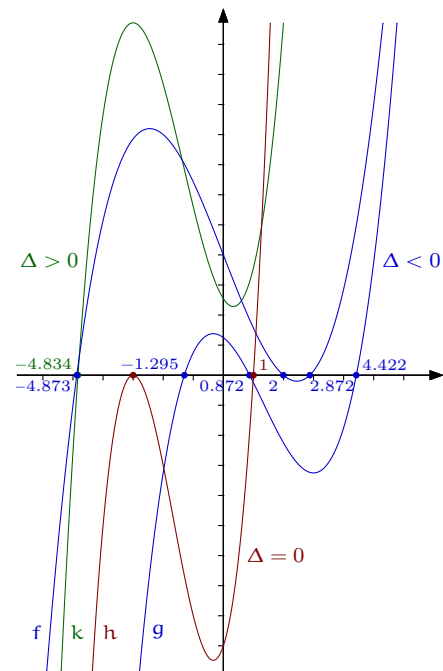
5.2.13 Beispiel

1. Für die Nullstellen der Funktion $f(x) := x^3 - 18x + 28$ ist $\Delta = -20 < 0$. Daher ist u eine dritte Wurzel von $-14 + 2i\sqrt{5}$, etwa $u = 1 - i\sqrt{5}$:

$$(1 - i\sqrt{5})^3 = 1 - 3i\sqrt{5} - 15 + 5i\sqrt{5} = -14 + 2i\sqrt{5}.$$

Die Nullstellen von f sind dann $x_1 = 2 \operatorname{Re}(u) = 2$ und $x_{2/3} = 2 \operatorname{Re}(u_{1/2} u) = \operatorname{Re}(-1 \pm i\sqrt{5} \pm \sqrt{15} \pm i\sqrt{3}) = -1 \pm \sqrt{15}$, in Übereinstimmung mit den Lösungen der quadratischen Gleichung $x^2 + 2x - 14 = 0$, die aus der kubischen durch Division mit dem Linearfaktor $x - 2$ entsteht.

Für den (seltenen) Fall, daß man die dritte Wurzel geschlossen angeben kann, ist die erste Version der Lösungsformeln für $\Delta < 0$ von Vorteil. Normalerweise aber findet man keine algebraische Ausdrücke für die Wurzeln. Dann ist die \cos -Version der Lösung vorzuziehen. Für die Nullstellen von $g(x) := \frac{1}{4}(x^3 - 4x^2 - 3x + 5)$ etwa, führt die Substitution $x = y + \frac{4}{3}$ auf die Normalform $y^3 - \frac{25}{3}y - \frac{101}{27} = 0$, mit $\Delta = -\frac{13 \cdot 149}{108} < 0$. Wir erhalten $\varphi_0 = \frac{1}{3} \arccos\left(\frac{101}{250}\right)$ und daraus $x_1 = \frac{10}{3} \cos\left(\frac{1}{3} \arccos\left(\frac{101}{250}\right)\right) + \frac{4}{3} \approx 4.422697$, $x_2 = -\frac{10}{3} \cos\left(\frac{1}{3} \arccos\left(\frac{101}{250}\right) - \frac{\pi}{3}\right) + \frac{4}{3} \approx -1.295415$ und $x_3 \approx 0.872717$.



2. $h(x) := x^3 + 5x^2 + 3x - 9 = 0$ führt mit der Substitution $x = y - \frac{5}{3}$ auf $y^3 - \frac{16}{3}y - \frac{128}{27} = 0$, mit $\Delta = -\left(\frac{16}{3}\right)^3 + \left(\frac{64}{27}\right)^2 = -\frac{16^3}{27^2} + \frac{64^2}{27^2} = 0$. Daher ist $y_1 = \frac{3 \cdot 128 \cdot 3}{27 \cdot 16} = \frac{8}{3}$, also $x_1 = \frac{8}{3} - \frac{5}{3} = 1$ und $y_{2/3} = -\frac{4}{3}$, d. h., $x_{2/3} = -3$. Das bedeutet $h(x) = (x - 1)(x + 3)^2$ (was man mit der systematischen Methode zum Auffinden rationaler Lösungen 11.8.7, bei diesem Beispiel, viel schneller erhalten hätte).

3. Für die Funktion $k(x) := \frac{27}{53}(x^3 + 4x^2 - 3x + 5)$ liefert die Substitution $x = y - \frac{4}{3}$ die Normalform $y^3 - \frac{25}{3}y + \frac{371}{27} = 0$, mit $\Delta = \frac{11^2 \cdot 23}{6^2 \cdot 3} > 0$. Das ergibt $u = \sqrt[3]{\frac{11}{6} \sqrt{\frac{23}{3}} - \frac{371}{54}} \approx -1.215110$ und $v = \sqrt[3]{\frac{11}{6} \sqrt{\frac{23}{3}} + \frac{371}{54}} \approx 2.286030$. Wir erhalten damit $x_1 \approx -3.501140 - \frac{4}{3} \approx -4.834473$ und $x_{2/3} = \frac{1}{2}(v - u) \pm i \frac{\sqrt{3}}{2}(u + v) - \frac{4}{3} \approx -0.417237 \pm 0.927444 i$.

4. $\ell(x) := x^3 + 6x^2 + 3(4+i)x + 9 + 7i$ führt mit der Substitution $x = y - 2$ auf $p = 3(4+i) - 12 = 3i$, $q = \frac{16 \cdot 27}{27} - 6(4+i) + 9 + 7i = 1 + i$ und $\Delta = \frac{1}{4}(1+i)^2 + i^3 = \frac{1}{2}i - i = -\frac{1}{2}i$. $y^3 + 3iy + 1 + i = 0$ ist die zu $\ell(x) = 0$ gehörende Normalform. Mit (5.67) erhalten wir $\sqrt{\Delta} = \frac{1}{2}(1-i)$ als eine mögliche Quadratwurzel. Daraus ergibt sich $u = \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} = \sqrt[3]{\frac{1}{2}(1-i) - \frac{1}{2}(1+i)} = \sqrt[3]{-i} = i$ und $v = 1$, also $x_1 = i - 3$. Die verbleibenden Lösungen von $\ell(x) = 0$ sind $x_2 = u_1 i - u_2 - 2 = -\frac{1}{2}(3 + \sqrt{3}) + \frac{1}{2}(\sqrt{3} - 1)i$ und $x_3 = u_2 i - u_1 - 2 = -\frac{1}{2}(3 - \sqrt{3}) - \frac{1}{2}(1 + \sqrt{3})i$.

5. $x^3 + 3ix + 2 + i = 0$, also $p = 3i$ und $q = 2 + i$, sowie $\Delta = \frac{3}{4}$. Wir benötigen für u eine dritte Wurzel von $z := -\frac{q}{2} + \sqrt{\Delta} = \frac{1}{2}(\sqrt{3} - 2) - \frac{1}{2}i = |z|e^{i\varphi}$. Mit $|z| = \sqrt{2 - \sqrt{3}} = \frac{1}{\sqrt{2}}(\sqrt{3} - 1)$ ist $\frac{\varphi}{3} = -\frac{1}{3}\arccos\left(\sqrt{2} \frac{\sqrt{3}-2}{\sqrt{3}-1}\right) = -\frac{1}{3}\arccos\left(\frac{\sqrt{2}}{4}(1 - \sqrt{3})\right) \approx -0.610865$. Wir erhalten $u = \sqrt[3]{|z|}e^{i\frac{\varphi}{3}} \approx 0.802926(\cos(0.610865) - i\sin(0.610865)) \approx 0.657719 - 0.460539i$ und $v = \frac{p}{3u} \approx -0.714357 + 1.020209i$, so daß schließlich $x_1 = u - v \approx 1.372076 - 1.480748i$ und $x_2 \approx -1.170726 + 0.691323i$ sowie $x_3 \approx -0.201350 + 0.789425i$ als Ergebnis folgt.

Dieses Beispiel zeigt, worin die Schwierigkeiten in der Anwendung der Cardanischen Formeln bestehen: Ist der Ausdruck $z = a + ib$ in der dritten Wurzel komplex, so ist eben eine dritte Wurzel aus einer komplexen Zahl zu finden. Das wird nicht immer in der Form eines geschlossenen Wurzelausdrucks, wie in Beispiel 4, möglich sein können. Vielmehr wird man sich über die Polardarstellung $z = |z|e^{i\varphi}$ mit einer Wurzel in der Form $\sqrt[3]{|z|}(\cos(\frac{\varphi}{3}) + i\sin(\frac{\varphi}{3}))$ behelfen müssen, in der man die trigonometrischen Funktionen im Allgemeinen nicht loswerden kann. Im Gegensatz zu Quadratwurzeln, wo man über $\cos(\frac{\varphi}{2}) = \sqrt{\frac{1}{2}(1 + \cos(\varphi))}$ und $\sin(\frac{\varphi}{2}) = \sqrt{\frac{1}{2}(1 - \cos(\varphi))}$ die Berechnung der Halbwinkelfunktionen auf den Kosinus zurückspielen kann, für den man über $\cos(\varphi) = \pm \frac{a}{\sqrt{a^2 + b^2}}$ einen algebraischen Ausdruck gewinnt (vergl. Aufgabe 5.2.18), ist das für $\cos(\frac{\varphi}{3})$ und $\sin(\frac{\varphi}{3})$ offenbar nicht möglich. Das liegt daran, daß $\cos(3\varphi) = 4\cos^3(\varphi) - 3\cos(\varphi)$ gilt, so daß $4\cos^3(\frac{\varphi}{3}) - 3\cos(\frac{\varphi}{3}) - \cos(\varphi) = 0$ die Lösungen einer kubischen Gleichung erfordert, um $\cos(\frac{\varphi}{3})$ als Funktion von $\cos(\varphi)$ darzustellen. Man könnte auf die Idee kommen, dafür (5.62) zu verwenden. Aber dabei dreht man sich im Kreis, denn der Ausdruck unter der dritten Wurzel ergibt genau $\cos(\varphi) \pm i\sin(\varphi) = e^{\pm i\varphi}$.

5.2.14 A Bestimmen Sie die 8-ten Wurzeln von $z = 5 + 12i$.

5.2.15 A Zeigen Sie: Ist v eine n -te Wurzel von x und w eine n -te Wurzel von z , dann ist vw eine n -te Wurzel von xz . Wenn wir v durch $\sqrt[n]{x}$ und w durch $\sqrt[n]{z}$ bezeichnen würden, ließe sich das durch $\sqrt[n]{x}\sqrt[n]{z} = \sqrt[n]{xz}$ wiedergeben. Machen Sie sich an folgender Rechnung klar, daß diese vermeintliche Gleichung für komplexe Zahlen im Allgemeinen *falsch* ist:

$$-1 = i^2 = \sqrt{-1}^2 = \sqrt{-1} \sqrt{-1} = \sqrt{(-1) \cdot (-1)} = \sqrt{1} = 1.$$

Wo liegt der Fehler in dieser Überlegung?

Zeigen Sie: Für den Betrag der n -ten Wurzel v von x gilt $|v| = \sqrt[n]{|x|}$.

5.2.16 A Zeigen Sie, daß der Ausdruck

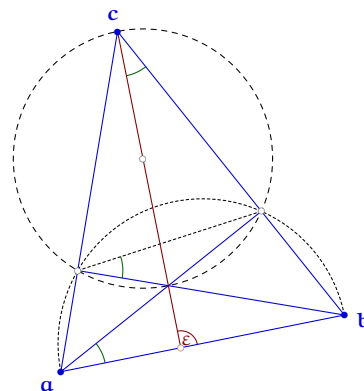
$$\frac{(e^{i\alpha} - e^{i\beta})^2 |e^{-i\alpha} - e^{i\beta}|^2}{|e^{i\alpha} - e^{i\beta}|^2 (e^{-i\alpha} - e^{i\beta})^2}$$

für $|\beta| \neq |\alpha|$ und $\alpha, \beta \in [-\pi, \pi]$ nicht von β abhängt. Geben Sie eine Interpretation dieses Ergebnisses an.

5.2.17 A

Verwenden Sie die Interpretation von Übung 5.2.16 und ergänzen Sie damit die Skizze für einen kurzen Beweis des Höhensatzes in Dreiecken:

In einem Dreieck schneiden sich die drei Höhen in einem gemeinsamen Punkt.



5.2.18 A Zeigen Sie, daß durch

$$\sqrt{z} := \sqrt{\frac{1}{2}(|z| + x)} + i v(y) \sqrt{\frac{1}{2}(|z| - x)} \quad (5.67)$$

eine Quadratwurzel der Zahl $z := x + iy$ gegeben ist. Eine weitere ist offensichtlich $-\sqrt{z}$. Berechnen Sie damit $\sqrt{12 - 5i}$ und $\sqrt{-4 - 5i}$.

Dabei ist $v(y) := 1$, falls $y \geq 0$ und $v(y) := -1$, falls $y < 0$ gilt. Diese Funktion unterscheidet sich von der sogenannten *Vorzeichenfunktion* sgn (auch als *Signum* bezeichnet) an der Stelle $y = 0$:

$$\text{sgn}(y) := \begin{cases} 1 & , y > 0, \\ 0 & , y = 0, \\ -1 & , y < 0. \end{cases} \quad (5.68)$$

5.2.19 A Zeigen Sie $\cos(3\varphi) = 4 \cos^3(\varphi) - 3 \cos(\varphi)$.

5.3 \mathbb{C}^n als Vektorraum

Der Schritt vom \mathbb{R}^n zum \mathbb{C}^n ist nicht sehr groß, denn wir haben in den Vektoren $\mathbf{x} = [x_1, x_2, \dots, x_n]$ nur komplexe Zahlen $x_i \in \mathbb{C}$ und in den Linearkombinationen

$$\lambda \mathbf{x} + \gamma \mathbf{y}$$

$\lambda, \gamma \in \mathbb{C}$ zuzulassen. Die Rechenregeln 5.1.1 können wir ansonsten wörtlich aus dem \mathbb{R}^n übernehmen.

5.3.1 Euklidisches Skalarprodukt und euklidische Norm im \mathbb{C}^n Das Skalarprodukt $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^n x_k y_k$ zwischen zwei Vektoren $\mathbf{x} = [x_1, x_2, \dots, x_n]$ und $\mathbf{y} = [y_1, y_2, \dots, y_n]$ läßt sich so nicht aus dem \mathbb{R}^n übertragen, denn sonst würde der wichtige Zusammenhang $\langle \mathbf{x} | \mathbf{x} \rangle = \sum_{k=1}^n x_k^2 = \|\mathbf{x}\|^2$ zwischen Skalarprodukt und Norm verloren gehen. Der Grund ist, daß $\sum_{k=1}^n x_k^2$ nicht mehr positiv zu sein braucht. Diese Eigenschaft hat das Quadrat x_k^2 einer komplexen Zahl x_k im Allgemeinen nicht mehr. Damit würde die Norm eine komplexe Größe sein, was nicht unserer Vorstellung von einem Längenbegriff entspricht. Weil wir aber auch im \mathbb{C}^n ein Skalarprodukt und eine Norm zur Verfügung haben wollen, passen wir die Begriffe aus dem \mathbb{R}^n an die Situation im \mathbb{C}^n an. Am einfachsten ist das für die Norm. Wir ersetzen in $\sum_{k=1}^n x_k^2$ die Summanden x_k^2 durch $|x_k|^2 = \overline{x_k} x_k \geq 0$, d. h., wir definieren die *euklidische Norm* im \mathbb{C}^n durch

$$\|\mathbf{x}\| := \sqrt{\sum_{k=1}^n |x_k|^2}. \quad (5.69)$$

Den oben angesprochenen Zusammenhang zwischen Skalarprodukt und Norm erhalten wir aufrecht, wenn wir das *euklidische Skalarprodukt* im \mathbb{C}^n durch

$$\langle \mathbf{x} | \mathbf{y} \rangle := \sum_{k=1}^n \overline{x_k} y_k \quad (5.70)$$

eingeführen: $\langle \mathbf{x} | \mathbf{x} \rangle = \sum_{k=1}^n \overline{x_k} x_k = \sum_{k=1}^n |x_k|^2 = \|\mathbf{x}\|^2$. Die Linearität ist jetzt nur noch in der zweiten Komponente gegeben, $\langle \mathbf{x} | \lambda \mathbf{y} + \gamma \mathbf{z} \rangle = \lambda \langle \mathbf{x} | \mathbf{y} \rangle + \gamma \langle \mathbf{x} | \mathbf{z} \rangle$, während die erste Komponente *antilinear* ist: $\langle \lambda \mathbf{x} + \gamma \mathbf{y} | \mathbf{z} \rangle = \overline{\lambda} \langle \mathbf{x} | \mathbf{z} \rangle + \overline{\gamma} \langle \mathbf{y} | \mathbf{z} \rangle$. Diese Eigenschaft wird als *Sesquilinearität* bezeichnet (das meint so etwas, wie ‚halb linear‘). Wie beim \mathbb{R}^n meinen wir, wenn nicht ausdrücklich etwas anderes gesagt wird, mit dem Skalarprodukt und der Norm auf \mathbb{C}^n jeweils die euklidische Version. Fassen wir die Eigenschaften des Skalarprodukts zusammen (vergl. 5.5.6):

$$\text{i) } \langle \mathbf{x} | \mathbf{x} \rangle \geq 0, \langle \mathbf{x} | \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}, \quad (\text{Definitheit})$$

$$\text{ii) } \langle \mathbf{x} | \mathbf{y} \rangle = \overline{\langle \mathbf{y} | \mathbf{x} \rangle}, \quad (\text{Antilinearität})$$

$$\text{iii) } \langle \mathbf{x} | t\mathbf{y} + s\mathbf{z} \rangle = t \langle \mathbf{x} | \mathbf{y} \rangle + s \langle \mathbf{x} | \mathbf{z} \rangle, \quad (\text{Sesquilinearität})$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $s, t \in \mathbb{C}$.

5.3.2 Beispiel Für $\mathbf{x} := [1, 1 + i, 2i]$ und $\mathbf{y} := [1 - 3i, 5 + 2i, 1 + 3i]$ ist $\|\mathbf{x}\| = \sqrt{1 + |1 + i|^2 + 4} = \sqrt{1 + 2 + 4} = \sqrt{7}$, $\|\mathbf{y}\| = \sqrt{1 + 9 + 25 + 4 + 1 + 9} = \sqrt{49} = 7$, $\langle \mathbf{x} | \mathbf{y} \rangle = 1 + 3i + (1 + i)(5 - 2i) + 2i(1 - 3i) = 14 + 8i$.

5.4 Der allgemeine Vektorraumbegriff*

Wir haben inzwischen einige Beispiele von Mengen gesehen, die wir jeweils als Vektorraum bezeichnet haben. Eine solche Menge, etwa \mathbb{R}^n , oder die Menge \mathcal{P} der Polynome, ist mit zwei grundlegenden Rechenoperationen ausgestattet: Zwei Elemente lassen sich addieren und jeder Vektor lässt sich strecken, d. h. lässt sich mit einer Zahl multiplizieren. Die Mengen haben darüber hinaus üblicherweise weitere Eigenschaften, aber was die Zugehörigkeit zu den Vektorräumen angeht, ist nicht mehr nötig. In der folgenden Definition geben wir allgemein an, was unter einem Vektorraum zu verstehen ist.

5.4.1 Definition Ein Vektorraum (V, \mathbb{K}) über \mathbb{K} besteht aus einer Menge V , deren Elemente $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ als Vektoren bezeichnet werden und einem Körper \mathbb{K} , den sogenannten Skalaren t, s, r, \dots . Darüber hinaus gibt es eine Addition $\mathbf{x} + \mathbf{y} \in V$ für zwei Vektoren \mathbf{x} und \mathbf{y} , sowie eine Skalarmultiplikation $t \cdot \mathbf{x} \in V$ zwischen einem Skalar t und einem Vektor \mathbf{x} , die die folgenden Eigenschaften aufweisen:

- i) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$. (Kommutativität)
- ii) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$. (Assoziativität)
- iii) Es gibt einen Vektor $\mathbf{0}$, den Nullvektor, mit der Eigenschaft $\mathbf{x} + \mathbf{0} = \mathbf{x}$.
- iv) Für alle \mathbf{x} gibt es einen Vektor $-\mathbf{x}$ mit $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$. (Existenz der Inversen)
- v) $t \cdot (\mathbf{x} + \mathbf{y}) = t \cdot \mathbf{x} + t \cdot \mathbf{y}$.
- vi) $(t + s) \cdot \mathbf{x} = t \cdot \mathbf{x} + s \cdot \mathbf{x}$.
- vii) $(ts) \cdot \mathbf{x} = t \cdot (s \cdot \mathbf{x})$.
- viii) Für das Einselement $1 \in \mathbb{K}$ gilt $1 \cdot \mathbf{x} = \mathbf{x}$.

Ziehen wir ein paar elementare Folgerungen aus dieser Definition.

- 1) Es gibt nur einen Nullvektor, denn für einen Vektor $\mathbf{0}'$ mit der Eigenschaft iii) gilt $\mathbf{0}' \stackrel{\text{iii)}}{=} \mathbf{0}' + \mathbf{0} \stackrel{\text{i)}}{=} \mathbf{0} + \mathbf{0}' \stackrel{\text{iii)}}{=} \mathbf{0}$.
- 2) Für jedes $\mathbf{x} \in V$ gibt es nur einen Vektor $-\mathbf{x}$, denn aus $\mathbf{x} + \mathbf{y} = \mathbf{0}$ folgt $-\mathbf{x} \stackrel{\text{iii)}}{=} -\mathbf{x} + \mathbf{0} = -\mathbf{x} + (\mathbf{x} + \mathbf{y}) \stackrel{\text{ii)}}{=} (-\mathbf{x} + \mathbf{x}) + \mathbf{y} \stackrel{\text{i)}}{=} (\mathbf{x} + (-\mathbf{x})) + \mathbf{y} \stackrel{\text{iv)}}{=} \mathbf{0} + \mathbf{y} \stackrel{\text{i)}}{=} \mathbf{y} + \mathbf{0} \stackrel{\text{iii)}}{=} \mathbf{y}$.
Daraus folgt insbesondere $-\mathbf{0} = \mathbf{0}$ und $-(-\mathbf{x}) = \mathbf{x}$, denn $\mathbf{0} \stackrel{\text{iii)}}{=} \mathbf{0} + \mathbf{0}$, bzw. $\mathbf{0} \stackrel{\text{iv)}}{=} \mathbf{x} + (-\mathbf{x}) \stackrel{\text{i)}}{=} (-\mathbf{x}) + \mathbf{x}$.
- 3) Für $0 \in \mathbb{K}$ gilt $0 \cdot \mathbf{x} = \mathbf{0}$, denn $0 \cdot \mathbf{x} = (0 + 0) \cdot \mathbf{x} \stackrel{\text{vi)}}{=} 0 \cdot \mathbf{x} + 0 \cdot \mathbf{x}$. Daraus folgt $\mathbf{0} \stackrel{\text{iv)}}{=} 0 \cdot \mathbf{x} + (-(0 \cdot \mathbf{x})) = (0 \cdot \mathbf{x} + 0 \cdot \mathbf{x}) + (-(0 \cdot \mathbf{x})) \stackrel{\text{ii)}}{=} 0 \cdot \mathbf{x} + (0 \cdot \mathbf{x} + (-(0 \cdot \mathbf{x}))) \stackrel{\text{iv)}}{=} 0 \cdot \mathbf{x} + \mathbf{0} \stackrel{\text{iii)}}{=} 0 \cdot \mathbf{x}$.
- 4) $-(t \cdot \mathbf{x}) = (-t) \cdot \mathbf{x} = t \cdot (-\mathbf{x})$, denn $t \cdot \mathbf{x} + (-t) \cdot \mathbf{x} \stackrel{\text{vi)}}{=} (t + (-t)) \cdot \mathbf{x} = 0 \cdot \mathbf{x} = \mathbf{0}$. Daraus folgt insbesondere auch $(-1) \cdot \mathbf{x} = -(1 \cdot \mathbf{x}) \stackrel{\text{viii)}}{=} -\mathbf{x}$. Das wiederum zeigt $t \cdot (-\mathbf{x}) = t \cdot ((-1) \cdot \mathbf{x}) \stackrel{\text{vii)}}{=} (-t) \cdot \mathbf{x}$.
- 5) $t \cdot \mathbf{0} = \mathbf{0}$; $t \cdot \mathbf{0} \stackrel{\text{iii)}}{=} t \cdot (\mathbf{0} + \mathbf{0}) = t \cdot (\mathbf{0} + (-\mathbf{0})) \stackrel{\text{v)}}{=} t \cdot \mathbf{0} + t \cdot (-\mathbf{0}) = t \cdot \mathbf{0} + (-(t \cdot \mathbf{0})) = \mathbf{0}$.

- 6) Aus $t \cdot x = 0$ folgt $t = 0$, oder $x = 0$: Zunächst sei $t \neq 0$. Dann existiert $t^{-1} \in \mathbb{K}$, so daß $0 = t^{-1} \cdot 0 = t^{-1} \cdot (t \cdot x) \stackrel{\text{vii)}}{=} (t^{-1}t) \cdot x = 1 \cdot x \stackrel{\text{viii)}}{=} x$. Ist dagegen $x \neq 0$, so muß $t = 0$ gelten, denn andernfalls würde, wie eben gezeigt, $x = 0$ folgen.

Nachdem wir jetzt die gewohnten Rechenregeln gefolgert haben, können wir es uns erlauben statt $x + (y + z)$ einfach $x + y + z$ zu schreiben. Genauso verkürzen wir $x + (-y)$ zu $x - y$, indem wir die Rechenoperation $-$ durch $x - y := x + (-y)$ einführen. Wenn keine Gefahr für Mißverständnisse besteht, schreiben wir künftig auch einfach tx statt $t \cdot x$. Meistens ist aus dem Kontext klar, über welchen Körper \mathbb{K} ein gegebener Vektorraum zu bilden ist. Daher werden wir meist von dem Vektorraum V , statt von (V, \mathbb{K}) sprechen können.

Eigentlich vermißt man eine weitere Eigenschaft, nämlich, daß der einzige Vektor x , der mit seinem Inversen $-x$ übereinstimmt, der Nullvektor ist. In dieser Allgemeinheit ist das jedoch nicht wahr. In der Definition ist nämlich auch der Fall enthalten, daß \mathbb{K} der kleinstmögliche Körper $\mathbb{F}_2 = \{0, 1\}$ ist, für den die Rechenregeln $0 + 0 = 0$, $0 + 1 = 1$, $0 \cdot 1 = 0$, $1 \cdot 1 = 1$ und insbesondere $1 + 1 = 0$ gelten. Für einen Vektorraum V über \mathbb{F}_2 gilt daher $0 = 0 \cdot x = (1 + 1) \cdot x = 1 \cdot x + 1 \cdot x = x + x$. Das bedeutet $x = -x$ für alle $x \in V$.

Allerdings haben wir diese Eigenschaft sofort zur Verfügung, sobald in \mathbb{K} nicht mehr $1 + 1 = 0$ gilt. Denn dann haben wir $0 = x - x = x + x = 1 \cdot x + 1 \cdot x = (1 + 1) \cdot x$. Aus $1 + 1 \neq 0$ folgt jetzt $x = 0$.

5.4.2 Definition Für einen Vektorraum (V, \mathbb{K}) heißt eine Teilmenge $W \subseteq V$ Teilraum von V , falls (W, \mathbb{K}) mit der Addition $+$ und der Skalarmultiplikation \cdot aus (V, \mathbb{K}) wieder ein Vektorraum über \mathbb{K} ist.

5.4.3 Satz Eine Teilmenge W von V ist genau dann ein Teilraum von (V, \mathbb{K}) , wenn für alle $t, s \in \mathbb{K}$ und alle $x, y \in W$ folgendes gilt:

$$s \cdot x + t \cdot y \in W. \quad (5.71)$$

Einen Vektor der Form $t_1 \cdot x_1 + t_2 \cdot x_2 + \dots + t_n \cdot x_n$ nennen wir *Linearkombination* der Vektoren x_1, x_2, \dots, x_n . Der Satz sagt dann, daß eine Teilmenge genau dann ein Teilraum ist, wenn keine Linearkombination seiner Elemente aus ihm hinausführt.

Beweis. Der Beweis ist sehr einfach. Wenn (W, \mathbb{K}) ein Vektorraum über \mathbb{K} ist, muß nichts gezeigt werden. Für die andere Richtung müssen wir die Vektorraumaxiome für W nachrechnen. Laut (5.71) haben wir bereits $x + y \in W$ und $t \cdot y \in W$ für alle $x, y \in W$ und $t \in \mathbb{K}$. Die Eigenschaften i) – viii) für (W, \mathbb{K}) werden jetzt von (V, \mathbb{K}) bereitgestellt. \square

Es ist an der Zeit, Beispiele für Vektorräume anzugeben. Wir fangen mit einem sehr allgemeinen an, aus dem alle für uns wichtigen mit Hilfe von Satz 5.4.3 leicht folgen.

5.4.4 Satz Sei X eine Menge und \mathbb{K} ein Körper. Dann ist die Menge $\mathcal{F}(X, \mathbb{K})$ aller Funktionen auf X mit Werten in \mathbb{K} ein Vektorraum über \mathbb{K} , wenn wir die Addition und die Skalarmultiplikation folgendermaßen definieren: Für alle $x \in X$, $t \in \mathbb{K}$, $f, g \in \mathcal{F}(X, \mathbb{K})$ gilt

$$(f + g)(x) := f(x) + g(x), \quad (5.72)$$

$$(t \cdot f)(x) := t \cdot f(x). \quad (5.73)$$

Beweis. Durch (5.72) und (5.73) werden offensichtlich wieder Funktionen von X nach \mathbb{K} definiert. Daher gilt $\mathbf{f} + \mathbf{g} \in V := \mathcal{F}(X, \mathbb{K})$ und $t \cdot \mathbf{f} \in V$. Nun zu den einzelnen Punkten: Auch wenn es sich eigentlich um eine Routineaufgabe handelt, machen wir uns dieses eine Mal die Mühe, die einzelnen Punkte nachzuweisen.

Zu i): $(\mathbf{f} + \mathbf{g})(x) = \mathbf{f}(x) + \mathbf{g}(x) = \mathbf{g}(x) + \mathbf{f}(x) = (\mathbf{g} + \mathbf{f})(x)$ für alle $x \in X$, da die Addition auf \mathbb{K} kommutativ ist. Daher stimmen die beiden Funktionen $\mathbf{f} + \mathbf{g}$ und $\mathbf{g} + \mathbf{f}$ in allen ihren Funktionswerten überein. Das bedeutet $\mathbf{f} + \mathbf{g} = \mathbf{g} + \mathbf{f}$.

Zu ii): $((\mathbf{f} + \mathbf{g}) + \mathbf{h})(x) = (\mathbf{f} + \mathbf{g})(x) + \mathbf{h}(x) = \mathbf{f}(x) + \mathbf{g}(x) + \mathbf{h}(x) = \mathbf{f}(x) + (\mathbf{g} + \mathbf{h})(x) = (\mathbf{f} + (\mathbf{g} + \mathbf{h}))(x)$, für alle $\mathbf{f}, \mathbf{g}, \mathbf{h} \in V$ und alle $x \in X$. Wie unter i) folgt die Behauptung.

Zu iii): Die Nullfunktion $\mathbf{0} \in V$, die durch $\mathbf{0}(x) := 0$ definiert ist, leistet das Gewünschte: $(\mathbf{f} + \mathbf{0})(x) = \mathbf{f}(x) + \mathbf{0}(x) = \mathbf{f}(x)$ für alle $x \in X$ bedeutet $\mathbf{f} + \mathbf{0} = \mathbf{f}$.

Zu iv): Wir definieren $-\mathbf{f}$ durch $(-\mathbf{f})(x) := -\mathbf{f}(x)$. Dann gilt $(\mathbf{f} + (-\mathbf{f}))(x) = \mathbf{f}(x) + (-\mathbf{f})(x) = \mathbf{f}(x) - \mathbf{f}(x) = 0 = \mathbf{0}(x)$ für alle $x \in X$. Das zeigt $\mathbf{f} + (-\mathbf{f}) = \mathbf{0}$.

Zu v): $(t \cdot (\mathbf{f} + \mathbf{g}))(x) = t \cdot (\mathbf{f} + \mathbf{g})(x) = t \cdot (\mathbf{f}(x) + \mathbf{g}(x)) = t \cdot \mathbf{f}(x) + t \cdot \mathbf{g}(x) = (t \cdot \mathbf{f})(x) + (t \cdot \mathbf{g})(x) = (t \cdot \mathbf{f} + t \cdot \mathbf{g})(x)$ zeigt $t \cdot (\mathbf{f} + \mathbf{g}) = t \cdot \mathbf{f} + t \cdot \mathbf{g}$.

Zu vi): $((t+s) \cdot \mathbf{f})(x) = (t+s) \cdot \mathbf{f}(x) = t \cdot \mathbf{f}(x) + s \cdot \mathbf{f}(x) = (t \cdot \mathbf{f})(x) + (s \cdot \mathbf{f})(x) = (t \cdot \mathbf{f} + s \cdot \mathbf{f})(x)$. Das zeigt $(t+s) \cdot \mathbf{f} = t \cdot \mathbf{f} + s \cdot \mathbf{f}$.

Zu vii): $((ts) \cdot \mathbf{f})(x) = (ts) \cdot \mathbf{f}(x) = t \cdot s \cdot \mathbf{f}(x) = t \cdot (s \cdot \mathbf{f})(x) = (t \cdot (s \cdot \mathbf{f}))(x)$. Das zeigt $(ts) \cdot \mathbf{f} = t \cdot (s \cdot \mathbf{f})$.

Zu viii): $(1 \cdot \mathbf{f})(x) = 1 \cdot \mathbf{f}(x) = \mathbf{f}(x)$ ergibt schließlich $1 \cdot \mathbf{f} = \mathbf{f}$. □

Wir haben in diesem Satz Funktionen \mathbf{f} ausnahmsweise in fetter Schrift wiedergegeben, um ihren Charakter als Vektor gemäß unserer bisherigen Übereinkunft hervorzuheben. Davon werden wir im Folgenden aber wieder Abstand nehmen.

Wenn es sich bei X um die endliche Menge $\{1, 2, \dots, n\}$ handelt, dann läßt sich jede Funktion $\mathbf{f} : X \rightarrow \mathbb{K}$ durch Angabe aller ihrer Funktionswerte $f_1 := \mathbf{f}(1), f_2 := \mathbf{f}(2), \dots, f_n := \mathbf{f}(n)$ eindeutig festlegen. Das bedeutet, \mathbf{f} wird durch das Tupel $[f_1, f_2, \dots, f_n] \in \mathbb{K}^n$ eindeutig bestimmt. Dann gehören zu $\mathbf{f} + \mathbf{g}$ und $t \cdot \mathbf{f}$ die Tupel $[f_1 + g_1, \dots, f_n + g_n] = [f_1, \dots, f_n] + [g_1, \dots, g_n]$ bzw. $[t \cdot f_1, \dots, t \cdot f_n] = t \cdot [f_1, \dots, f_n]$, mit der Addition bzw. Skalarmultiplikation auf \mathbb{K}^n , die uns bereits geläufig ist. Satz 5.4.4 stellt jetzt sicher, daß es sich bei \mathbb{K}^n um einen Vektorraum gemäß Definition 5.4.1 handelt.

Die Menge \mathcal{P} der Polynome auf \mathbb{C} ist eine Teilmenge von $\mathcal{F}(\mathbb{C}, \mathbb{C})$, dem Vektorraum aller Funktionen von \mathbb{C} nach \mathbb{C} , gemäß Satz 5.4.4. Da für Polynome p und q auch $t \cdot p + s \cdot q$ ein Polynom ist, haben wir die Teilraumeigenschaft 5.71 von Satz 5.4.3 nachgewiesen. Damit ist \mathcal{P} ein Vektorraum über \mathbb{C} .

5.5 Vektorräume mit Norm und Skalarprodukt

Im Folgenden werden wir die Eigenschaften eines Skalarprodukts und den Zusammenhang mit einer Norm untersuchen. Dabei werden die Überlegungen an keiner Stelle einfacher, wenn wir

uns nur auf die Vektorräume \mathbb{R}^n oder \mathbb{C}^n beziehen. Einzig nötig ist ein Vektorraum V über $\mathbb{K} = \mathbb{R}$, oder $\mathbb{K} = \mathbb{C}$, für den wir zusätzlich Eigenschaften annehmen werden.

5.5.1 Definition (Norm) Eine Norm auf einem Vektorraum (V, \mathbb{K}) ist eine Abbildung

$$\|\cdot\|: V \rightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

$$i) \quad \|\mathbf{x}\| \geq 0, \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}, \quad (\text{Definitheit})$$

$$ii) \quad \|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|, \quad (\text{Homogenität})$$

$$iii) \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \quad (\text{Dreiecksungleichung})$$

für alle $\mathbf{x}, \mathbf{y} \in V, \lambda \in \mathbb{K}$.

5.5.2 Definition (Skalarprodukt) Ein Skalarprodukt auf dem Vektorraum (V, \mathbb{K}) über $\mathbb{K} = \mathbb{R}$, oder $\mathbb{K} = \mathbb{C}$ ist eine Abbildung

$$\langle \cdot | \cdot \rangle: V \times V \rightarrow \mathbb{K}$$

mit folgenden Eigenschaften:

$$i) \quad \langle \mathbf{x} | \mathbf{x} \rangle \geq 0, \langle \mathbf{x} | \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}, \quad (\text{Definitheit})$$

$$ii) \quad \langle \mathbf{x} | \mathbf{y} \rangle = \overline{\langle \mathbf{y} | \mathbf{x} \rangle}, \quad (\text{Anti)symmetrie})$$

$$iii) \quad \langle \mathbf{z} | \lambda \mathbf{x} + \gamma \mathbf{y} \rangle = \lambda \langle \mathbf{z} | \mathbf{x} \rangle + \gamma \langle \mathbf{z} | \mathbf{y} \rangle, \quad (\text{Sesquilinearität})$$

für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \lambda, \gamma \in \mathbb{K}$.

Eine solche Abbildung $\langle \cdot | \cdot \rangle$ nennt man auch *sesquilinear*, oder *Sesquilinearform*, wenn es sich bei \mathbb{K} um \mathbb{C} handelt, denn sie ist nur in der zweiten Komponente linear, in der ersten aber *antilinear*, wie man sofort aus ii) und iii) schließt:

$$\begin{aligned} \langle \lambda \mathbf{x} + \gamma \mathbf{y} | \mathbf{z} \rangle &= \overline{\langle \mathbf{z} | \lambda \mathbf{x} + \gamma \mathbf{y} \rangle} = \overline{\lambda \langle \mathbf{z} | \mathbf{x} \rangle + \gamma \langle \mathbf{z} | \mathbf{y} \rangle} = \bar{\lambda} \overline{\langle \mathbf{z} | \mathbf{x} \rangle} + \bar{\gamma} \overline{\langle \mathbf{z} | \mathbf{y} \rangle} \\ &= \bar{\lambda} \langle \mathbf{x} | \mathbf{z} \rangle + \bar{\gamma} \langle \mathbf{y} | \mathbf{z} \rangle. \end{aligned}$$

Für $\mathbb{K} = \mathbb{R}$ ist ii) einfach die Symmetrie, aus der mit iii) die Linearität in beiden Komponenten von $\langle \cdot | \cdot \rangle$, die *Bilinearität*, folgt.

5.5.3 Satz (CAUCHY-SCHWARZ-Ungleichung) Für alle $\mathbf{x}, \mathbf{y} \in V$ gilt

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \langle \mathbf{x} | \mathbf{x} \rangle^{1/2} \langle \mathbf{y} | \mathbf{y} \rangle^{1/2}. \quad (5.74)$$

Gleichheit gilt genau dann, wenn \mathbf{x} und \mathbf{y} linear abhängig sind.

Beweis. Wir gehen von $\mathbb{K} = \mathbb{C}$ aus, da die Beweisführung leicht auf die einfachere Situation $\mathbb{K} = \mathbb{R}$ übertragen werden kann. Für den Fall $\mathbf{x} = \mathbf{0}$, oder $\mathbf{y} = \mathbf{0}$ ist die Behauptung trivial. Daher können wir für unsere Überlegungen von $\mathbf{y} \neq \mathbf{0}$ ausgehen. Aus i) folgt $\langle \mathbf{x} - \lambda \mathbf{y} | \mathbf{x} - \lambda \mathbf{y} \rangle \geq$

0 für alle $\mathbf{x}, \mathbf{y} \in V$ und $\lambda \in \mathbb{C}$. Wir haben also $0 \leq \langle \mathbf{x} | \mathbf{x} \rangle - \bar{\lambda} \langle \mathbf{y} | \mathbf{x} \rangle - \lambda \langle \mathbf{x} | \mathbf{y} \rangle + |\lambda|^2 \langle \mathbf{y} | \mathbf{y} \rangle$. Wir wählen $\lambda := \langle \mathbf{y} | \mathbf{x} \rangle / \langle \mathbf{y} | \mathbf{y} \rangle$ und erhalten damit

$$0 \leq \langle \mathbf{x} | \mathbf{x} \rangle - 2 \frac{|\langle \mathbf{x} | \mathbf{y} \rangle|^2}{\langle \mathbf{y} | \mathbf{y} \rangle} + \frac{|\langle \mathbf{x} | \mathbf{y} \rangle|^2}{\langle \mathbf{y} | \mathbf{y} \rangle} = \langle \mathbf{x} | \mathbf{x} \rangle - \frac{|\langle \mathbf{x} | \mathbf{y} \rangle|^2}{\langle \mathbf{y} | \mathbf{y} \rangle}.$$

Multiplikation mit $\langle \mathbf{y} | \mathbf{y} \rangle$ und Umstellen nach $|\langle \mathbf{x} | \mathbf{y} \rangle|^2$ ergibt die quadrierte Fassung der CAUCHY-SCHWARZschen Ungleichung. Die Monotonie der Wurzel (vergl. die Lösung zu Übung 5.2.5) zeigt die behauptete Ungleichung.

Es ist klar, daß in der Ungleichung das Gleichheitszeichen gilt, wenn \mathbf{x} und \mathbf{y} linear abhängig sind (man überlege sich das). Wir müssen auch die Umkehrung zeigen. D. h. wir gehen von $|\langle \mathbf{x} | \mathbf{y} \rangle| = \langle \mathbf{x} | \mathbf{x} \rangle^{1/2} \langle \mathbf{y} | \mathbf{y} \rangle^{1/2}$ aus und müssen zeigen, daß $\gamma \mathbf{x} + \delta \mathbf{y} = \mathbf{0}$ für geeignete $\gamma, \delta \in \mathbb{C}$, $(\gamma, \delta) \neq (0, 0)$ gilt. Ist $\mathbf{y} = \mathbf{0}$, so können wir $\gamma = 0$ und $\delta = 1$ wählen (entsprechend, falls $\mathbf{x} = \mathbf{0}$ sein sollte). Wir dürfen also wieder von $\mathbf{x} \neq \mathbf{0}$ und $\mathbf{y} \neq \mathbf{0}$ ausgehen. Aus der EULER-Darstellung $\langle \mathbf{x} | \mathbf{y} \rangle = |\langle \mathbf{x} | \mathbf{y} \rangle| e^{i\varphi}$ ergibt sich mit $\gamma := 1$ und $\delta := -\frac{\langle \mathbf{x} | \mathbf{x} \rangle^{1/2}}{\langle \mathbf{y} | \mathbf{y} \rangle^{1/2}} e^{-i\varphi}$

$$\begin{aligned} \langle \mathbf{x} + \delta \mathbf{y} | \mathbf{x} + \delta \mathbf{y} \rangle &= \langle \mathbf{x} | \mathbf{x} \rangle + 2 \operatorname{Re} \delta \langle \mathbf{x} | \mathbf{y} \rangle + |\delta|^2 \langle \mathbf{y} | \mathbf{y} \rangle \\ &= \langle \mathbf{x} | \mathbf{x} \rangle - 2 \operatorname{Re} \frac{\langle \mathbf{x} | \mathbf{x} \rangle^{1/2}}{\langle \mathbf{y} | \mathbf{y} \rangle^{1/2}} e^{-i\varphi} \langle \mathbf{x} | \mathbf{y} \rangle + \langle \mathbf{x} | \mathbf{x} \rangle \\ &= 2 \langle \mathbf{x} | \mathbf{x} \rangle - 2 \frac{\langle \mathbf{x} | \mathbf{x} \rangle^{1/2}}{\langle \mathbf{y} | \mathbf{y} \rangle^{1/2}} |\langle \mathbf{x} | \mathbf{y} \rangle| = 0. \end{aligned}$$

Aus der Definitheit des Skalarprodukts folgt $\mathbf{x} + \delta \mathbf{y} = \mathbf{0}$, d. h. \mathbf{x} und \mathbf{y} sind linear abhängig. \square

5.5.4 Korollar Durch $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ wird eine Norm auf V definiert.

Beweis. Die Eigenschaften i) und ii) sind leicht einzusehen. Wir müssen die Dreiecksungleichung nachweisen:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y} | \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \langle \mathbf{x} | \mathbf{y} \rangle + \langle \mathbf{y} | \mathbf{x} \rangle + \|\mathbf{y}\|^2 \\ &= \|\mathbf{x}\|^2 + 2 \operatorname{Re} \langle \mathbf{x} | \mathbf{y} \rangle + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2 |\operatorname{Re} \langle \mathbf{x} | \mathbf{y} \rangle| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2 |\langle \mathbf{x} | \mathbf{y} \rangle| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2 \|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Die Monotonie der Wurzel liefert schließlich den Beweis der Behauptung. \square

Mit diesem Ergebnis lautet die CAUCHY-SCHWARZ-Ungleichung jetzt

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|. \quad (5.75)$$

Vektorräume mit so reichhaltiger Struktur haben eigene Namen erhalten:

5.5.5 Definition Ein Vektorraum V über \mathbb{K} , der mit einer Norm $\|\cdot\|$ versehen ist, heißt normierter Raum, den wir, wenn wir ausführlich sein wollen, durch $(V, \|\cdot\|)$ kennzeichnen. Hat er ein Skalarprodukt $\langle \cdot | \cdot \rangle$, so wird er als Skalarprodukt-Raum bezeichnet, mit der ausführlichen Notation $(V, \langle \cdot | \cdot \rangle)$.

Haben diese Räume eine weitere Eigenschaft, eine Vollständigkeit gegenüber Grenzwertbildung, so werden sie als *Banachraum* bzw. *Hilbertraum* bezeichnet. Diese zusätzliche Eigenschaft ist für die endlichdimensionalen Räume, die wir im Rahmen der linearen Algebra überwiegend behandeln, immer erfüllt. Ein Hilbertraum ist immer auch ein Banachraum, da das Skalarprodukt eine kanonische Norm erzeugt. Die Umkehrung gilt im Allgemeinen nicht.

5.5.6 Das Skalarprodukt auf \mathbb{C}^n Als Anwendung behandeln wir noch einmal das kanonische Skalarprodukt 5.70 auf \mathbb{C}^n . Für $\mathbf{x} = [x_1, \dots, x_n]^t \in \mathbb{C}^n$ und $\mathbf{y} = [y_1, \dots, y_n]^t \in \mathbb{C}^n$ ist es folgendermaßen definiert:

$$\langle \mathbf{x} | \mathbf{y} \rangle := \sum_{k=1}^n \overline{x_k} y_k.$$

Wir prüfen die Bedingungen von 5.5.2 nach. $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$ ist klar.

$$\langle \mathbf{x} | \mathbf{x} \rangle = \sum_{k=1}^n |x_k|^2 = 0 \Leftrightarrow |x_k| = 0 \text{ für } k = 1, \dots, n \Leftrightarrow \mathbf{x} = [0, \dots, 0]^t = \mathbf{0}.$$

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^n \overline{x_k} y_k = \sum_{k=1}^n \overline{\overline{x_k} y_k} = \sum_{k=1}^n \overline{y_k} x_k = \overline{\langle \mathbf{y} | \mathbf{x} \rangle}.$$

$$\langle \mathbf{x} | t\mathbf{y} + s\mathbf{z} \rangle = \sum_{k=1}^n \overline{x_k} (ty_k + sz_k) = t \sum_{k=1}^n \overline{x_k} y_k + s \sum_{k=1}^n \overline{x_k} z_k = t \langle \mathbf{x} | \mathbf{y} \rangle + s \langle \mathbf{x} | \mathbf{z} \rangle.$$

Unsere allgemeinen Überlegungen zum Skalarprodukt gelten jetzt für diese spezielle Version. Wir wissen also, daß durch $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ eine Norm auf \mathbb{C}^n gegeben ist, für die die CAUCHY-SCHWARZ-Ungleichung und die Dreiecksungleichung gilt. Schreiben wir uns das einmal explizit auf:

$$\left| \sum_{k=1}^n \overline{x_k} y_k \right| \leq \sqrt{\sum_{k=1}^n |x_k|^2} \sqrt{\sum_{k=1}^n |y_k|^2},$$

$$\sqrt{\sum_{k=1}^n |x_k + y_k|^2} \leq \sqrt{\sum_{k=1}^n |x_k|^2} + \sqrt{\sum_{k=1}^n |y_k|^2}.$$

Man sieht, daß diese konkrete Realisierung durch die Koordinaten, die man durchaus als weniger abstrakt als unsere Überlegungen aus den letzten Abschnitten empfinden könnte, den Beweis dieser Ungleichungen nicht erleichtert hätte (wer das anders sieht, kann ja einmal versuchen, die Beweise in dieser Koordinatenversion zu führen).

Wir werden immer wieder auf kreis- oder kugelförmige Umgebungen eines Punktes \mathbf{x}_0 stoßen, für die wir eine einheitliche Notation benötigen:

5.5.7 Definition Auf einem normierten Raum $(V, \|\cdot\|)$ bezeichnen wir die Mengen

$$U_r(\mathbf{x}_0) := \{ \mathbf{x} \in V \mid \|\mathbf{x} - \mathbf{x}_0\| < r \} \quad (5.76)$$

$$\bar{U}_r(\mathbf{x}_0) := \{ \mathbf{x} \in V \mid \|\mathbf{x} - \mathbf{x}_0\| \leq r \} \quad (5.77)$$

als offene bzw. abgeschlossene Kugeln mit Mittelpunkt \mathbf{x}_0 und Radius r .

In der Analysis werden wir diesen Umgebungen auch begegnen, nur wird dort traditionell das Symbol Epsilon ε für den Radius verwendet und von der ε -Umgebung des Punktes \mathbf{x}_0 gesprochen ($\varepsilon > 0$ kennzeichnet in der Mathematik üblicherweise eine kleine Größe). Normalerweise meint man damit immer die offene Kugel um \mathbf{x}_0 mit Radius $\varepsilon > 0$.

Für $(\mathbb{R}^n, \|\cdot\|)$ mit der euklidischen Norm $\| [x_1, \dots, x_n]^t \| = \sqrt{\sum_{k=1}^n x_k^2}$, handelt es sich bei

$U_r(\mathbf{x}_0)$ für $n = 1$ um das offene Intervall $(x_0 - r, x_0 + r)$, für $n = 2$ um die offene Kreisscheibe mit Radius r und Mittelpunkt $\mathbf{x}_0 \in \mathbb{R}^2$ und endlich für $n = 3$ tatsächlich um eine Kugel im landläufigen Sinne, allerdings ohne die Punkte der Kugelschale. Durch die Kugelschale $K_r(\mathbf{x}_0) := \{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| = r \}$ unterscheidet sich $\bar{U}_r(\mathbf{x}_0)$ von $U_r(\mathbf{x}_0)$ (vergl. (5.1.16)).

Versehen wir \mathbb{R}^n mit anderen Normen, dann werden die Mengen $U_r(\mathbf{x}_0)$ bzw. $\bar{U}_r(\mathbf{x}_0)$ mangels besserer Bezeichnung gemäß der Definition weiterhin als Kugeln bezeichnet, auch wenn sie gar nicht mehr unserer Vorstellung von Kugeln entsprechen.

5.5.8 A Eine wichtige Klasse von Normen sind die sogenannten L^p -Normen $\|\cdot\|_p$. Für $1 \leq p \leq \infty$ ist ihre Definition unter iii) bzw. iv) für den Vektorraum \mathbb{R}^2 wiedergegeben. Die L^2 -Norm ist die uns schon bekannte euklidische Norm. Skizzieren Sie die abgeschlossenen Einheitskugeln $\bar{U}_1(\mathbf{0})$ für die L^1 -, die L^2 -, die L^3 - und die L^∞ -Norm. Zeigen Sie, daß es sich bei $\|\cdot\|_1$ und $\|\cdot\|_\infty$ tatsächlich um Normen handelt.

i) $\bar{U}_1(\mathbf{0}) = \{ [x_1, x_2]^t \in \mathbb{R}^2 \mid |x_1| + |x_2| \leq 1 \}$ gehört zur L^1 -Norm:

$$\|[x_1, x_2]^t\|_1 := |x_1| + |x_2|.$$

ii) $\bar{U}_1(\mathbf{0}) = \{ [x_1, x_2]^t \in \mathbb{R}^2 \mid |x_1|^2 + |x_2|^2 \leq 1 \}$ gehört zur L^2 -, oder euklidischen Norm:

$$\|[x_1, x_2]^t\|_2 := \sqrt{|x_1|^2 + |x_2|^2}.$$

iii) $\bar{U}_1(\mathbf{0}) = \{ [x_1, x_2]^t \in \mathbb{R}^2 \mid |x_1|^p + |x_2|^p \leq 1 \}$ gehört zur L^p -Norm:

$$\|[x_1, x_2]^t\|_p := \sqrt[p]{|x_1|^p + |x_2|^p}, p \geq 1.$$

iv) $\bar{U}_1(\mathbf{0}) = \{ [x_1, x_2]^t \in \mathbb{R}^2 \mid \max\{|x_1|, |x_2|\} \leq 1 \}$ gehört zur L^∞ -Norm:

$$\|[x_1, x_2]^t\|_\infty := \max\{|x_1|, |x_2|\}.$$

Die Bezeichnung $\|\cdot\|_\infty$ kommt nicht von ungefähr, denn es gilt $\lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty$.

5.5.9 A Zeigen Sie: Für eine Norm gilt die sog. *umgekehrte Dreiecksungleichung*

$$\|\mathbf{x} - \mathbf{y}\| \geq \left| \|\mathbf{x}\| - \|\mathbf{y}\| \right|. \quad (5.78)$$

5.5.10 A (Satz von RIESZ) Eine *Linearform* ℓ auf einem Vektorraum V über dem Körper \mathbb{K} ist eine lineare Abbildung von V nach \mathbb{K} . Zeigen Sie: Jede Linearform ℓ auf $V = \mathbb{C}^n$ ist von der Form

$$\ell(\mathbf{x}) = \langle \mathbf{y}_\ell | \mathbf{x} \rangle, \quad (5.79)$$

mit einem durch ℓ eindeutig bestimmten Vektor $\mathbf{y}_\ell \in \mathbb{C}^n$. *Hinweis: Entwickeln Sie \mathbf{x} in der kanonischen Basis von \mathbb{C}^n .*

5.5.11 A (Polarisationsgleichung) Ein Skalarprodukt definiert über $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ die kanonisch zugeordnete Norm. Umgekehrt kann man das Skalarprodukt aus dieser Norm auch wieder zurückgewinnen. Zeigen Sie dazu die sogenannte *Polarisationsgleichung*

$$\langle \mathbf{x} | \mathbf{y} \rangle = \frac{1}{4} [\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2] \quad (5.80)$$

$$\langle \mathbf{x} | \mathbf{y} \rangle = \frac{1}{4} [\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 - i(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2)] \quad (5.81)$$

Dabei gilt (5.80) für Vektorräume über \mathbb{R} und (5.81) für Vektorräume über \mathbb{C} .

5.5.12 A Zeigen Sie: Für ein Skalarprodukt $\langle \cdot | \cdot \rangle$ auf V gilt:

$$\langle \mathbf{x} | \mathbf{y} \rangle = 0 \text{ f. a. } \mathbf{y} \in V \Leftrightarrow \mathbf{x} = \mathbf{0}.$$

Schließen Sie daraus: Aus $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{z} | \mathbf{y} \rangle$ für alle $\mathbf{y} \in V$ folgt $\mathbf{x} = \mathbf{z}$.

5.5.13 A Folgern Sie aus $\|\|\mathbf{a}\|^2 \mathbf{b} - \langle \mathbf{a} | \mathbf{b} \rangle \mathbf{a}\|^2 \geq 0$ die CAUCHY-SCHWARZ-Ungleichung.

5.5.14 Die HÖLDER- und MINKOWSKI-Ungleichung* Es seien $\mathbf{x} := [x_1, \dots, x_n]^t$ und $\mathbf{y} := [y_1, \dots, y_n]^t$ Vektoren aus \mathbb{C}^n (oder \mathbb{R}^n), sowie $1 < p < \infty$. Der Nachweis, daß

$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{k=1}^n |x_k|^p} = \left[\sum_{k=1}^n |x_k|^p \right]^{\frac{1}{p}} \quad (5.82)$$

eine Norm definiert, steht und fällt mit dem Beweis der Dreiecksungleichung. Das war schon für $p = 2$ nicht ganz einfach, wo die CAUCHY-SCHWARZsche Ungleichung $|\langle \mathbf{x} | \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$ das entscheidende Werkzeug war. Diese Rolle übernimmt im allgemeinen Fall die HÖLDER-Ungleichung

$$\sum_{k=1}^n |x_k y_k| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q. \quad (5.83)$$

Dabei ist $q > 1$ die Lösung der Gleichung $\frac{1}{p} + \frac{1}{q} = 1$. Für $p = 2$ ergibt sich daraus die CAUCHY-SCHWARZsche Ungleichung:

$$|\langle \mathbf{x} | \mathbf{y} \rangle| = \left| \sum_{k=1}^n \overline{x_k} y_k \right| \leq \sum_{k=1}^n |x_k y_k| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2.$$

Ausgangspunkt für (5.83) ist die YOUNG'sche Ungleichung in der Form

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad (5.84)$$

für $a, b > 0$ und $\frac{1}{p} + \frac{1}{q} = 1$, die wir an anderer Stelle beweisen werden (Satz 11.9.16). Wir definieren $a_i := \frac{|x_i|}{\|\mathbf{x}\|_p}$, $b_i := \frac{|y_i|}{\|\mathbf{y}\|_q}$ und haben dann nach (5.84)

$$\begin{aligned} \frac{1}{\|\mathbf{x}\|_p \|\mathbf{y}\|_q} \sum_{i=1}^n |x_i y_i| &= \sum_{i=1}^n a_i b_i \leq \frac{1}{p} \sum_{i=1}^n a_i^p + \frac{1}{q} \sum_{i=1}^n b_i^q = \frac{1}{p} \sum_{i=1}^n \frac{|x_i|^p}{\|\mathbf{x}\|_p^p} + \frac{1}{q} \sum_{i=1}^n \frac{|y_i|^q}{\|\mathbf{y}\|_q^q} \\ &= \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

Daraus folgt unmittelbar (5.83). Mit dieser Ungleichung läßt sich jetzt die sogenannte MINKOWSKI-Ungleichung beweisen, die einfach die Dreiecksungleichung für $\|\cdot\|_p$ darstellt. Unter Verwendung von $q(p-1) = p$ und $\frac{1}{q} = \frac{p-1}{p}$ schätzen wir folgendermaßen ab:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_p^p &= \sum_{i=1}^n |x_i + y_i| |x_i + y_i|^{p-1} \leq \sum_{i=1}^n |x_i| |x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i| |x_i + y_i|^{p-1} \\ &\stackrel{5.83}{\leq} \|\mathbf{x}\|_p \left[\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right]^{\frac{1}{q}} + \|\mathbf{y}\|_p \left[\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right]^{\frac{1}{q}} \\ &= (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p) \left[\sum_{i=1}^n |x_i + y_i|^p \right]^{\frac{1}{p}(p-1)} = (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p) \cdot \|\mathbf{x} + \mathbf{y}\|_p^{p-1}. \end{aligned}$$

Das läßt sich sofort in die MINKOWSKI-Ungleichung umformen:

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p \quad (5.85)$$

Den Vektorraum $(\mathbb{C}^n, \|\cdot\|_p)$ bezeichnet man als L^p -Raum.

6 Matrizen

In einer Gießerei werden Legierungen aus Kupfer (Cu), Zinn (Sn) und Zink (Zn) benötigt. Als Restposten alter Aufträge liegen noch Legierungen mit 80% Cu, 15% Sn, 5% Zn, sowie 85% Cu, 12% Sn, 3% Zn und 75% Cu, 10% Sn, 15% Zn vor. Kann daraus durch Zusammenschmelzen geeigneter Anteile dieser Legierungen der neue Auftrag mit 81% Cu, 12% Sn und 7% Zn bedient werden?

Der erste Schritt zu einer Lösung dieser Aufgabe besteht darin, die einzelnen Legierungen übersichtlich zusammenzufassen. Wir bedienen uns dabei der Vektorrechnung, indem wir jeder Legierung einen Vektor zuordnen, dessen Einträge die prozentualen Anteile an Cu, Sn und Zn sind. Zur ersten Legierung gehört demnach der Vektor $[80, 15, 5]$, zur zweiten $[85, 12, 3]$ und zur dritten $[75, 10, 15]$. Der Vektor des Auftrags ist $[81, 12, 7]$.

Wir versuchen aus x Anteilen der ersten Legierung, y Anteilen der zweiten und z Anteilen der dritten die gewünschte Legierung zu mischen:

$$x \begin{bmatrix} 80 \\ 15 \\ 5 \end{bmatrix} + y \begin{bmatrix} 85 \\ 12 \\ 3 \end{bmatrix} + z \begin{bmatrix} 75 \\ 10 \\ 15 \end{bmatrix} = \begin{bmatrix} 80x + 85y + 75z \\ 15x + 12y + 10z \\ 5x + 3y + 15z \end{bmatrix} = \begin{bmatrix} 81 \\ 12 \\ 7 \end{bmatrix}.$$

Daraus erhalten wir das lineare Gleichungssystem

$$\begin{aligned} 80x + 85y + 75z &= 81 \\ 15x + 12y + 10z &= 12 \\ 5x + 3y + 15z &= 7, \end{aligned}$$

für das wir ein systematisches Lösungsverfahren benötigen.

6.1 Lineare Gleichungssysteme und das Gauss-Verfahren

Ein *lineares Gleichungssystem* (LGS) besteht im Allgemeinen aus m Gleichungen der Form

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i, \quad i = 1, \dots, m$$

für die n Unbekannten $x_1, \dots, x_n \in \mathbb{K}$. Dabei sind für $i = 1, \dots, m$ und $j = 1, \dots, n$ die Zahlen $a_{ij} \in \mathbb{K}$ – die *Koeffizienten* des Gleichungssystems – fest gewählt. Ausgeschrieben:

$$\begin{array}{cccccccl} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \quad (6.1)$$

Sind die Zahlen b_1, \dots, b_m auf der rechten Seite alle gleich Null, so heißt das Gleichungssystem *homogen*, andernfalls *inhomogen*.

Um zu dem Lösungsverfahren für (6.1) zu gelangen, das als GAUSS-Verfahren bekannt ist, machen wir uns klar, welche Operationen eine Gleichung in eine äquivalente überführt: Zu einer Gleichung kann auf beiden Seiten dieselbe Zahl addiert werden. Außerdem kann man beide Seiten mit ein und derselben Zahl multiplizieren, solange es sich dabei nicht um Null handelt. Wenn wir diese beiden Möglichkeiten geschickt kombinieren, können wir (6.1) soweit vereinfachen, daß die Existenz einer Lösung leicht zu erkennen und diese gegebenenfalls auch einfach zu bestimmen ist. Das Verfahren verläuft nach folgendem Schema. Wir dürfen o. B. d. A. davon ausgehen, daß $a_{11} \neq 0$ ist, denn andernfalls können wir uns die Gleichungen so vertauscht denken, daß das erste Element der ersten Gleichung nicht verschwindet (sollten alle Elemente a_{1j} der ersten Spalte gleich Null sein, dann wenden wir unser Verfahren auf die erste Spalte an, die wenigstens einen nicht verschwindenden Koeffizienten a_{ij} aufweist). Sollte $a_{21} = 0$ sein, dann ist für den Moment an der zweiten Gleichung nichts zu ändern. Andernfalls dürfen wir die erste Gleichung mit a_{21} und die zweite mit a_{11} multiplizieren:

$$\begin{aligned} a_{21}a_{11}x_1 + a_{21}a_{12}x_2 + \dots + a_{21}a_{1n}x_n &= a_{21}b_1 \\ a_{11}a_{21}x_1 + a_{11}a_{22}x_2 + \dots + a_{11}a_{2n}x_n &= a_{11}b_2. \end{aligned}$$

Jetzt können wir auf beiden Seiten der zweiten Gleichung die Zahl $a_{21}b_1$ abziehen. Auf der rechten erhalten wir dabei $a_{11}b_2 - a_{21}b_1$. Der Trick ist nun, für die linke Seite die Zahl $a_{21}b_1$ durch den Ausdruck $a_{21}a_{11}x_1 + a_{21}a_{12}x_2 + \dots + a_{21}a_{1n}x_n$ zu ersetzen, von dem die erste Gleichung ja behauptet, daß er mit $a_{21}b_1$ übereinstimmt. Wenn das Gleichungssystem lösbar ist, kann das zu keinem Widerspruch führen, da die erste Gleichung dann durch geeignete Zahlen x_1, \dots, x_n erfüllt werden kann. Wie man sieht ist der Sinn dieses Verfahrens, die Unbekannte x_1 aus der zweiten Gleichung zu eliminieren:

$$\begin{aligned} a_{21}a_{11}x_1 + a_{21}a_{12}x_2 + \dots &= a_{21}b_1 \\ (a_{11}a_{22} - a_{21}a_{12})x_2 + \dots &= a_{11}b_2 - a_{21}b_1. \end{aligned}$$

Wenn wir die erste Gleichung durch Division mit a_{21} wieder in ihre Ausgangsform zurück verwandeln, ergibt sich

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots &= b_1 \\ (a_{11}a_{22} - a_{21}a_{12})x_2 + \dots &= a_{11}b_2 - a_{21}b_1 \\ a_{31}x_1 + a_{32}x_2 + \dots &= b_3 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \dots &= b_m. \end{aligned}$$

Diesen Vorgang werden wir künftig meist dadurch abkürzen, daß wir einfach das a_{21} -fache der ersten Gleichung von der modifizierten zweiten abziehen, ohne dafür die erste Gleichung tatsächlich zu verändern. Nun wenden wir dieses Verfahren auf die dritte, die vierte etc. an bis

wir bei der letzten angelangt sind. Wir erhalten dabei ein neues Gleichungssystem der Form

$$\begin{array}{ccccccccc}
 a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
 & & \tilde{a}_{22}x_2 & + & \tilde{a}_{23}x_3 & + & \cdots & + & \tilde{a}_{2n}x_n & = & \tilde{b}_1 \\
 & & \tilde{a}_{32}x_2 & + & \tilde{a}_{33}x_3 & + & \cdots & + & \tilde{a}_{3n}x_n & = & \tilde{b}_3 \\
 & & \vdots & & \vdots & & & & \vdots & & \vdots \\
 & & \tilde{a}_{m2}x_2 & + & \tilde{a}_{m3}x_3 & + & \cdots & + & \tilde{a}_{mn}x_n & = & \tilde{b}_m,
 \end{array}$$

mit neuen Koeffizienten \tilde{a}_{ij} , die in der ersten Spalte, bis auf den ersten, alle verschwinden. Jetzt wenden wir dieses Verfahren auf die zweite Spalte an. Dabei arbeiten wir nun mit der zweiten Gleichung. Wir können o. B. d. A. wieder $\tilde{a}_{22} \neq 0$ annehmen, denn andernfalls vertauschen wir die Gleichungen 2, ..., m, bis der Eintrag \tilde{a}_{22} ungleich Null ist. Sollten alle Koeffizienten $\tilde{a}_{22}, \dots, \tilde{a}_{m2}$ verschwinden, so machen wir mit der dritten Spalte weiter, vertauschen gegebenenfalls die Gleichungen bis $\tilde{a}_{23} \neq 0$ gilt, usw. Wie oben beschrieben bringen wir die Koeffizienten \tilde{a}_{32} bis \tilde{a}_{m2} , aber auch a_{12} zum Verschwinden (bzw. \tilde{a}_{33} bis \tilde{a}_{m3} und a_{13}). Das Gleichungssystem hat mittlerweile die Form

$$\begin{array}{ccccccccc}
 a_{11}x_1 & & & + & \hat{a}_{13}x_3 & + & \cdots & + & \hat{a}_{1n}x_n & = & \hat{b}_1 \\
 & \tilde{a}_{22}x_2 & + & \tilde{a}_{23}x_3 & + & \cdots & + & \tilde{a}_{2n}x_n & = & \tilde{b}_1 \\
 & & \hat{a}_{33}x_3 & + & \cdots & + & \hat{a}_{3n}x_n & = & \hat{b}_3 \\
 & & \vdots & & & & \vdots & & \vdots \\
 & & \hat{a}_{m3}x_3 & + & \cdots & + & \hat{a}_{mn}x_n & = & \hat{b}_m,
 \end{array}$$

bzw., falls die zweite Spalte bereits bei der Bearbeitung der ersten mit verschwunden ist:

$$\begin{array}{ccccccccc}
 a_{11}x_1 & + & a_{12}x_2 & & & + & \hat{a}_{14}x_4 & + & \cdots & + & \hat{a}_{1n}x_n & = & \hat{b}_1 \\
 & & & \tilde{a}_{23}x_3 & + & \tilde{a}_{24}x_4 & + & \cdots & + & \tilde{a}_{2n}x_n & = & \tilde{b}_1 \\
 & & & & \hat{a}_{34}x_4 & + & \cdots & + & \hat{a}_{3n}x_n & = & \hat{b}_3 \\
 & & & & \vdots & & & & \vdots & & \vdots \\
 & & & & \hat{a}_{m4}x_4 & + & \cdots & + & \hat{a}_{mn}x_n & = & \hat{b}_m.
 \end{array}$$

\hat{a}_{ij} und \hat{b}_i sind die Koeffizienten, die sich bei diesem zweiten Durchgang ergeben haben. Auf diese Weise fahren wir fort, bis die letzte Spalte abgearbeitet ist. Dabei kann es durchaus passieren, daß bei einem Arbeitsschritt eine ganze Gleichung verschwindet. Das bedeutet, daß von den m Ausgangsgleichungen wenigstens eine überflüssig ist und daß von den m Unbekannten, die durch die m Gleichungen festgelegt werden sollen tatsächlich höchstens $m - 1$ zu bestimmen sind. Pro verschwundener Gleichung erhöht sich die Anzahl der Unbekannten, die nicht festgelegt werden können um eins. Daß solche Situationen auftreten ist zu erwarten. Man muß sich ja nur vorstellen, daß man zwei Gleichungen für drei Unbekannte hat, sagen wir

$$\begin{aligned}
 2x_1 + 3x_2 - 4x_3 &= -3 \\
 5x_1 - 2x_2 + 3x_3 &= 10.
 \end{aligned}$$

Sicher werden wir nichts Neues erfahren, wenn wir das Doppelte der ersten von der zweiten abziehen und das Ergebnis als dritte Gleichung verwenden. Obwohl es sich dann um drei Gleichungen

chungen handelt, liefert eine von ihnen keine weiteren Informationen. Man kann sich vorstellen, daß solche Beispiele mit wachsender Anzahl von Gleichungen beliebig verwickelt werden können. Das GAUSS-Verfahren entdeckt sie jedoch alle.

Es ist auch möglich, daß bei einem Arbeitsschritt nur die linke Seite, nicht aber die rechte verschwindet. In diesem Fall ist das Gleichungssystem nicht lösbar, denn unter der Annahme, daß es lösbar ist (und nur unter dieser Prämisse können wir das Verfahren rechtfertigen) sind wir zu einem Widerspruch in Form einer unlösbaren Gleichung gelangt. Das bedeutet, daß die Annahme der Lösbarkeit nicht wahr sein kann.

Nach Beendigung des Verfahrens ist die ursprünglich rechteckige Gestalt des Gleichungssystems in eine Stufenform verwandelt worden, aus der die Lösungen, falls sie existieren, leicht bestimmt werden können. Ein typisches Ergebnis könnte etwa folgendermaßen aussehen:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 & + \bar{a}_{15}x_5 + \bar{a}_{16}x_6 & = \bar{b}_1 \\ & \bar{a}_{23}x_3 & + \bar{a}_{25}x_5 + \bar{a}_{26}x_6 = \bar{b}_2 \\ & & \bar{a}_{34}x_4 + \bar{a}_{35}x_5 + \bar{a}_{36}x_6 = \bar{b}_3 \\ & & & \bar{a}_{47}x_7 = \bar{b}_4, \end{array}$$

oder mit konkreten Zahlen

$$\begin{array}{rcl} x_1 - 2x_2 & + 7x_5 - 5x_6 & = 1 \\ & x_3 & - x_5 - 8x_6 = 4 \\ & & x_4 + 9x_5 + 7x_6 = 6 \\ & & & x_7 = 2. \end{array}$$

Hier wären die gesuchten Größen x_1 , x_3 , x_4 und x_7 für *jede Wahl* von x_2 , x_5 und x_6 eindeutig zu berechnen:

$$x_1 = 1 + 2x_2 - 7x_5 + 5x_6, \quad x_3 = 4 + x_5 + 8x_6, \quad x_4 = 6 - 9x_5 - 7x_6, \quad x_7 = 2.$$

Anders gesagt: Wir können das System nach x_1 , x_3 , x_4 und x_7 auflösen, während x_2 , x_5 und x_6 nicht weiter bestimmbar sind. Jede Wahl von x_2 , x_5 und x_6 führt nach unserer Lösungsformel zu einer Lösung. Es gibt daher unendlich viele.

Und wie sehen diese Lösungen aus?

Wenn wir vereinbaren, die Lösungen als Elemente $[x_1, x_2, x_3, x_4, x_5, x_6, x_7]$ aus den Vektorraum \mathbb{K}^7 aufzufassen (in diesem Beispiel), dann erhalten wir alle Lösungen als Elemente einer Menge L , nämlich der Menge aller $[x_1, x_2, x_3, x_4, x_5, x_6, x_7] \in \mathbb{K}^7$, für die x_1 , x_3 , x_4 und x_7 wie oben als Funktionen der beliebig wählbaren x_2 , x_5 und x_6 bestimmt sind:

$$L = \left\{ \left[\begin{array}{c} 1 + 2x_2 - 7x_5 + 5x_6 \\ x_2 \\ 4 \quad + x_5 + 8x_6 \\ 6 \quad - 9x_5 - 7x_6 \\ x_5 \\ x_6 \\ 2 \end{array} \right] \mid x_2, x_5, x_6 \in \mathbb{K} \right\}.$$

Um die Lösungsstruktur zu erkennen, ergänzen wir die Koeffizienten, die Eins oder Null sind:

$$L = \left\{ \left[\begin{array}{c} 1 + 2x_2 - 7x_5 + 5x_6 \\ 0 + 1x_2 + 0x_5 + 0x_6 \\ 4 + 0x_2 + 1x_5 + 8x_6 \\ 6 + 0x_2 - 9x_5 - 7x_6 \\ 0 + 0x_2 + 1x_5 + 0x_6 \\ 0 + 0x_2 + 0x_5 + 1x_6 \\ 2 + 0x_2 + 0x_5 + 0x_6 \end{array} \right] \mid x_2, x_5, x_6 \in \mathbb{K} \right\}.$$

Berücksichtigen wir nun noch die Rechenregeln für Vektoren in \mathbb{K}^7 , so läßt sich das systematisch und übersichtlich folgendermaßen darstellen:

$$L = \left\{ \left[\begin{array}{c} 1 \\ 0 \\ 4 \\ 6 \\ 0 \\ 0 \\ 2 \end{array} \right] + x_2 \left[\begin{array}{c} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right] + x_5 \left[\begin{array}{c} -7 \\ 0 \\ 1 \\ -9 \\ 1 \\ 0 \\ 0 \end{array} \right] + x_6 \left[\begin{array}{c} 5 \\ 0 \\ 8 \\ -7 \\ 0 \\ 1 \\ 0 \end{array} \right] \mid x_2, x_5, x_6 \in \mathbb{K} \right\}.$$

Durch Einsetzen bestätigt man, daß $[1, 0, 4, 6, 0, 0, 2]$ eine Lösung des inhomogenen Gleichungssystems ist. Die Vektoren $[2, 1, 0, 0, 0, 0, 0]$, $[-7, 0, 1, -9, 1, 0, 0]$ und $[5, 0, 8, -7, 0, 1, 0]$ sind Lösungen des homogenen Systems

$$\begin{aligned} x_1 - 2x_2 + 7x_5 - 5x_6 &= 0 \\ x_3 - x_5 - 8x_6 &= 0 \\ x_4 + 9x_5 + 7x_6 &= 0 \\ x_7 &= 0. \end{aligned}$$

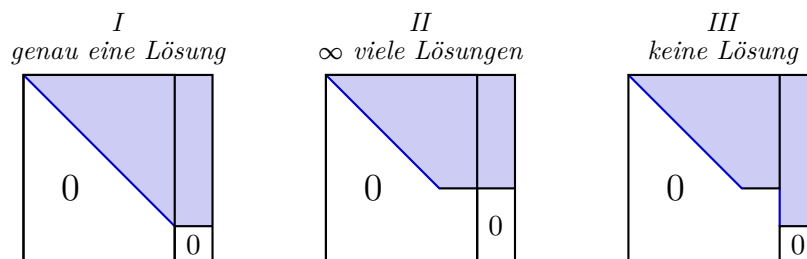
Diese Zweiteilung ist hier kein Zufall, sondern spiegelt die wesentliche Struktur der Lösungen linearer Gleichungssysteme wieder. Zum Schluß wollen wir uns noch klar machen, wie denn die eindeutige Lösbarkeit zu erkennen ist. Bei dem hier vorgestellten vollständigen GAUSS-Verfahren, bei dem nicht nur die Elemente unter der Diagonalen, sondern auch die oberhalb zum Verschwinden gebracht werden, ist das denkbar einfach: Das Verfahren muß bei einer voll besetzten Diagonalen enden. Die gesuchte Lösung steht dann auf der rechten Seite. Das könnte z. B. folgendermaßen aussehen

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 4 \\ x_3 &= 2 \\ x_4 &= 3, \end{aligned}$$

wenn wir ursprünglich von vier Gleichungen ausgegangen wären.

Es gibt drei typische Zustände für das GAUSS-Verfahren, die (eventuell nach Vertauschen von geeigneten Spalten) immer zu erreichen sind. An ihnen ist die *Anzahl* möglicher Lösungen sofort zu erkennen, auch wenn das für die Lösungen selbst i. A. noch nicht gilt. Es gibt durchaus Situationen, in denen diese Information bereits völlig ausreichend ist (vergl. 6.4)

Diese Zustände sind in der folgenden Skizze schematisch wiedergegeben. Die stark gezeichneten Linien meinen dabei Einträge $\neq 0$ während die unterlegten Flächen beliebige Einträge symbolisieren.



Ein Beispiel für die Situation *genau eine Lösung* ist etwa

$$\begin{aligned} x_1 - 2x_2 \quad \quad - 2x_4 &= 5 \\ x_2 + x_3 + 2x_4 &= 4 \\ x_3 + 3x_4 &= 2 \\ x_4 &= 3. \end{aligned}$$

Wenn man nur an der Anzahl der Lösungen interessiert ist, dann ist man mit dieser sog. *oberen Dreiecksform* fertig. Die letzte Gleichung legt nämlich x_4 als 3 fest. Das in die vorletzte Gleichung eingesetzt, ermöglicht es x_3 als $2 - 3 \cdot 3 = -7$ zu bestimmen. Auf diese Weise könnte man fortfahren und auch noch x_2 und x_1 ausrechnen. Wenn man die konkrete Lösung gar nicht wissen muß, begnügt man sich damit festzustellen, daß dieses Verfahren immer durchzuführen ist und immer zu genau einem Ergebnis führt.

Die Situation ∞ viele Lösungen könnte folgendermaßen aussehen:

$$\begin{aligned} x_1 + 2x_3 - x_4 \quad \quad + 7x_5 - 5x_6 - 2x_2 &= 2 \\ x_3 + 2x_4 \quad \quad - x_5 - 8x_6 &= 3 \\ x_4 \quad \quad + 9x_5 + 7x_6 &= 1 \\ x_7 &= 4. \end{aligned}$$

Hier ist x_7 durch die Zahl 4 bestimmt. x_4 läßt sich in Abhängigkeit von x_5 und x_6 berechnen: $x_4 = 1 - 9x_5 - 7x_6$. Mit dieser Lösung könnte man aus der zweiten Gleichung auf x_3 und damit endlich auf x_1 schließen. Auf diese Weise haben wir uns noch einmal klar gemacht, daß das Gleichungssystem auf jeden Fall lösbar ist. Da die Variablen x_2 , x_5 und x_6 aber nicht weiter bestimmt werden können, liefert *jede* ihrer Belegungen mit konkreten Zahlen eine Lösung. Das zeigt deutlich, daß es sich um unendlich viele Lösungen handelt. Wieder ist es ausreichend,

dieses Ergebnis an der oberen Dreiecksform abzulesen, wenn man sich nur über die Anzahl der Lösungen informieren will.

Schließlich die Situation *keine Lösung*. Sie ist leicht daran zu erkennen, daß die linke Seite einer Gleichung bei einer Umformung des GAUSS-Verfahrens komplett verschwindet, während die rechte Seite von Null verschieden bleibt.

6.2 Die Matrix zum LGS

Wir führen den Begriff *Matrix* und eine Multiplikation zwischen Matrix und Vektor ein, um die Lösungsstruktur linearer Gleichungssysteme besser herauszuarbeiten. Diese neuen Begriffe sind dabei zunächst nur bequeme Abkürzungen, die uns unsere Aufgabe erleichtern. Aber wie so oft in der Mathematik werden diese Dinge zusehends ein Eigenleben entwickeln, das weit über die ursprüngliche Aufgabenstellung hinausreicht.

Für das LGS (6.1) definieren wir die $m \times n$ -Matrix A der Koeffizienten $a_{11}, a_{12}, \dots, a_{mn}$ als das rechteckige Zahlenschema

$$A := \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}, \quad (6.2)$$

mit m Zeilen und n Spalten. Einen Vektor \mathbf{x} schreiben wir ab jetzt als Spaltenvektor, oder anders gesagt, als eine $n \times 1$ -Matrix:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Die sogenannte *Transponierte* A^t einer $m \times n$ -Matrix A entsteht, indem man die Zeilen zu Spalten macht, oder indem man die Einträge außerhalb der Diagonalen an dieser spiegelt. Man erhält die $n \times m$ -Matrix

$$A^t := \begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{bmatrix}. \quad (6.3)$$

Insbesondere ist also $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$.

Eine $m \times n$ -Matrix wird nach folgender Vorschrift mit einem Vektor $\mathbf{x} \in \mathbb{K}^n$ multipliziert

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} := \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}. \quad (6.4)$$

Der Ergebnisvektor ist aus \mathbb{K}^m . Man sieht, daß die Multiplikation gerade so konstruiert ist, daß die Einträge dieses Vektors die linke Seite des LGS (6.1) wiedergeben. Führen wir noch $\mathbf{b} := [b_1, b_2, b_3, \dots, b_m]^t \in \mathbb{K}^m$ ein, so läßt sich (6.1) in die übersichtliche Form

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (6.5)$$

bringen (wie üblich schreiben wir meist $\mathbf{A}\mathbf{x}$ statt $\mathbf{A} \cdot \mathbf{x}$). Ist die rechte Seite \mathbf{b} nicht der Nullvektor, so nennen wir das LGS *inhomogen*, andernfalls *homogen*. Eine Lösung \mathbf{x} des inhomogenen Systems heißt *inhomogene*, eine des homogenen *homogene Lösung*. Die folgende Regel läßt sich einfach nachrechnen:

$$\mathbf{A}(\mathbf{s}\mathbf{x} + \mathbf{t}\mathbf{y}) = \mathbf{s}\mathbf{A}\mathbf{x} + \mathbf{t}\mathbf{A}\mathbf{y} \quad (6.6)$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ und $s, t \in \mathbb{K}$, denn sie läßt sich auf das Distributivgesetz $a_{ij}(sx_j + ty_j) = sa_{ij}x_j + ta_{ij}y_j$ gewöhnlicher Zahlen $s, t, x_j, y_j, a_{ij} \in \mathbb{K}$ zurückführen. Diese Gleichung zeigt, daß die Multiplikation mit der Bildung von Linearkombinationen verträglich ist, in dem Sinne, daß das Produkt mit einer Linearkombination die Linearkombination der Produkte ist. Wir sprechen künftig in diesem Zusammenhang von der *Linearität* der Multiplikation von Matrix und Vektor.

Jetzt läßt sich die Lösungsstruktur eines inhomogenen LGSs leicht verstehen. Ist \mathbf{x}_1 eine inhomogene Lösung des LGSs und \mathbf{x}_0 eine homogene – es gilt also $\mathbf{A}\mathbf{x}_1 = \mathbf{b}$ und $\mathbf{A}\mathbf{x}_0 = \mathbf{0}$ – so ist $\mathbf{x}_1 + t\mathbf{x}_0$ wieder eine inhomogene Lösung:

$$\mathbf{A}(\mathbf{x}_1 + t\mathbf{x}_0) = \mathbf{A}\mathbf{x}_1 + t\mathbf{A}\mathbf{x}_0 = \mathbf{b} + t \cdot \mathbf{0} = \mathbf{b}.$$

Man kann zu einer inhomogenen Lösung beliebige homogene addieren, ohne die Eigenschaft inhomogene Lösung zu sein zu stören. Um es ganz deutlich zu sagen:

Eine beliebige Lösung \mathbf{x} eines inhomogenen LGSs läßt sich immer aus einer inhomogenen Lösung \mathbf{x}_1 durch Addition einer geeigneten homogenen Lösung \mathbf{x}_0 gewinnen: $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$.

Ist nämlich \mathbf{x} eine beliebige inhomogene Lösung, dann gilt $\mathbf{A}(\mathbf{x} - \mathbf{x}_1) = \mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_1 = \mathbf{b} - \mathbf{b} = \mathbf{0}$. Also ist $\mathbf{x}_0 := \mathbf{x} - \mathbf{x}_1$ eine homogene Lösung und $\mathbf{x} = \mathbf{x}_1 + \mathbf{x} - \mathbf{x}_1 = \mathbf{x}_1 + \mathbf{x}_0$.

Diese einfache Lösungsstruktur begegnet uns bei allen Arten linearer Gleichungen, etwa bei linearen Rekurrenzgleichungen 9, oder bei linearen Differentialgleichungen.

6.3 Das Schema zum Gauss-Verfahren

6.3.1 Beispiel Wir lösen das inhomogene LGS

$$\begin{array}{ll} \text{I} & 2x_1 + 5x_2 + 7x_3 + 3x_4 + 3x_5 = 0 \\ \text{II} & 3x_1 + 6x_2 + 6x_3 + 3x_4 + 3x_5 = 6 \\ \text{III} & 2x_1 + 4x_2 + 4x_3 + 2x_4 + 3x_5 = 3 \\ \text{IV} & x_1 + 3x_2 + 5x_3 + 2x_4 + 3x_5 = -3. \end{array}$$

Dabei stehen uns die folgenden *Zeilenumformungen* zur Verfügung (vergl. 6.1):

1. Multiplikation einer Zeile mit einer Zahl $a \neq 0$.
2. Addiere / subtrahiere eine Zeile zu / von einer anderen.
3. Vertausche zwei Zeilen / Spalten.

Hier könnte das bedeuten: Teile Gleichung II durch 3 und vertausche II mit I. Ziehe das Doppelte der Gleichung I von II ab. Wir werden das beim *GAUSS-Verfahren*, das wir an diesem Beispiel vorführen wollen, folgendermaßen notieren: $\text{II} : 3, \text{I} \leftrightarrow \text{II}, \text{II} - 2 \cdot \text{I}$, etc.

Bei allen Rechenoperationen sind letztlich nur die Koeffizienten der Variablen x_1, x_2, \dots, x_5 betroffen. Wenn wir etwa das Doppelte der vierten Gleichung von der dritten abziehen wollen, $\text{III} - 2 \cdot \text{IV}$, so erhalten wir

$$(2 - 2) \cdot x_1 + (4 - 6) \cdot x_2 + (4 - 10) \cdot x_3 + (2 - 4) \cdot x_4 + (3 - 6) \cdot x_5 = 3 + 6.$$

Das GAUSS-Verfahren erfordert viele dieser Rechenschritte. Um den Schreibaufwand dabei in Grenzen zu halten, läßt man die Variablen x_1, \dots, x_5 einfach weg und notiert das Ergebnis einer Umformung nur als Zahlenreihe

$$0 \quad -2 \quad -6 \quad -2 \quad -3 \quad 9.$$

Dabei steht an der ersten Position der Koeffizient von x_1 , an der zweiten der von x_2 usw. und an der letzten die rechte Seite der Gleichung. Daher besteht der erste Schritt beim GAUSS-Verfahren darin, das Gleichungssystem in ein Zahlenschema zu verwandeln, in dem nur noch die Koeffizienten auftauchen:

	x_1	x_2	x_3	x_4	x_5		
I	2	5	7	3	3	0	$\text{I} \leftrightarrow \text{II}$
II	3	6	6	3	3	6	$\text{II} : 3$
III	2	4	4	2	3	3	
IV	1	3	5	2	3	-3	
I	1	2	2	1	1	2	
II	2	5	7	3	3	0	$\text{II} - 2 \cdot \text{I}$
III	2	4	4	2	3	3	$\text{III} - 2 \cdot \text{I}$
IV	1	3	5	2	3	-3	$\text{IV} - \text{I}$

Nachdem Zeile I mit Zeile II vertauscht wurde, ist die neu entstandene Zeile I die Arbeitszeile. Mit ihr werden die anderen verändert. Dabei wird versucht, durch geschickte Addition oder Subtraktion geeigneter Vielfacher des ersten Elements die restlichen Spaltenelemente zum Verschwinden zu bringen.

	x_1	x_2	x_3	x_4	x_5	
I	1	2	2	1	1	2
II	0	1	3	1	1	-4
III	0	0	0	0	1	-1
IV	0	1	3	1	2	-5
						IV - II
I	1	0	-4	-1	-1	10
II	0	1	3	1	1	-4
III	0	0	0	0	1	-1
IV	0	0	0	0	1	-1
						I + III
						II - III
I	1	0	-4	-1	0	9
II	0	1	3	1	0	-3
III	0	0	0	0	1	-1

Anschließend ist Zeile II die Arbeitszeile, mit der die nicht verschwindenden Spalteneinträge 2 und 1 der zweiten Spalte beseitigt werden. Im vorletzten Schritt ist Zeile IV überflüssig geworden, da sie mit Zeile III übereinstimmt. Das Verfahren ist beendet, wenn jede weitere Zeilenumformung bereits beseitigte Einträge wieder zum Erscheinen bringen würde. Die Lösung des LGS erhalten wir, wenn wir das System in die Gleichungen zurück übersetzen, die es repräsentiert:

$$\begin{array}{rclcl}
 \text{I} & x_1 & -4x_3 - x_4 & = & 9 \\
 \text{II} & & x_2 + 3x_3 + x_4 & = & -3 \\
 \text{III} & & & & x_5 = -1.
 \end{array}$$

x_1, x_2 lassen sich in Abhängigkeit von x_3 und x_4 angeben. $x_5 = -1$ ist konstant. Damit können wir den Lösungsvektor $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5]^t$ folgendermaßen aufschreiben:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 9 + 4x_3 + 1x_4 \\ -3 - 3x_3 - 1x_4 \\ 0 + 1x_3 + 0x_4 \\ 0 + 0x_3 + 1x_4 \\ -1 + 0x_3 + 0x_4 \end{bmatrix} = \begin{bmatrix} 9 \\ -3 \\ 0 \\ 0 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Das bedeutet, die Lösungsmenge L ist

$$L = \left\{ \begin{bmatrix} 9 \\ -3 \\ 0 \\ 0 \\ -1 \end{bmatrix} + t \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \mid t, s \in \mathbb{R} \right\}.$$

6.3.2 A Finden Sie die Lösung für das Beispiel der Gießerei aus der Einführung

$$\begin{aligned}80x + 85y + 75z &= 81 \\15x + 12y + 10z &= 12 \\5x + 3y + 15z &= 7.\end{aligned}$$

6.3.3 A Finden Sie alle Lösungen des Gleichungssystems

$$\begin{aligned}2x + 3y - 16z + 14w &= 16 \\7x + 7y - 34z + 20w &= -1 \\-3x - y + 2z + 8w &= 33 \\31x + 22y - 94z + 14w &= -151.\end{aligned}$$

6.3.4 A Ein magisches Quadrat

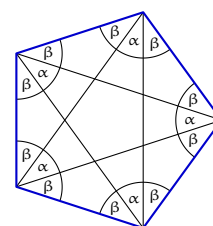
$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

hat gleiche Spalten-, Zeilen- und Diagonalsumme s und nur Einträge aus \mathbb{N} . Das führt auf die acht linearen Gleichungen

$$\begin{aligned}a + b + c &= s \\d + e + f &= s \\g + h + i &= s \\a + d + g &= s \\b + e + h &= s \\c + f + i &= s \\a + e + i &= s \\c + e + g &= s.\end{aligned}$$

Finden sie ein magisches Quadrat mit lauter verschiedenen Einträgen.

6.3.5 A Stellen Sie für die beiden Winkel α und β ein lineares Gleichungssystem auf und lösen Sie es.



6.3.6 Die Grenzen des GAUSS-Verfahrens Das GAUSS-Verfahren besteht im Wesentlichen in der geschickten Anwendung *äquivalenter Zeilenumformungen*: Multiplikation einer Zeile mit einer Zahl $c \neq 0$ und Addition des Vielfachen einer anderen Zeile. Sie heißen äquivalent, weil sich durch sie die Lösungsmenge des Systems nicht ändert. Das ist eine elementare Überlegung.

Weil sie aber so grundlegend ist, soll sie doch kurz vorgeführt werden. Wir gehen von einem Gleichungssystem

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1}x_1 & + & a_{i2}x_2 & + & \cdots & + & a_{in}x_n = b_i \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{k1}x_1 & + & a_{k2}x_2 & + & \cdots & + & a_{kn}x_n = b_k \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

aus und addieren das c -Fache der i -ten Gleichung zur k -ten:

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1}x_1 & + & a_{i2}x_2 & + & \cdots & + & a_{in}x_n = b_i \\ \vdots & & \vdots & & \vdots & & \vdots \\ (a_{k1} + ca_{i1})x_1 & + & (a_{k2} + ca_{i2})x_2 & + & \cdots & + & (a_{kn} + ca_{in})x_n = b_k + cb_i \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

Ist $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$ eine Lösung des ersten Gleichungssystems, dann löst \mathbf{x} auch das zweite. Dort hat sich ja nur die k -te Gleichung verändert, alle anderen behielten ihre ursprüngliche Form. Sie bleiben also weiterhin gültig. Es ist demnach nur zu zeigen, daß die neue k -te Gleichung erfüllt wird. Das ist simpel, denn dafür müssen wir ihre Terme nur geeignet umsortieren:

$$\begin{aligned} & (a_{k1} + ca_{i1})x_1 + (a_{k2} + ca_{i2})x_2 + \cdots + (a_{kn} + ca_{in})x_n \\ &= a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n + c(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) \\ &= b_k + c b_i. \end{aligned}$$

Nun sei \mathbf{x} eine Lösung des zweiten Systems. Jetzt müssen wir zeigen, daß die k -te Gleichung des ersten Systems erfüllt wird, denn die anderen sind ja weiterhin in beiden Systemen identisch. Insbesondere gilt die i -te Gleichung, so daß wir $0 = c(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) - cb_i$ in der folgenden Rechnung verwenden dürfen:

$$\begin{aligned} & a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n \\ &= a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n + c(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) - c b_i \\ &= (a_{k1} + ca_{i1})x_1 + (a_{k2} + ca_{i2})x_2 + \cdots + (a_{kn} + ca_{in})x_n - c b_i \\ &= b_k + c b_i + c b_i = b_k. \end{aligned}$$

Damit haben wir gezeigt, daß \mathbf{x} genau dann eine Lösung des ersten Systems ist, wenn \mathbf{x} eine Lösung des zweiten darstellt. Dasselbe gilt natürlich auch, wenn die Zeilenumformung nur in der Multiplikation mit einer Zahl $c \neq 0$ besteht.

Streng genommen dürfte das GAUSS-Verfahren in jedem Schritt nur in einer solchen äquivalenten Zeilenumformung bestehen. Wir haben aber mit einer Zeile i und einem Arbeitselement $a_{ir} \neq 0$ immer gleich alle Einträge a_{kr} oberhalb ($k < i$) und alle unterhalb ($k > i$) in einem Arbeitsgang eliminiert. Das ist jedoch gerechtfertigt, denn diesen Vorgang kann man sich leicht in

eine mehrfache Anwendung elementarer Zeilenumformungen aufgelöst denken, wie wir sie oben beschrieben haben. Dieses Vorgehen ist natürlich vernünftig, da das GAUSS-Verfahren sonst in endlose Schreibarbeit ausufern würde. Man kann sogar von diesem starren Verfahren abweichen und ‚Schwächen‘ eines Gleichungssystems ausnutzen, die darin bestehen können, daß durch geschickte Kombination von Gleichungen ein beträchtlicher Teil der Koeffizienten mit einem Mal zum Verschwinden gebracht werden können. Bei dem Gleichungssystem

	x_1	x_2	x_3	x_4	x_5	
I	3	1	0	0	0	1
II	1	1	6	1	1	1
III	4	4	24	2	3	3
IV	-1	-1	0	0	1	-3
						III - 4 II IV + II

bietet es sich natürlich an, nicht sofort mit dem Arbeitselement 3 der ersten Zeile zu arbeiten, sondern zunächst die günstige Lage der Koeffizienten in der zweiten, dritten und vierten Zeile auszunutzen:

I	3	1	0	0	0	1	
II	1	1	6	1	1	1	3 II - I
III	0	0	0	-2	-1	-1	III \leftrightarrow IV
IV	0	0	6	1	2	-2	
I	3	1	0	0	0	1	
II	0	2	18	3	3	2	
III	0	0	6	1	2	-2	
IV	0	0	0	-2	-1	-1	

Das ist solange gerechtfertigt, wie wir bei solchen Abkürzungen nur äquivalente Zeilenumformungen verwenden. Wenn man aber sehr geniale, weil sehr komplexe Möglichkeiten der Vereinfachung entdeckt, kann es leicht passieren, daß man den Überblick verliert und unbemerkt Zeilenumformungen durchführt, die nicht mehr äquivalent sind. Wie so etwas aussehen könnte, zeigt folgendes einfache Beispiel:

6.3.7 Beispiel

I	3	1	0	0	0	1	I + II
II	1	1	6	1	1	1	II + I
III	1	2	2	1	3	3	III - IV
IV	1	-1	0	0	1	-3	IV - III
I	4	2	6	1	1	2	
II	4	2	6	1	1	2	II - I
III	0	3	2	1	2	6	
IV	0	-3	-2	-1	-2	-6	IV + III
I	4	2	6	1	1	2	
III	0	3	2	1	2	6	

Wie durch ein Wunder sind wir auf einmal zwei Gleichungen losgeworden. Da wir bei diesem ‚Trick‘ aber gar nicht auf eine spezielle Form der Zeileneinträge angewiesen sind, sollten wir mißtrauisch werden und unser Verfahren noch einmal überdenken. Dabei ist es gar nicht so offensichtlich, wo die beiden Gleichungen eigentlich geblieben sind. Schließlich haben wir ja nur Gleichungen addiert und subtrahiert. Eine Lösung des Ausgangssystems ist daher auch eine des entstandenen Systems. Da dieses aber nur noch aus zwei Gleichungen besteht, kann es *mehr* Lösungen als das Ausgangssystem haben. Unsere Umformungen sind daher wohl nicht mehr äquivalent. Das erkennt man immer daran, daß es einem nicht gelingt, das Verfahren auf die äquivalenten Zeilenumformungen zurückzuführen: Zur ersten Gleichung haben wir die zweite addiert, was eine äquivalente Umformung wäre, würden wir nicht gleichzeitig die erste zur zweiten addieren. Das läßt sich nicht mehr in ein Nacheinander äquivalenter Zeilenumformungen auflösen. Dadurch, daß wir in einem ersten Schritt die zweite Gleichung zur ersten addieren, haben wir die erste verändert, so daß wir die zweite Zeilenumformung $II + I$ mangels unveränderter ersten Gleichung gar nicht mehr ausführen können.

Dieses Beispiel ist zugegebenermaßen sehr durchsichtig und der beschriebene Fehler daher leicht zu vermeiden. Das ist aber nicht immer so. Die Situation kann durchaus verwickelter sein und dann zu nicht äquivalenten Zeilenumformungen führen, die nicht mehr so leicht zu entdecken sind. Wir rechnen ein Beispiel auf zwei Weisen vollständig durch. Einmal, unter (zu) geschickter Ausnutzung von Abkürzungen und einmal nach dem Standardverfahren, wie wir es in Abschnitt 6.1 eingeführt haben.

6.3.8 Beispiel Wir lösen das homogene Gleichungssystem

$$\begin{aligned} 2x_1 + x_2 + 3x_3 + x_4 &= 0 \\ x_1 + 2x_2 + x_3 + 2x_4 + 3x_5 &= 0 \\ -x_1 + x_2 + 3x_4 + x_5 &= 0. \end{aligned}$$

Wie üblich machen wir das mit dem schematischen GAUSS-Verfahren:

I	2	1	3	1	0	I + 2 III	I	2	1	3	1	0	I + 2 III
II	1	2	1	2	3	2 II - I	II	1	2	1	2	3	II + III
III	-1	1	0	3	1	III + II	III	-1	1	0	3	1	
I	0	3	3	7	2	I + II - 2 III	I	0	3	3	7	2	I - II
II	0	3	-1	3	6		II	0	3	1	5	4	
III	0	3	1	5	4	III - II	III	-1	1	0	3	1	III \leftrightarrow I
I	0	0	0	0	0		I	-1	1	0	3	1	
II	0	3	-1	3	6	2 II + III	II	0	3	1	5	4	II - III/2
III	0	0	2	2	-2	III/2	III	0	0	2	2	-2	III/2
I	0	0	0	0	0		I	-1	1	0	3	1	-3 I + II
II	0	6	0	8	10	II/2	II	0	3	0	4	5	
III	0	0	1	1	-1		III	0	0	1	1	-1	
I	0	0	0	0	0		I	3	0	0	-5	2	
II	0	3	0	4	5		II	0	3	0	4	5	
III	0	0	1	1	-1		III	0	0	1	1	-1	

Das linke System führt auf die Lösungen

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1x_1 + 0x_4 + 0x_5 \\ 0x_1 - \frac{4}{3}x_4 - \frac{5}{3}x_5 \\ 0x_1 - 1x_4 + 1x_5 \\ 0x_1 + 1x_4 + 1x_5 \\ 0x_1 + 0x_4 + 1x_5 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \frac{x_4}{3} \begin{bmatrix} 0 \\ -4 \\ -3 \\ 3 \\ 0 \end{bmatrix} + \frac{x_5}{3} \begin{bmatrix} 0 \\ -5 \\ 3 \\ 0 \\ 3 \end{bmatrix},$$

mit beliebigen x_1 , x_4 und x_5 , während das rechte

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \frac{5}{3}x_4 - \frac{2}{3}x_5 \\ -\frac{4}{3}x_4 - \frac{5}{3}x_5 \\ -1x_4 + 1x_5 \\ 1x_4 + 0x_5 \\ 0x_4 + 1x_5 \end{bmatrix} = \frac{x_4}{3} \begin{bmatrix} 5 \\ -4 \\ -3 \\ 3 \\ 0 \end{bmatrix} + \frac{x_5}{3} \begin{bmatrix} -2 \\ -5 \\ 3 \\ 0 \\ 3 \end{bmatrix},$$

mit beliebigen x_4 und x_5 ergibt. Keiner der drei Basisvektoren $\mathbf{a} := [1, 0, 0, 0, 0]^t$, $\mathbf{b} := [0, -4, -3, 3, 0]^t$ und $\mathbf{c} := [0, -5, 3, 0, 3]^t$ des linken Schemas ist ein Lösung des Ausgangssystems, wie man durch Einsetzen leicht erkennt. Dagegen sind die Basisvektoren $\mathbf{d} := [5, -4, -3, 3, 0]^t$ und $\mathbf{e} := [-2, -5, 3, 0, 3]^t$ des rechten Schemas sehr wohl Lösungen der Aufgabe. Wie hängen die beiden Lösungsmengen zusammen? Die vermeintliche Lösungsmenge $\{r\mathbf{a} + s\mathbf{b} + t\mathbf{c} \mid r, s, t \in \mathbb{R}\}$, die das Verfahren auf der linken Seite erzeugt, umfaßt die tatsächliche Lösungsmenge $\mathcal{L} := \{u\mathbf{d} + v\mathbf{e} \mid u, v \in \mathbb{R}\}$, die das traditionelle GAUSS-Verfahren hervorbringt, denn es gilt $\mathbf{d} = 5\mathbf{a} + \mathbf{b}$ und $\mathbf{e} = -2\mathbf{a} + \mathbf{c}$. Sie hat aber mehr Elemente, denn \mathbf{a} ist offensichtlich keine Lösung des Ausgangssystems und kann daher auch nicht in \mathcal{L} liegen. Der Grund für dieses Ergebnis ist nicht in den Umformungen I+2 III und 2 II-I der ersten beiden Gleichungen zu suchen, denn das kann in äquivalente Zeilenumformungen aufgelöst werden: Zuerst 2 II-I und dann I+2 III. Das Problem macht in der dritten Gleichung der Schritt III+II, sollte man annehmen. Die Kombination mit 2 II-I läßt sich aber wieder durch äquivalente Umformungen erreichen: Zuerst III+II und dann 2 II-I. Dasselbe gilt für die erste und die dritte Zeilenumformung: Zuerst I+2 III und dann III+II. Was das Problem erzeugt ist demnach die Kombination aller drei Umformungen in einem Arbeitsschritt. Man kann sich vorstellen, daß man bei mehr Gleichungen noch viel undurchsichtigere Konstellationen produzieren kann (ein neuer Gesichtspunkt dazu ist im Beispiel 6.4.8 zu finden). Sie alle haben aber ein Erkennungsmerkmal, nämlich, daß nicht klar ist, welche Zeile die eigentliche Arbeitszeile ist, mit der andere Zeilen verändert werden. Und das ist auch schon die Strategie, mit der solche Situationen einfach zu vermeiden sind: Die Arbeitszeile darf im gleichen Arbeitsschritt nie selbst verändert werden (allenfalls darf sie mit einer geeigneten Zahl $\alpha \neq 0$ multipliziert werden, um dafür nicht das Gleichungssystem noch einmal aufschreiben zu müssen).

6.4 Lineare Unabhängigkeit

6.4.1 Definition Eine Teilmenge $\mathcal{B} \subset V$ des Vektorraums V über \mathbb{K} heißt linear unabhängig (mitunter durch l. u. abgekürzt), falls jede Nullkombination aus \mathcal{B} , d. h. jede Gleichung der Form

$$\lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \cdots + \lambda_n \mathbf{b}_n = \mathbf{0}, \quad (6.7)$$

mit Vektoren $\mathbf{b}_1, \dots, \mathbf{b}_n$ von \mathcal{B} und Koeffizienten $\lambda_1, \dots, \lambda_n$ aus \mathbb{K} nur durch $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ zu erfüllen ist. Eine solche Nullkombination bezeichnen wir als trivial. Damit können wir die lineare Unabhängigkeit von \mathcal{B} kurz und prägnant folgendermaßen ausdrücken: \mathcal{B} ist genau dann linear unabhängig, wenn jede Nullkombination aus \mathcal{B} trivial ist. \mathcal{B} heißt linear abhängig (l. a.), falls \mathcal{B} nicht linear unabhängig ist.

Enthält $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ nur endlich viele Elemente, dann läßt sich die lineare Unabhängigkeit durch Lösen einer einzigen linearen (Vektor-)Gleichung überprüfen, nämlich $\lambda_1 \mathbf{b}_1 + \cdots + \lambda_n \mathbf{b}_n = \mathbf{0}$. Diese deckt alle möglichen Nullkombinationen ab. Für $V = \mathbb{C}^n$ ist sie zu einem homogenen linearen Gleichungssystem aus n Gleichungen äquivalent, das aus den n Vektorkomponenten entsteht. Es wird mit dem GAUSS-Verfahren gelöst. Betrachten wir das folgende Beispiel.

6.4.2 Beispiel

$$\mathcal{B} := \left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 1 \\ 9 \end{bmatrix} \right\}.$$

Die Nullkombination

$$\lambda_1 \begin{bmatrix} 1 \\ 3 \\ 2 \\ 3 \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ 2 \\ 2 \\ 7 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 4 \\ 1 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

führt auf das homogene lineare Gleichungssystem

$$\begin{array}{rrcr} \lambda_1 & - & \lambda_2 & = & 0 \\ 3\lambda_1 & + & 2\lambda_2 & + & 4\lambda_3 = 0 \\ 2\lambda_1 & + & 2\lambda_2 & + & \lambda_3 = 0 \\ 3\lambda_1 & + & 7\lambda_2 & + & 9\lambda_3 = 0, \end{array}$$

das wir wie üblich behandeln – allerdings nur soweit, bis wir die Anzahl der Lösungen erkennen können, denn die Lösungen selbst sind normalerweise uninteressant. Sind die Vektoren nämlich linear unabhängig, dann ist die Lösung der Nullvektor $[\lambda_1, \lambda_2, \lambda_3]^t = \mathbf{0}$. Da $\mathbf{0}$ immer eine Lösung darstellt, sind die Vektoren genau dann linear unabhängig, wenn es nur eine Lösung

(nämlich $\mathbf{0}$) gibt. Das ist zweifelsfrei an einer oberen Dreiecksform und nicht verschwindenden Diagonaleinträgen über die gesamte Breite des Gleichungssystems abzulesen (vergl. Seite 104):

	λ_1	λ_2	λ_3	
I	1	-1	0	
II	3	2	4	II - 3I
III	2	3	1	III - 2I
IV	3	7	9	IV - 3I
I	1	-1	0	
II	0	5	4	
III	0	5	1	III - II
IV	0	10	9	IV - 2II
I	1	-1	0	
II	0	5	4	
III	0	0	-3	
IV	0	0	1	

Zeile III als Gleichung zurück übersetzt lautet einfach $-3\lambda_3 = 0$, also $\lambda_3 = 0$. Das in die Gleichung für Zeile II eingesetzt ergibt $5\lambda_2 + 4\lambda_3 = 5\lambda_2 = 0$, d. h. $\lambda_2 = 0$. Diese Ergebnisse in die erste Gleichung $\lambda_1 - \lambda_2 = 0$ eingesetzt ergibt schließlich auch $\lambda_1 = 0$. Damit ist nachgerechnet, daß eine Nullkombination der Vektoren $[1, 3, 2, 3]^t$, $[-1, 2, 3, 7]$ und $[0, 4, 1, 9]^t$ trivial sein muß. Sie sind damit linear unabhängig.

Diese Rechnung haben wir hier noch einmal zur Demonstration vorgeführt, um daran zu erinnern, wieso eine vollbesetzte Diagonale zur linearen Unabhängigkeit der beteiligten Vektoren führt. In einer konkreten Rechnung ist das aber nicht mehr nötig, da die Überlegung jedes mal nach genau diesem Schema verläuft und nie mit Überraschungen aufwarten wird.

6.4.3 Beispiel

$$\mathcal{C} := \left\{ \begin{bmatrix} 1 \\ -1 \\ 6 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -11 \\ -2 \\ 2 \\ 19 \\ 7 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 10 \\ 6 \\ 21 \\ -9 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 5 \\ -4 \\ 0 \end{bmatrix} \right\}.$$

Eine Nullkombination dieser Vektoren führt, wie im Beispiel 6.4.2, auf das Schema

	λ_1	λ_2	λ_3	λ_4	λ_5	
I	1	-3	10	2	-11	
II	-1	2	6	3	-2	II + I
III	6	0	21	5	2	III - 6I
IV	4	1	-9	-4	19	IV - 4I
V	2	1	1	0	7	V - 2I

	λ_1	λ_2	λ_3	λ_4	λ_5	
I	1	-3	10	2	-11	
II	0	-1	16	5	-13	
III	0	18	-39	-7	68	III + 18 II
IV	0	13	-49	-12	63	IV + 13 II
V	0	7	-19	-4	29	V + 7 II
I	1	-3	10	2	-11	
II	0	-1	16	5	-13	
III	0	0	249	83	-166	III/83
IV	0	0	159	53	-106	IV/53
V	0	0	93	31	-62	V/31
I	1	-3	10	2	-11	
II	0	-1	16	5	-13	
III	0	0	3	1	-2	
IV	0	0	3	1	-2	IV - III
V	0	0	3	1	-2	V - III
I	1	-3	10	2	-11	
II	0	-1	16	5	-13	
III	0	0	3	1	-2	
IV	0	0	0	0	0	
V	0	0	0	0	0	

Die dritte Zeile bedeutet $3\lambda_3 + \lambda_4 - 2\lambda_5 = 0$. Sie ist z. B. für $\lambda_4 = -3$, $\lambda_5 = 0$ und $\lambda_3 = 1$ zu erfüllen. Eingesetzt in die Gleichung $-\lambda_2 + 16\lambda_3 + 5\lambda_4 - 13\lambda_5 = 0$ für Zeile II erhalten wir $\lambda_2 = 1$. Das in die Gleichung $\lambda_1 - 3\lambda_2 + 10\lambda_3 + 2\lambda_4 - 11\lambda_5 = 0$ für die erste Zeile eingesetzt liefert schließlich $\lambda_1 = -1$. Mit

$$-\begin{bmatrix} 1 \\ -1 \\ 6 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 10 \\ 6 \\ 21 \\ -9 \\ 1 \end{bmatrix} - 3 \begin{bmatrix} 2 \\ 3 \\ 5 \\ -4 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} -11 \\ -2 \\ 2 \\ 19 \\ 7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

haben wir daher eine nicht triviale Nullkombination der Vektoren von \mathcal{C} gefunden. Es gibt sogar noch weitere, aber für die lineare Abhängigkeit von Vektoren ist es nur wichtig *eine* zu finden. Es ist sogar ausreichend zweifelsfrei nachzuweisen, daß es eine geben muß. Dafür ist es normalerweise gar nicht nötig, sie tatsächlich auszurechnen. Hier genügt es, wenn man das GAUSS-Verfahren bis zur oberen Dreiecksform durchführt und aus dessen stumpfer Form auf die Existenz nicht trivialer Nullkombinationen zu schließen (vergl. Situation II auf Seite 104).

6.4.4 Beispiel Enthält $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n, \dots\}$ abzählbar unendlich viele Elemente, dann muß man die Gleichungen $\lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n = \mathbf{0}$ für alle $n \in \mathbb{N}$ lösen. Sollte allerdings nur eine dieser Gleichungen eine nicht triviale Lösung aufweisen, so ist \mathcal{B} linear abhängig. Wir wählen als Beispiel im Vektorraum \mathcal{P} aller reellen (oder komplexen) Polynome die Menge \mathcal{B} aller *Monome*, also aller Polynome p_k , die durch $p_k(x) := x^k$ definiert sind: $\mathcal{B} := \{p_k \mid k \in \mathbb{N}_0\}$.

Wir nehmen an, es gäbe eine nicht triviale Nullkombination $\lambda_0 p_0 + \lambda_1 p_1 + \dots + \lambda_n p_n = \sum_{k=0}^n \lambda_k p_k = 0$. Durch $p := \sum_{k=0}^n \lambda_k p_k$ ist insbesondere ein Polynom $p \in \mathcal{P}$ definiert: $p(x) = \sum_{k=0}^n \lambda_k p_k(x) = \sum_{k=0}^n \lambda_k x^k$. Laut unserer Annahme hätte es die merkwürdige Eigenschaft $p(x) = 0$ für alle $x \in \mathbb{R}$, d. h. dieses Polynom ist das sogenannte *Nullpolynom*. Wir müssen zeigen, daß alle Koeffizienten λ_k verschwinden. Das folgt aus einem etwas allgemeineren Ergebnis, nämlich der Eindeutigkeit der Koeffizienten eines Polynoms (Näheres im Abschnitt 11.8 über Polynome, Satz 11.8.2). Da das Nullpolynom mit den Koeffizienten $\lambda_k = 0$ darstellbar ist, gibt es gar keine anderen. Das heißt, es muß $\lambda_k = 0$ für $k = 0, \dots, n$ gelten. Damit haben wir nachgewiesen, daß \mathcal{B} linear unabhängig ist.

Die folgende Übung zeigt, daß es auch einen anderen Weg gibt die lineare Unabhängigkeit von $\{p_k \mid k \in \mathbb{N}_0\}$ zu überprüfen, wenn man dafür auf die elementare Ableitungsregel $p'_k(x) = k x^{k-1}$ für Polynome zurückgreift.

6.4.5 A Zeigen Sie, daß $\mathcal{B} := \{p_k \mid k \in \mathbb{N}_0\}$ linear unabhängig ist. Gehen Sie dabei folgendermaßen vor:

Starten Sie mit einer Nullkombination $0 = \lambda_0 p_0 + \lambda_1 p_1 + \dots + \lambda_n p_n$. Durch wiederholtes Ableiten und Auswerten an der Stelle $x = 0$ erhalten Sie daraus ein lineares Gleichungssystem für die λ_i , das bereits in Diagonalform vorliegt.

6.4.6 Beispiel. Eine linear unabhängige Menge kann sogar mehr als nur abzählbar unendlich viele Elemente enthalten. Bezeichnen wir für jedes $0 \neq t \in \mathbb{R}$ mit e_t die Exponentialfunktionen $e_t(x) := e^{tx}$, so ist die Menge $\mathcal{E} := \{e_t \mid t \in \mathbb{R} \setminus \{0\}\}$ linear unabhängig. Offensichtlich enthält \mathcal{E} so viele Elemente wie $\mathbb{R} \setminus \{0\}$ (und \mathbb{R} läßt sich nicht mehr abzählen, wie wir aus den Aufgaben wissen $\rightarrow \dots$). Um das einzusehen, müssen wir zeigen, daß jede Nullkombination $\lambda_1 e_{t_1} + \lambda_2 e_{t_2} + \dots + \lambda_n e_{t_n} = 0$ trivial ist. Dabei sind die Faktoren t_i paarweise verschieden. Machen wir eine kleine Anleihe an unser Schulwissen, nämlich $e'_t(x) = t e^{tx}$ (vergl. (11.40)). Damit läßt sich die Nullkombination beliebig oft ableiten. Wir erhalten $t_1 \lambda_1 e_{t_1} + t_2 \lambda_2 e_{t_2} + \dots + t_n \lambda_n e_{t_n} = 0$, dann $t_1^2 \lambda_1 e_{t_1} + t_2^2 \lambda_2 e_{t_2} + \dots + t_n^2 \lambda_n e_{t_n} = 0$ und nach der k -ten Ableitung $t_1^k \lambda_1 e_{t_1} + t_2^k \lambda_2 e_{t_2} + \dots + t_n^k \lambda_n e_{t_n} = 0$. An einer Stelle $x \in \mathbb{R}$ ausgewertet bedeutet das

$$t_1^k \lambda_1 e^{t_1 x} + t_2^k \lambda_2 e^{t_2 x} + \dots + t_n^k \lambda_n e^{t_n x} = 0.$$

Wählen wir dabei die Stelle $x = 0$, so erhalten wir wegen $e^0 = 1$ das folgende System linearer Gleichungen für die Unbekannten λ_i :

$$\begin{aligned} t_1 \lambda_1 + t_2 \lambda_2 + \dots + t_n \lambda_n &= 0 \\ t_1^2 \lambda_1 + t_2^2 \lambda_2 + \dots + t_n^2 \lambda_n &= 0 \\ &\vdots \\ t_1^n \lambda_1 + t_2^n \lambda_2 + \dots + t_n^n \lambda_n &= 0. \end{aligned}$$

Die Lösung erfordert eine geschickte Anwendung des GAUSS-Verfahrens. Wir führen das aber nicht für allgemeines n vor. Die Idee ist bereits für $n = 4$ zu erkennen und kann dann leicht

verallgemeinert werden:

	λ_1	λ_2	λ_3	λ_4	
I	t_1	t_2	t_3	t_4	
II	t_1^2	t_2^2	t_3^2	t_4^2	II - t_1 I
III	t_1^3	t_2^3	t_3^3	t_4^3	III - t_1 II
IV	t_1^4	t_2^4	t_3^4	t_4^4	IV - t_1 III
I	t_1	t_2	t_3	t_4	
II	0	$t_2(t_2 - t_1)$	$t_3(t_3 - t_1)$	$t_4(t_4 - t_1)$	
III	0	$t_2^2(t_2 - t_1)$	$t_3^2(t_3 - t_1)$	$t_4^2(t_4 - t_1)$	III - t_2 II
IV	0	$t_2^3(t_2 - t_1)$	$t_3^3(t_3 - t_1)$	$t_4^3(t_4 - t_1)$	IV - t_2 III
I	t_1	t_2	t_3	t_4	
II	0	$t_2(t_2 - t_1)$	$t_3(t_3 - t_1)$	$t_4(t_4 - t_1)$	
III	0	0	$t_3(t_3 - t_2)(t_3 - t_1)$	$t_4(t_4 - t_2)(t_4 - t_1)$	
IV	0	0	$t_3^2(t_3 - t_2)(t_3 - t_1)$	$t_4^2(t_4 - t_2)(t_4 - t_1)$	IV - t_3 III
I	t_1	t_2	t_3	t_4	
II	0	$t_2(t_2 - t_1)$	$t_3(t_3 - t_1)$	$t_4(t_4 - t_1)$	
III	0	0	$t_3(t_3 - t_2)(t_3 - t_1)$	$t_4(t_4 - t_2)(t_4 - t_1)$	
IV	0	0	0	$t_4(t_4 - t_3)(t_4 - t_2)(t_4 - t_1)$	

Das ist eine obere Dreiecksform mit nicht verschwindender Diagonalen. Daher hat unser Gleichungssystem die einzig mögliche Lösung $\lambda_1 = \dots = \lambda_n = 0$. Das bedeutet, daß alle Nullkombinationen in \mathcal{E} trivial sind. \mathcal{E} ist also, wie behauptet, linear unabhängig.

Wir wollen noch die Verbindung unserer Definition der linearen Unabhängigkeit mit einer gängigen Vorstellung aufzeigen. Linear unabhängige Vektoren sollten die Eigenschaft haben, daß sich keiner der beteiligten durch die anderen linear kombinieren läßt.

6.4.7 Lemma Eine Teilmenge \mathcal{B} eines Vektorraumes V ist genau dann linear unabhängig, wenn kein Vektor aus \mathcal{B} eine Linearkombination der übrigen Vektoren aus \mathcal{B} ist.

Beweis. Ist \mathcal{B} linear unabhängig und $\mathbf{b} \in \mathcal{B}$ eine Linearkombination von anderen Vektoren $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathcal{B}$, also $\mathbf{b} = \lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n$, dann würde $\lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n - 1 \cdot \mathbf{b} = \mathbf{0}$ eine nicht triviale Nullkombination mit Vektoren aus \mathcal{B} ergeben, im Widerspruch zur linearen Unabhängigkeit von \mathcal{B} . Daher kann kein Element von \mathcal{B} eine Linearkombination der übrigen Elemente aus \mathcal{B} sein.

Für die Umkehrung gehen wir davon aus, daß kein Element aus \mathcal{B} eine Linearkombination der übrigen Elemente von \mathcal{B} sein kann und nehmen an, $\lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \dots + \lambda_n \mathbf{b}_n = \mathbf{0}$ wäre eine nicht triviale Nullkombination von Vektoren $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n \in \mathcal{B}$. Das bedeutet, daß wenigstens einer der Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_n$ von Null verschieden sein müßte. Wir können o. B. d. A. $\lambda_1 \neq 0$ annehmen. Dann ließe sich die Nullkombination nach \mathbf{b}_1 auflösen: $\mathbf{b}_1 = -\frac{\lambda_2}{\lambda_1} \mathbf{b}_2 - \dots - \frac{\lambda_n}{\lambda_1} \mathbf{b}_n$. \mathbf{b}_1 wäre eine Linearkombination der übrigen Vektoren aus \mathcal{B} , im Widerspruch zu unserer Annahme. \square

6.4.8 Beispiel Mit den Ergebnissen unserer Untersuchungen zur linearen Unabhängigkeit können wir jetzt noch etwas besser verstehen, was der eigentliche Grund dafür war, daß die Modifikation des GAUSS-Verfahrens auf Seite 112 falsche Ergebnisse lieferte. Dazu fassen wir die drei Zeilen des Gleichungssystems als Zeilenvektoren z_1, z_2 und z_3 auf. Im ersten Arbeitsschritt haben wir aus $\mathcal{Z}_1 := \{z_1, z_2, z_3\}$ die Menge $\mathcal{Z}_2 := \{z_1 + 2z_3, 2z_2 - z_1, z_3 + z_2\}$ mit den neuen Zeilenvektoren $\tilde{z}_1 := z_1 + 2z_3, \tilde{z}_2 := 2z_2 - z_1$ und $\tilde{z}_3 := z_3 + z_2$ des zweiten Gleichungssystems gemacht. Egal ob \mathcal{Z}_1 linear unabhängig ist oder nicht, die Menge \mathcal{Z}_2 ist sicher linear abhängig, denn $1 \cdot (z_1 + 2z_3) + 1 \cdot (2z_2 - z_1) - 2 \cdot (z_3 + z_2)$ ist eine nicht triviale Nullkombination ihrer Elemente. Das bedeutet, wenn man die zweite Zeile \tilde{z}_2 zur ersten \tilde{z}_1 addiert und das Doppelte der Dritten \tilde{z}_3 von ihr abzieht, kann man sie zum Verschwinden bringen, unabhängig davon, was ursprünglich in z_1, z_2 und z_3 stand. Wäre also der Übergang von \mathcal{Z}_1 zu \mathcal{Z}_2 durch äquivalente Zeilenumformungen zustande gekommen, dann hätten wir eine universelle Methode entdeckt, wie wir in jedem Gleichungssystem eine Zeile zum Verschwinden bringen können. Wiederholte Anwendung unseres Verfahrens würde schließlich dafür sorgen, daß wir jedes Gleichungssystem auf zwei Gleichungen reduzieren könnten. Das ist offensichtlich unsinnig, denn dann wäre kein Gleichungssystem mit mehr als zwei Unbekannten eindeutig lösbar.

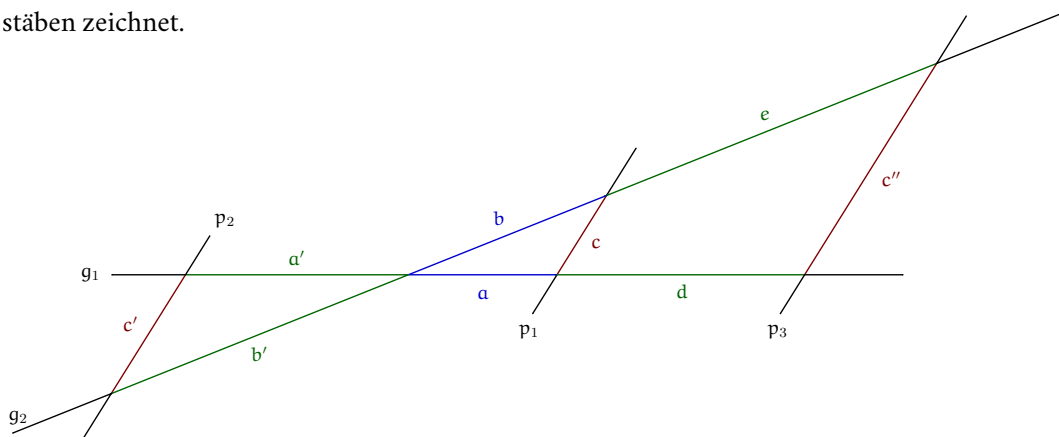
6.4.9 A

$$\text{i) } \mathcal{B} := \left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \\ 12 \end{bmatrix}, \begin{bmatrix} -10 \\ 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 10 \\ 11 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \\ -2 \\ 5 \end{bmatrix} \right\}, \quad \mathcal{C} := \left\{ \begin{bmatrix} 1 \\ 14 \\ 19 \\ -60 \\ 0 \end{bmatrix}, \begin{bmatrix} 9 \\ -3 \\ -12 \\ 63 \\ -87 \end{bmatrix}, \begin{bmatrix} 11 \\ 5 \\ 18 \\ -45 \\ -19 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 6 \\ -16 \\ -4 \end{bmatrix} \right\}.$$

Zeigen Sie, daß \mathcal{B} linear unabhängig und \mathcal{C} linear abhängig ist.

*ii) $\mathcal{S} := \{s_t \mid t \in \mathbb{R}_+\}$ mit $s_t(x) := \sin(tx)$ ist linear unabhängig.

6.4.10 Die Strahlensätze Eine schöne Anwendung des Begriffs der linearen Unabhängigkeit sind die Strahlensätze der Geometrie: Parallele Geraden p_1, p_2 und p_3 unterteilen ein Geradenbündel g_1, g_2 in Abschnitte a, a', b, b', c, c' und $a + d, b + e$, die in gleichen Verhältnis zueinander stehen. Das ist die Begründung für die nicht weiter verwunderliche Tatsache, daß sich die Seitenverhältnisse in einem Dreieck nicht ändern, wenn man es in verschiedenen Maßstäben zeichnet.

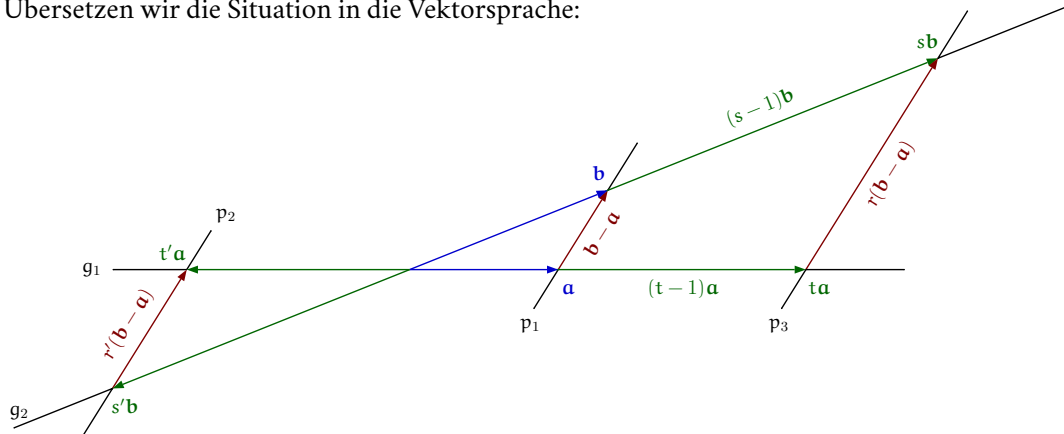


Der Strahlensatz für die Situation der Skizze besagt:

$$\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}, \quad \frac{a}{b} = \frac{a'}{b'} = \frac{a+d}{b+e} = \frac{d}{e}, \quad \frac{a}{c} = \frac{a'}{c'}, \quad \frac{b}{c} = \frac{b'}{c'}.$$

Aus $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$ folgt $\frac{a}{b} = \frac{a'}{b'}$, sowie $\frac{a}{c} = \frac{a'}{c'}$ und $\frac{b}{c} = \frac{b'}{c'}$. Die Gleichung $\frac{a}{b} = \frac{d}{e}$, die etwas aus dem Rahmen zu fallen scheint, folgt einfach aus der Anwendung des Strahlensatzes auf die rechte Seite der Skizze: $\frac{a}{b} = \frac{a+d}{b+e}$, also $a(b+e) = b(a+d)$, oder $ae = bd$.

Übersetzen wir die Situation in die Vektorsprache:



Wir legen den Ursprung in den Schnittpunkt von g_1 und g_2 und wählen die linear unabhängigen Vektoren \mathbf{a} und \mathbf{b} als ihre Richtungsvektoren. Die parallelen Geraden p_1 , p_2 und p_3 haben denselben Richtungsvektor $\mathbf{b} - \mathbf{a}$. Für das linke Dreieck gilt $t'\mathbf{a} = s'\mathbf{b} + r'(\mathbf{b} - \mathbf{a})$, also $(t' + r')\mathbf{a} + (s' - r')\mathbf{b} = \mathbf{0}$. Das ist eine Gleichung für die beiden Unbekannten r' und s' . Allerdings ist es eine Vektorgleichung, genauer ist es eine *Nullkombination* der beiden linear unabhängigen Vektoren \mathbf{a} und \mathbf{b} . Da diese trivial sein muß, folgt einfach $t' + r' = 0$ und $s' - r' = 0$, d. h. $t' = -r' = s'$. \mathbf{b} muß also mit demselben Faktor t' gestreckt werden wie \mathbf{a} , um die Schnittpunkte von p_2 mit g_2 und g_1 zu liefern. Das Verhältnis $\frac{a}{b}$ der Seitenlängen $a := \|\mathbf{a}\|$ und $b := \|\mathbf{b}\|$ stimmt offensichtlich mit dem der Längen $a' := \|t'\mathbf{a}\|$ und $b' := \|t'\mathbf{b}\|$ überein:

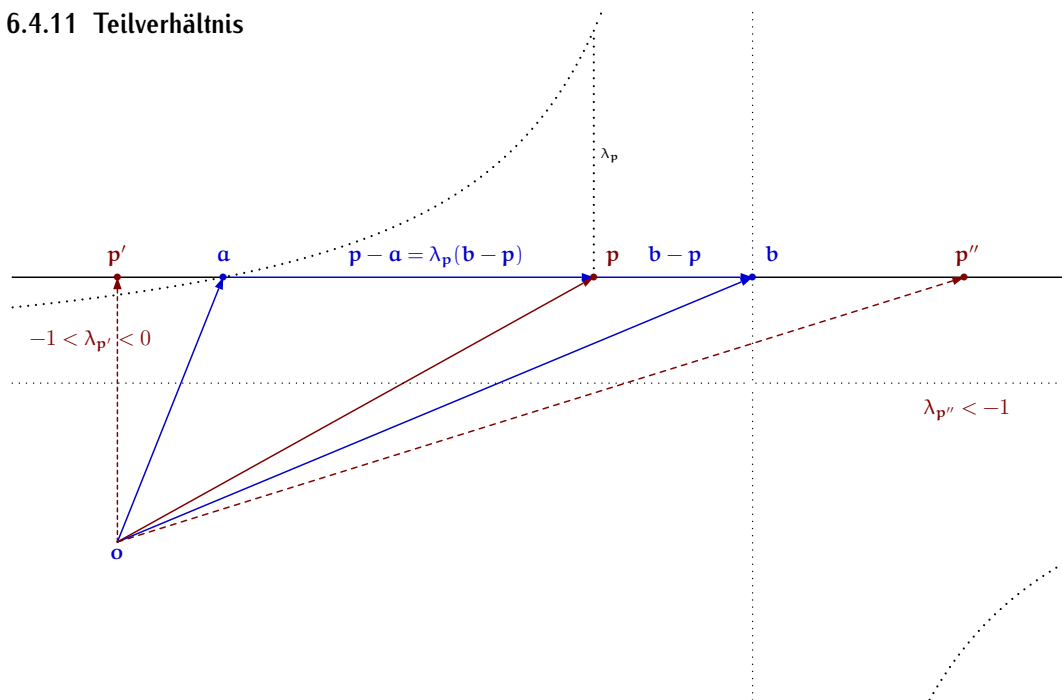
$$\frac{a'}{b'} = \frac{\|t'\mathbf{a}\|}{\|t'\mathbf{b}\|} = \frac{|t'| \|\mathbf{a}\|}{|t'| \|\mathbf{b}\|} = \frac{\|\mathbf{a}\|}{\|\mathbf{b}\|} = \frac{a}{b}.$$

Mit $c := \|\mathbf{b} - \mathbf{a}\|$ und $c' := \|t'(\mathbf{b} - \mathbf{a})\|$ ergeben sich die verbleibenden Verhältnisse genauso einfach, etwa

$$\frac{a'}{c'} = \frac{\|t'\mathbf{a}\|}{\|t'(\mathbf{b} - \mathbf{a})\|} = \frac{|t'| \|\mathbf{a}\|}{|t'| \|\mathbf{b} - \mathbf{a}\|} = \frac{\|\mathbf{a}\|}{\|\mathbf{b} - \mathbf{a}\|} = \frac{a}{c}.$$

Für die Dreiecke der rechten Seite folgert man wie oben $t = s$. Mit $d := \|(t-1)\mathbf{a}\|$ und $e := \|(t-1)\mathbf{b}\|$ zeigt sich dann auch $\frac{a}{b} = \frac{d}{e}$.

6.4.11 Teilverhältnis



Für das Geradenstück zwischen \mathbf{a} und \mathbf{b} wählen wir (in Anlehnung an die Intervall-Notation in \mathbb{R}) die Schreibweise (\mathbf{a}, \mathbf{b}) . Sollen die Randpunkte enthalten sein, dann schreiben wir $[\mathbf{a}, \mathbf{b}]$.

Für einen Punkt \mathbf{p} auf der Geraden durch \mathbf{a} und \mathbf{b} , der weder mit \mathbf{a} noch mit \mathbf{b} übereinstimmt, definieren wir das *Teilverhältnis* $T(\mathbf{a}, \mathbf{p}, \mathbf{b})$ für die Teilung von (\mathbf{a}, \mathbf{b}) durch \mathbf{p} über die Gleichung

$$\mathbf{p} - \mathbf{a} = T(\mathbf{a}, \mathbf{p}, \mathbf{b})(\mathbf{b} - \mathbf{p}). \quad (6.8)$$

Wenn \mathbf{a} und \mathbf{b} fest gegeben sind, schreiben wir auch einfach $\lambda_{\mathbf{p}}$ statt $T(\mathbf{a}, \mathbf{p}, \mathbf{b})$. Es gilt

$$|\lambda_{\mathbf{p}}| = \frac{\|\mathbf{p} - \mathbf{a}\|}{\|\mathbf{b} - \mathbf{p}\|},$$

d. h. bis auf ein mögliches Vorzeichen ist $\lambda_{\mathbf{p}}$ das Verhältnis der Längen der von \mathbf{p} erzeugten Teilstrecken. Liegt \mathbf{p} zwischen \mathbf{a} und \mathbf{b} , d. h. gilt $\mathbf{p} \in (\mathbf{a}, \mathbf{b})$, dann ist $\lambda_{\mathbf{p}} > 0$. Liegt es in Richtung $\mathbf{a} - \mathbf{b}$ außerhalb von (\mathbf{a}, \mathbf{b}) (in der Skizze links von \mathbf{a}), dann ist das an $-1 < \lambda_{\mathbf{p}} < 0$ zu erkennen und liegt es auf der anderen Seite, an $\lambda_{\mathbf{p}} < -1$.

Ein klassisches Beispiel ist der *goldene Schnitt*:

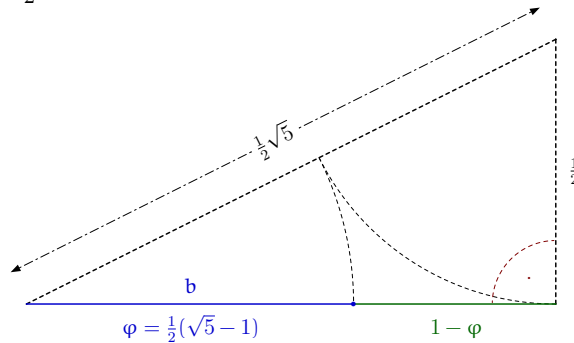
Eine Strecke wird im Verhältnis des goldenen Schnitts geteilt, wenn sich die kürzere Strecke zur längeren wie die längere zur ganzen verhält.

Wenn also eine Strecke der Länge a durch Teilstrecken der Längen b und $a - b < b$ in diesem Verhältnis geteilt werden soll, muß $\varphi := \frac{b}{a}$ die Gleichungen

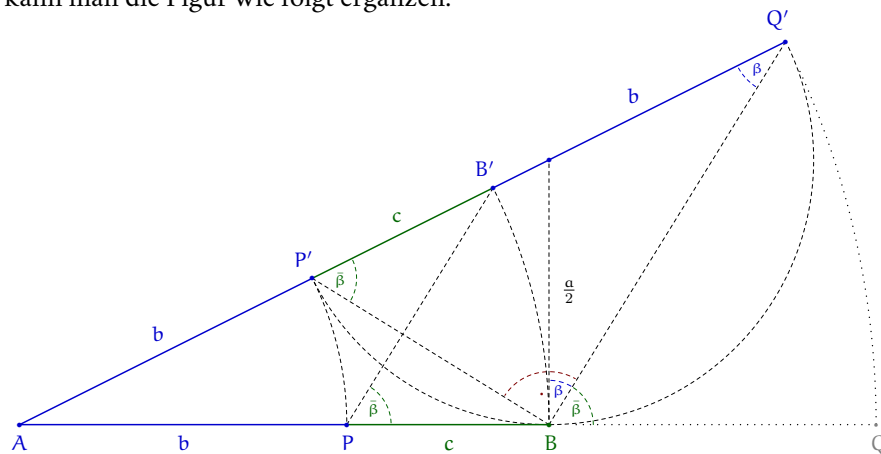
$$\varphi = \frac{b}{a} = \frac{a-b}{b} = \frac{a}{b} - 1 = \frac{1}{\varphi} - 1, \quad \text{also} \quad \varphi^2 + \varphi - 1 = 0$$

erfüllen. Die einzige positive Lösung ist $\varphi = \frac{1}{2}(\sqrt{5} - 1) \approx 0.618$. Das ist also das Verhältnis der längeren zur ganzen Strecke. Das müßte mit dem Verhältnis $\frac{a-b}{b}$ der kürzeren Strecke zur längeren übereinstimmen: $\frac{1}{\varphi} - 1 = \frac{2}{\sqrt{5}-1} - 1 = \frac{2(\sqrt{5}+1)}{5-1} - 1 = \frac{1}{2}(\sqrt{5} + 1) - 1 = \varphi$. Auch $\Phi := \frac{1}{2}(\sqrt{5} + 1) = \frac{1}{\varphi} = 1 + \varphi \approx 1.618$ wird als (großer) goldener Schnitt bezeichnet.

Die Skizze zeigt eine der Konstruktionen des goldenen Schnitts für $a = 1$, die darauf beruht, eine Strecke der Länge $\frac{1}{2}(\sqrt{5} - 1)$ zu konstruieren:



Für einen Beweis dieser Konstruktion, der nur die klassischen Hilfsmittel *Zirkel und Lineal* verwendet, kann man die Figur wie folgt ergänzen:

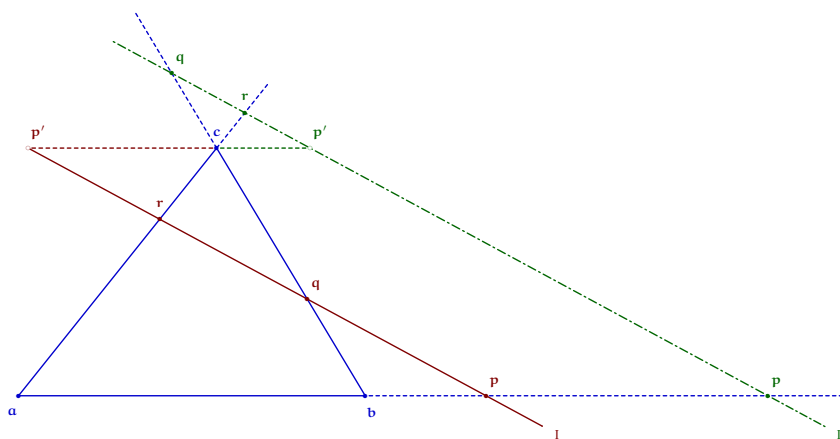


Dabei habe die Strecke AB die Länge $a = b + c$. Laut Konstruktion gilt das dann auch für $P'Q'$. Offensichtlich hat $P'B'$ die Länge c , so daß $B'Q'$ die Länge b haben muß. Aus der Symmetrie der Anordnung ist $PB' \parallel BQ'$ ersichtlich. Damit lassen sich die Strahlensätze anwenden: $\frac{c}{b} = \frac{b}{b+c}$, also $\frac{a-b}{b} = \frac{b}{a}$.

6.4.12 Satz (Menelaos) Die drei Schnittpunkte p, q, r einer Geraden mit den Seiten eines Dreiecks abc teilen diese mit Teilverhältnissen $T(a, p, b)$, $T(b, q, c)$ und $T(c, r, a)$, für die

$$T(a, p, b)T(b, q, c)T(c, r, a) = -1 \quad (6.9)$$

gilt. Umgekehrt folgt für drei Punkt p, q und r auf den Seiten des Dreiecks aus dieser Beziehung, daß sie auf einer Geraden liegen.



Beweis. Aus der Skizze ergibt sich mit Hilfe der Strahlensätze für beide Situationen I und II

$$\begin{aligned} T(\mathbf{a}, \mathbf{p}, \mathbf{b})T(\mathbf{b}, \mathbf{q}, \mathbf{c})T(\mathbf{c}, \mathbf{r}, \mathbf{a}) &= -\frac{\|\mathbf{p} - \mathbf{a}\|}{\|\mathbf{b} - \mathbf{p}\|} \frac{\|\mathbf{q} - \mathbf{b}\|}{\|\mathbf{c} - \mathbf{q}\|} \frac{\|\mathbf{r} - \mathbf{c}\|}{\|\mathbf{a} - \mathbf{r}\|} \\ &= -\frac{\|\mathbf{p} - \mathbf{a}\|}{\|\mathbf{b} - \mathbf{p}\|} \frac{\|\mathbf{b} - \mathbf{p}\|}{\|\mathbf{c} - \mathbf{p}'\|} \frac{\|\mathbf{c} - \mathbf{p}'\|}{\|\mathbf{p} - \mathbf{a}\|} = -1. \end{aligned}$$

In der Situation I ist nur $T(\mathbf{a}, \mathbf{p}, \mathbf{b})$ negativ, in der Situation II sind alle drei Teilverhältnisse negativ.

Für die Umkehrung dürfen wir o. B. d. A. von $\mathbf{a} = \mathbf{0}$ ausgehen. Mit $\lambda_p = T(\mathbf{a}, \mathbf{p}, \mathbf{b})$, $\lambda_q = T(\mathbf{b}, \mathbf{q}, \mathbf{c})$ und $\lambda_r = T(\mathbf{c}, \mathbf{r}, \mathbf{a})$ haben wir

$$\mathbf{p} = \frac{\lambda_p}{1 + \lambda_p} \mathbf{b}, \quad \mathbf{r} = \frac{1}{1 + \lambda_r} \mathbf{c} \quad \text{und} \quad \mathbf{q} = \frac{1}{1 + \lambda_q} \mathbf{b} + \frac{\lambda_q}{1 + \lambda_q} \mathbf{c}.$$

Wenn wir $\mathbf{q} = (1 - \beta)\mathbf{p} + \beta\mathbf{r} = \mathbf{p} + \beta(\mathbf{r} - \mathbf{p})$ für ein geeignetes β zeigen können, ist nachgewiesen, daß \mathbf{p} , \mathbf{q} und \mathbf{r} auf einer Geraden liegen. Dafür wählen wir $\beta := \frac{\lambda_q}{1 + \lambda_q}(1 + \lambda_r)$. Dann gilt

$$1 - \beta = \frac{1 - \lambda_q \lambda_r}{1 + \lambda_q} = \frac{\lambda_p - \lambda_p \lambda_q \lambda_r}{\lambda_p(1 + \lambda_q)} = \frac{1 + \lambda_p}{\lambda_p(1 + \lambda_q)},$$

woraus $\mathbf{q} = (1 - \beta)\mathbf{p} + \beta\mathbf{r}$ jetzt leicht folgt. \square

6.4.13 A Zeigen Sie, daß der Inkreismittelpunkt \mathbf{m} eines Dreiecks mit den Ecken \mathbf{a} , \mathbf{b} und \mathbf{c} durch

$$\mathbf{m} = \frac{\|\mathbf{b} - \mathbf{c}\|}{u} \cdot \mathbf{a} + \frac{\|\mathbf{c} - \mathbf{a}\|}{u} \cdot \mathbf{b} + \frac{\|\mathbf{a} - \mathbf{b}\|}{u} \cdot \mathbf{c} \quad (6.10)$$

gegeben ist. Dabei bezeichnet $u := \|\mathbf{a} - \mathbf{b}\| + \|\mathbf{b} - \mathbf{c}\| + \|\mathbf{c} - \mathbf{a}\|$ den Umfang des Dreiecks.

Schneiden Sie dafür die Winkelhalbierenden durch \mathbf{a} und \mathbf{c} . Als Richtungsvektoren der Geraden eignen sich $\frac{\mathbf{b} - \mathbf{a}}{\|\mathbf{b} - \mathbf{a}\|} + \frac{\mathbf{c} - \mathbf{a}}{\|\mathbf{c} - \mathbf{a}\|}$ bzw. $\frac{\mathbf{b} - \mathbf{c}}{\|\mathbf{b} - \mathbf{c}\|} - \frac{\mathbf{c} - \mathbf{a}}{\|\mathbf{c} - \mathbf{a}\|}$.

Zur Lösung der Gleichung verwenden Sie, daß die Menge $\{\mathbf{b} - \mathbf{a}, \mathbf{b} - \mathbf{c}\}$ linear unabhängig ist.

Berechnen Sie \mathbf{m} für $\mathbf{a} := [0, 0]^t$, $\mathbf{b} := [14, 0]^t$ und $\mathbf{c} := [9, 12]^t$. Bestimmen Sie auch den Radius des Inkreises und den Flächeninhalt des Dreiecks (Vergl. auch Aufgabe 5.2.7).

6.5 Basis und Dimension

6.5.1 Definition Für einen Vektorraum V über \mathbb{K} sei eine Menge $\mathcal{B} \subseteq V$ gegeben. Die lineare Hülle $\text{lh}(\mathcal{B})$ von \mathcal{B} ist die Menge aller Linearkombinationen, die sich mit Elementen aus \mathcal{B} bilden lassen. \mathcal{B} heißt erzeugend, falls $\text{lh}(\mathcal{B}) = V$ gilt. \mathcal{B} heißt Basis von V , falls \mathcal{B} erzeugend und linear unabhängig ist. V heißt endlich erzeugt, falls \mathcal{B} nur endlich viele Elemente enthält.

Von einer Basis \mathcal{B} wird man erwarten, daß sich jedes Element aus V durch die Basisvektoren aufbauen läßt. Das ist gewährleistet, weil \mathcal{B} erzeugend ist. Die lineare Unabhängigkeit sorgt dafür, daß \mathcal{B} nicht unnötig viele Elemente enthält (vergl. Lemma 6.4.7) und daß die Darstellung durch die Basisvektoren eindeutig ist.

6.5.2 Lemma Sei \mathcal{B} eine Basis für den Vektorraum V . Dann läßt sich jeder Vektor aus V eindeutig als Linearkombination von Elementen aus \mathcal{B} darstellen.

Beweis. Sei $\mathbf{x} \in V$. Da \mathcal{B} erzeugend ist, sind alle Elemente von V Linearkombinationen von Vektoren aus \mathcal{B} : $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{b}_k$, $\mathbf{b}_k \in \mathcal{B}$, $k = 1, \dots, n$. Ließe sich \mathbf{x} auch noch auf eine andere Weise durch Vektoren von \mathcal{B} kombinieren, etwa $\mathbf{x} = \sum_{k=1}^m y_k \tilde{\mathbf{b}}_k$, $\tilde{\mathbf{b}}_k \in \mathcal{B}$, $k = 1, \dots, m$, dann wäre durch $\mathbf{0} = \sum_{k=1}^n x_k \mathbf{b}_k - \sum_{k=1}^m y_k \tilde{\mathbf{b}}_k$ eine nicht triviale Nullkombination von Vektoren aus \mathcal{B} gegeben. Da \mathcal{B} linear unabhängig ist, erhalten wir einen Widerspruch. \square

6.5.3 Satz (Basis-Austausch-Satz) V sei durch eine Basis $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ endlich erzeugt. In der Entwicklung eines Vektors $\mathbf{x} \in V$ sei der Basisvektor \mathbf{b}_j vorhanden. Ersetzt man \mathbf{b}_j in \mathcal{B} durch \mathbf{x} , so ist die entstandene Menge $\mathcal{B}' := \{\mathbf{b}_k \mid k = 1, \dots, n, k \neq j\} \cup \{\mathbf{x}\}$ ebenfalls eine Basis für V .

Beweis. Laut Voraussetzung gilt $\mathbf{x} = \sum_{k \neq j}^n x_k \mathbf{b}_k + x_j \mathbf{b}_j$, $x_j \neq 0$. Wir können o. B. d. A. von $x_j = 1$ ausgehen, indem wir \mathbf{x} durch \mathbf{x}/x_j ersetzen. Zunächst zeigen wir, daß \mathcal{B}' linear unabhängig ist. Aus

$$\mathbf{0} = \sum_{k \neq j}^n \lambda_k \mathbf{b}_k + \lambda_j \mathbf{x} = \sum_{k \neq j}^n \lambda_k \mathbf{b}_k + \lambda_j \sum_{k \neq j}^n x_k \mathbf{b}_k + \lambda_j \mathbf{b}_j = \sum_{k \neq j}^n (\lambda_k + \lambda_j x_k) \mathbf{b}_k + \lambda_j \mathbf{b}_j$$

folgt, wegen der linearen Unabhängigkeit von \mathcal{B} , zuerst $\lambda_j = 0$ und dann $0 = \lambda_k + \lambda_j x_k = \lambda_k$ für $k \neq j$. Daher ist jede Nullkombination aus \mathcal{B}' trivial und \mathcal{B}' linear unabhängig.

\mathcal{B}' ist erzeugend: Jedes $\mathbf{y} \in V$ hat eine Darstellung $\mathbf{y} = \sum_{k \neq j}^n y_k \mathbf{b}_k + y_j \mathbf{b}_j$. Mit $\mathbf{b}_j = \mathbf{x} - \sum_{k \neq j}^n x_k \mathbf{b}_k$ erhalten wir daraus $\mathbf{y} = \sum_{k \neq j}^n (y_k - y_j x_k) \mathbf{b}_k + y_j \mathbf{x} \in \text{lh}(\mathcal{B}')$. Das zeigt $V = \text{lh}(\mathcal{B}')$. \square

6.5.4 Korollar Für eine Basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ von V kann eine linear unabhängige Menge $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ nicht mehr Elemente als \mathcal{B} enthalten. Hat sie gleich viele Elemente, so ist sie ebenfalls eine Basis. Insbesondere hat jede Basis von V dieselbe Anzahl von Basisvektoren.

Beweis. Wir nehmen $m > n$ an. Der Vektor \mathbf{c}_1 läßt sich in der Basis \mathcal{B} darstellen. O. B. d. A. sei der Basisvektor \mathbf{b}_1 in dieser Entwicklung vertreten (andernfalls nummerieren wir \mathcal{B} geeignet um). Nach Satz 6.5.3 ist $\mathcal{B}_1 := \{\mathbf{c}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ eine Basis von V . In dieser Basis läßt sich \mathbf{c}_2 darstellen:

$$\mathbf{c}_2 = \alpha_1 \mathbf{c}_1 + \sum_{k=2}^n \alpha_k \mathbf{b}_k.$$

Da \mathcal{C} linear unabhängig ist, muß für $k \geq 2$ wenigstens ein α_k von Null verschieden sein. Wir können o. B. d. A. $\alpha_2 \neq 0$ annehmen. Damit ist $\mathcal{B}_2 := \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{b}_3, \dots, \mathbf{b}_n\}$ eine Basis für V . Auf diese Weise können wir fortfahren, bis auch der Basisvektor \mathbf{b}_n durch \mathbf{c}_n ersetzt worden ist. Dann ist $\mathcal{B}_n := \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ eine Basis von V , in der die überzähligen Vektoren $\mathbf{c}_{n+1}, \dots, \mathbf{c}_m$ darstellbar sein müßten. Das ist ein Widerspruch zur linearen Unabhängigkeit von \mathcal{C} (vergl. Lemma 6.4.7). Daher muß $n = m$ gelten. Dann ist \mathcal{C} selbst eine Basis von V . Jetzt seien \mathcal{B} und \mathcal{C} zwei Basen. Dann kann die linear unabhängige Menge \mathcal{C} nach den bisherigen Ergebnissen nicht mehr Elemente als \mathcal{B} enthalten – weniger aber auch nicht, da die Argumentation mit vertauschten Rollen von \mathcal{B} und \mathcal{C} wiederholt werden kann. \square

6.5.5 Satz Jeder endlich erzeugte Vektorraum hat eine endliche Basis.

Beweis. Der Vektorraum V (o. B. d. A. $\neq \{0\}$) wird durch die Vektoren einer Menge $\mathcal{C} := \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ linear erzeugt. Ist \mathcal{C} linear unabhängig, so handelt es sich bereits um eine Basis. Andernfalls gibt es nach Lemma 6.4.7 einen Vektor aus \mathcal{C} , sagen wir \mathbf{c}_m , der durch die anderen linear kombiniert werden kann. Dann sollte auch die Menge $\mathcal{C}_1 := \{\mathbf{c}_1, \dots, \mathbf{c}_{m-1}\}$ noch erzeugend sein. Denn jeder Vektor $\mathbf{x} \in V$ läßt sich als Linearkombination der Vektoren $\mathbf{c}_1, \dots, \mathbf{c}_{m-1}$ und \mathbf{c}_m gewinnen. Ersetzen wir darin \mathbf{c}_m durch seine Linearkombination aus den Vektoren $\mathbf{c}_1, \dots, \mathbf{c}_{m-1}$, so sind im Ergebnis auch nur diese $m - 1$ Vektoren aus \mathcal{C}_1 an der Linearkombination von \mathbf{x} beteiligt.

Nun, da \mathcal{C}_1 ebenfalls erzeugend ist, können wir die Argumentation wiederholen: Entweder ist \mathcal{C}_1 bereits linear unabhängig und damit eine Basis, oder wir gehen (o. B. d. A.) zu der erzeugenden Menge $\mathcal{C}_2 := \{\mathbf{c}_1, \dots, \mathbf{c}_{m-2}\}$ über. Dieser Vorgang muß spätestens nach $m - 1$ Schritten stoppen und eine linear unabhängige Menge liefern. Laut Konstruktion ist sie erzeugend und daher eine Basis von V . \square

6.5.6 Definition Ein endlich erzeugter Vektorraum V heißt endlichdimensional. Die Dimension $\dim V$ von V ist die Anzahl der Basisvektoren einer Basis. V bezeichnen wir als unendlichdimensional, falls er keine Basis mit endlich vielen Elementen hat.

$\dim V$ ist die größte Anzahl an Elementen, die eine linear unabhängige Teilmenge von V haben kann.

6.5.7 Korollar Hat ein Vektorraum linear unabhängige Mengen mit beliebig, oder gar unendlich vielen Elementen, so ist er unendlich dimensional.

Beweis. Gäbe es eine endliche Basis, sagen wir mit n Elementen, so dürfte keine linear unabhängige Menge mehr als n Elemente enthalten, im Widerspruch zur Annahme. \square

Wenn nichts anderes gesagt wird, gehen wir im Folgenden immer von endlichdimensionalen Vektorräumen aus.

6.5.8 Satz (Basisergänzungssatz) *In einem endlichdimensionalen Vektorraum lässt sich jede linear unabhängige Menge zu einer Basis ergänzen.*

Beweis. $\mathcal{C} := \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ sei eine linear unabhängige Menge in einem Vektorraum V der Dimension n . Wir können von $m < n$ ausgehen, da andernfalls \mathcal{C} bereits eine Basis wäre (Korollar 6.5.4). Dann ist $V_0 := \text{lh } \mathcal{C}$ ein echter Teilraum von V (weil andernfalls \mathcal{C} eine Basis von V mit weniger Elementen als $\dim V$ wäre). Es gibt daher ein $\mathbf{x}_1 \notin V_0$. Dann ist $\mathcal{C}_1 := \mathcal{C} \cup \{\mathbf{x}_1\}$ ebenfalls linear unabhängig. Andernfalls gäbe es nämlich eine nicht triviale Nullkombination $\mathbf{0} = \sum_{k=1}^m \lambda_k \mathbf{c}_k + \lambda_0 \mathbf{x}_1$. Dabei müsste $\lambda_0 \neq 0$ (o. B. d. A. = 1) sein, da \mathcal{C} ja linear unabhängig ist. Wir könnten dann nach \mathbf{x}_1 auflösen und erhielten \mathbf{x}_1 als Linearkombination der Vektoren aus \mathcal{C} , was zum Widerspruch $\mathbf{x}_1 \in V_0$ führen würde. \mathcal{C}_1 ist also tatsächlich linear unabhängig. Enthält diese Menge bereits n Elemente, so handelt es sich um eine Basis. Andernfalls können wir das Verfahren wiederholen und erhalten linear unabhängige Mengen $\mathcal{C}_2 := \mathcal{C}_1 \cup \{\mathbf{x}_2\} = \mathcal{C} \cup \{\mathbf{x}_1, \mathbf{x}_2\}$, usw. Nach $n - m$ Schritten haben wir eine linear unabhängige Menge $\mathcal{C}_{n-m} := \mathcal{C} \cup \{\mathbf{x}_1, \dots, \mathbf{x}_{n-m}\}$ mit n Elementen, also eine Basis, die die linear unabhängigen Vektoren aus \mathcal{C} enthält. \square

Bemerkung: Die linear unabhängige Menge, von der in dem Satz die Rede ist, kann durchaus auch nur ein Element enthalten.

6.5.9 Beispiel Die Vektorräume \mathbb{R}^n und \mathbb{C}^n haben, wie erwartet, die Dimension n . Wir müssen dafür nur eine Basis mit n Elementen finden. Ohne groß darüber nachzudenken verwenden wir eine solche Basis bereits, wenn wir einen Vektor $\mathbf{x} \in \mathbb{R}^n$ oder $\mathbf{x} \in \mathbb{C}^n$ durch seinen Koordinatenvektor $[x_1, x_2, x_3, \dots, x_n]^t$ angeben. Es gilt nämlich

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + x_n \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \sum_{i=1}^n x_i \mathbf{e}_i.$$

Dabei ist \mathbf{e}_i , $i = 1, 2, \dots, n$, der Spaltenvektor, der an der Stelle i den Eintrag 1 und an allen anderen Stellen den Eintrag 0 hat.

$$\mathcal{E} := \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\} = \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_{n-1}, \mathbf{e}_n \}. \quad (6.11)$$

Diese Vektoren sind die Richtungsvektoren der Koordinatenachsen. Die Zahl x_i mißt den Anteil, den die Richtung e_i am Vektor x hat. Offensichtlich sind diese Vektoren erzeugend und linear unabhängig, also eine Basis. Wir bezeichnen sie mit \mathcal{E} und nennen sie die *kanonische Basis* von \mathbb{R}^n bzw. \mathbb{C}^n :

6.5.10 Beispiel Für \mathbb{C}^n definiert jede Menge aus n linear unabhängigen Vektoren eine Basis. Beispielsweise ist

$$\mathcal{B} := \left\{ \begin{bmatrix} 1 \\ 4 \\ i \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} i \\ 1 \\ 3 \end{bmatrix} \right\},$$

eine Basis für den \mathbb{C}^3 . Überzeugen wir uns davon, indem wir das homogene Gleichungssystem lösen, das zu einer Nullkombination der drei Vektoren aus \mathcal{B} gehört:

I	1	1	i	
II	4	2	1	II - 4 I
III	i	0	3	III - i I
I	1	1	i	
II	0	-2	1 - 4i	
III	0	-i	4	2 III - i II
I	1	1	i	
II	0	-2	1 - 4i	
III	0	0	4 - i	

Jeder Vektor x aus \mathbb{C}^3 muß sich jetzt auf eindeutige Weise als Linearkombination der Basisvektoren b_1, b_2, b_3 von \mathcal{B} schreiben lassen: $x = \sum_{k=1}^3 x_k b_k$. Man nennt das *den Vektor x nach der Basis \mathcal{B} entwickeln*. Um das konkret durchzuführen, müssen die *Entwicklungskoeffizienten* x_k gefunden werden. Dabei handelt es sich um ein lineares Gleichungssystem, das mit dem GAUSS-Verfahren gelöst werden kann. Als Beispiel wählen wir $x := [3, 1, i]^t$:

$$x_1 \begin{bmatrix} 1 \\ 4 \\ i \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} i \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ i \end{bmatrix} \Leftrightarrow \begin{array}{rrcr} x_1 & + & x_2 & + & ix_3 & = & 3 \\ 4x_1 & + & 2x_2 & & x_3 & = & 1 \\ ix_1 & & & + & 3x_3 & = & i \end{array}$$

	x_1	x_2	x_3		
I	1	1	i	3	
II	4	2	1	1	II - 4 I
III	i	0	3	i	III - i I
I	1	1	i	3	2 I + II
II	0	-2	1 - 4i	-11	
III	0	-i	4	-2i	2 III - i II
I	2	0	1 - 2i	-5	(4 - i)I - (1 - 2i)III
II	0	-2	1 - 4i	-11	(4 - i)II - (1 - 4i)III
III	0	0	4 - i	7i	

	x_1	x_2	x_3	
I	$2(4-i)$	0	0	$-34-2i$
II	0	$-2(4-i)$	0	$-72+4i$
III	0	0	$4-i$	$7i$
I	17	0	0	$-67-21i$
II	0	17	0	$146+28i$
III	0	0	17	$-7+28i$

Die Lösung ist $x_1 = -\frac{1}{17}(67+21i)$, $x_2 = \frac{1}{17}(146+28i)$ und $x_3 = \frac{1}{17}(-7+28i)$. Die Entwicklung von \mathbf{x} nach der Basis \mathcal{B} haben wir damit berechnet:

$$\begin{bmatrix} 3 \\ 1 \\ i \end{bmatrix} = -\frac{1}{17}(67+21i) \begin{bmatrix} 1 \\ 4 \\ i \end{bmatrix} + \frac{1}{17}(146+28i) \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \frac{1}{17}(-7+28i) \begin{bmatrix} i \\ 1 \\ 3 \end{bmatrix}.$$

6.5.11 Komplexes LGS mit reellem GAUSS-Verfahren Das letzte Beispiel zeigt, daß sich am GAUSS-Verfahren für komplexe Zahlen nichts ändert – außer, daß es recht mühsam werden kann. Es gibt aber eine Methode, ein komplexes Gleichungssystem $C\mathbf{z} = \mathbf{w}$ in ein reelles zu übersetzen, allerdings zu dem Preis, daß aus einem komplexen $m \times n$ -System ein reelles $2m \times 2n$ -System wird. Dafür zerlegt man die Matrix C in ihren Real- und Imaginärteil A bzw. B (reelle $m \times n$ -Matrizen) und genauso den gesuchten Vektor $\mathbf{z} \in \mathbb{C}^n$ sowie die Inhomogenität $\mathbf{w} \in \mathbb{C}^m$:

$$C = A + iB, \quad \mathbf{z} = \mathbf{x} + i\mathbf{y}, \quad \mathbf{w} = \mathbf{u} + i\mathbf{v}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^m.$$

Dann gilt $C\mathbf{z} = (A + iB)(\mathbf{x} + i\mathbf{y}) = A\mathbf{x} - B\mathbf{y} + i(B\mathbf{x} + A\mathbf{y}) = \mathbf{u} + i\mathbf{v}$, d. h., wir erhalten ein gekoppeltes, reelles Gleichungssystem: $A\mathbf{x} - B\mathbf{y} = \mathbf{u}$ und $B\mathbf{x} + A\mathbf{y} = \mathbf{v}$. Das bauen wir zu einem reellen $2m \times 2n$ -System um

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}. \quad (6.12)$$

Dieses (viermal so große) reelle Gleichungssystem lösen wir mit dem reellen GAUSS-Verfahren.

6.5.12 Beispiel Ein komplexes 3×3 -Gleichungssystem:

$$C := \begin{bmatrix} 1+i & -i & 2 \\ i & 1+2i & 2+4i \\ 3 & 1+2i & 4+i \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 3 & 1 & 4 \end{bmatrix} + i \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \\ 0 & 2 & 1 \end{bmatrix} = A + iB,$$

$$\mathbf{w} := \begin{bmatrix} 16+3i \\ -8+34i \\ 15+18i \end{bmatrix} = \begin{bmatrix} 16 \\ -8 \\ 15 \end{bmatrix} + i \begin{bmatrix} 3 \\ 34 \\ 18 \end{bmatrix} = \mathbf{u} + i\mathbf{v}.$$

Das führt auf das reelle 6×7 -Schema für das GAUSS-Verfahren:

	x_1	x_2	x_3	y_1	y_2	y_3		
I	1	0	2	-1	1	0	16	
II	0	1	2	-1	-2	-4	-8	
III	3	1	4	0	-2	-1	15	III - 3 I
IV	1	-1	0	1	0	2	3	IV - I
V	1	2	4	0	1	2	34	V - I
VI	0	2	1	3	1	4	18	
I	1	0	2	-1	1	0	16	
II	0	1	2	-1	-2	-4	-8	
III	0	1	-2	3	-5	-1	-33	III - II
IV	0	-1	-2	2	-1	2	-13	IV + II
V	0	2	2	1	0	2	18	V - 2 II
VI	0	2	1	3	1	4	18	VI - 2 II
I	1	0	2	-1	1	0	16	2 I + III
II	0	1	2	-1	-2	-4	-8	2 II + III
III	0	0	-4	4	-3	3	-25	
IV	0	0	0	1	-3	-2	-21	
V	0	0	-2	3	4	10	34	2 V - III
VI	0	0	-3	5	5	12	34	4 VI - 3 III
I	2	0	0	2	-1	3	7	I - 2 X
II	0	2	0	2	-7	-5	-41	II - 2 IV
III	0	0	-4	4	-3	3	-25	III - 4 IV
IV	0	0	0	1	-3	-2	-21	
V	0	0	0	2	11	17	93	V - 2 IV
VI	0	0	0	8	29	39	211	VI - 8 IV
I	2	0	0	0	5	7	49	(17 I - 5 V)/2
II	0	2	0	0	-1	-1	1	(17 II + V)/2
III	0	0	-4	0	9	11	59	(17 III - 9 V)/2
IV	0	0	0	1	-3	-2	-21	17 IV + 3 V
V	0	0	0	0	17	21	135	
VI	0	0	0	0	53	55	379	-(17 VI - 53 V)/178
I	17	0	0	0	0	7	79	(I - 7 VI)/17
II	0	17	0	0	0	2	76	(II - 2 VI)/17
III	0	0	-34	0	0	-1	-106	-(III + VI)/34
IV	0	0	0	17	0	29	48	(IV - 29 VI)/17
V	0	0	0	0	17	21	135	(V - 21 VI)/17
VI	0	0	0	0	0	1	4	
I	1	0	0	0	0	0	3	
II	0	1	0	0	0	0	4	
III	0	0	1	0	0	0	3	
IV	0	0	0	1	0	0	-4	
V	0	0	0	0	1	0	3	
VI	0	0	0	0	0	1	4	

Daher ist $z = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix} + i \begin{bmatrix} -4 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 - 4i \\ 4 + 3i \\ 3 + 4i \end{bmatrix}.$

Kontrolle:

$$\begin{aligned}
 Cz &= \begin{bmatrix} 1+i & -i & 2 \\ i & 1+2i & 2+4i \\ 3 & 1+2i & 4+i \end{bmatrix} \begin{bmatrix} 3-4i \\ 4+3i \\ 3+4i \end{bmatrix} \\
 &= \begin{bmatrix} (1+i)(3-4i) - i(4+3i) + 2(3+4i) \\ i(3-4i) + (1+2i)(4+3i) + (2+4i)(3+4i) \\ 3(3-4i) + (1+2i)(4+3i) + (4+i)(3+4i) \end{bmatrix}
 \end{aligned}$$

$$= \begin{bmatrix} 7-i+3-4i+6+8i \\ 4+3i-2+11i-10+20i \\ 9-12i-2+11i+8+19i \end{bmatrix} = \begin{bmatrix} 16+3i \\ -8+34i \\ 15+18i \end{bmatrix} = \mathbf{w}.$$

Ändern wir das Beispiel so ab, daß C einen nichttrivialen Kern hat:

$$C := \begin{bmatrix} 1+i & -i & 2 \\ i & 1+2i & 2+4i \end{bmatrix}, \quad \mathbf{w} := \begin{bmatrix} 16+3i \\ -8+34i \end{bmatrix}.$$

	x_1	x_2	x_3	y_1	y_2	y_3		
I	1	0	2	-1	1	0	16	
II	0	1	2	-1	-2	-4	-8	
III	1	-1	0	1	0	2	3	III - I
IV	1	2	4	0	1	2	34	IV - I
I	1	0	2	-1	1	0	16	
II	0	1	2	-1	-2	-4	-8	
III	0	-1	-2	2	-1	2	-13	III + II
IV	0	2	2	1	0	2	18	IV - 2 II
I	1	0	2	-1	1	0	16	I + IV
II	0	1	2	-1	-2	-4	-8	II + IV
III	0	0	0	1	-3	-2	-21	III \leftrightarrow IV
IV	0	0	-2	3	4	10	34	
I	1	0	0	2	5	10	50	I - 2 IV
II	0	1	0	2	2	6	26	II - 2 IV
III	0	0	-2	3	4	10	34	III - 3 IV
IV	0	0	0	1	-3	-2	-21	
I	1	0	0	0	11	14	92	
II	0	1	0	0	8	10	68	
III	0	0	-2	0	13	16	97	
IV	0	0	0	1	-3	-2	-21	

Das führt auf

$$\begin{array}{rcl} x_1 & = & 92 - 11 y_2 - 14 y_3 \\ x_2 & = & 68 - 8 y_2 - 10 y_3 \\ x_3 & = & -48.5 + 6.5 y_2 + 8 y_3 \\ y_1 & = & -21 + 3 y_2 + 2 y_3 \\ y_2 & = & 0 + 1 y_2 + 0 y_3 \\ y_3 & = & 0 + 0 y_2 + 1 y_3 \end{array} \quad \text{also} \quad \mathbf{z} = \begin{bmatrix} 92 - 21i \\ 68 \\ -48.5 \end{bmatrix} + y_2 \begin{bmatrix} 3i - 11 \\ i - 8 \\ 6.5 \end{bmatrix} + y_3 \begin{bmatrix} 2i - 14 \\ -10 \\ i + 8 \end{bmatrix},$$

$y_2, y_3 \in \mathbb{R}.$

Das ist zwar ein Ergebnis, aber die allgemeine Lösungsform $\mathbf{z} = \mathbf{z}_1 + t \cdot \mathbf{z}_0$, $t \in \mathbb{C}$, läßt sich nicht ganz leicht erkennen. Die spezielle inhomogene Lösung $\mathbf{z}_1 = [92 - 21i, 68, -18.5]^t$ ist natürlich gut sichtbar, aber die homogene Lösung \mathbf{z}_0 nicht, wenn man nicht schon weiß, daß bei diesem einfachen Beispiel der Kern von C eindimensional ist. Daher müssen $[3i - 11, i - 8, 6.5]^t$ und $[2i - 14, -10, i + 8]^t$ linear abhängig sein (der Streckungsfaktor ist $\frac{2}{13}(8 + i)$), so daß man z. B. den ersten Vektor als \mathbf{z}_0 wählen kann. Bei größeren Systemen mit größeren Kernen wird die Analyse der homogenen Lösungen allerdings weiteren Rechenaufwand erfordern. Es wäre daher wünschenswert, das Verfahren so abzuändern, daß die komplexen Lösungen ohne weitere Nachbearbeitung herauskommen. Dafür sollte man das GAUSS-Verfahren derart steuern, daß die Ausgangsform in ihrer Struktur erhalten wird, daß das Ergebnis also wieder die Gestalt von (6.12) aufweist:

	x_1	x_2	x_3	y_1	y_2	y_3		
I	1	0	2	-1	1	0	16	I + III
II	0	1	2	-1	-2	-4	-8	II + III
III	0	0	0	1	-3	-2	-21	
IV	0	0	-2	3	4	10	34	IV - 3 III

	x_1	x_2	x_3	y_1	y_2	y_3		
I	1	0	2	0	-2	-2	-5	13 I + 2 IV
II	0	1	2	0	-5	-6	-29	13 II + 5 IV
III	0	0	0	1	-3	-2	-21	13 III + 3 IV
IV	0	0	-2	0	13	16	97	
I	13	0	22	0	0	6	129	
II	0	13	16	0	0	2	108	
III	0	0	-6	13	0	22	18	
IV	0	0	-2	0	13	16	97	

Das lässt sich jetzt zwanglos in eine komplexe Lösung übersetzen:

$$\begin{aligned}
 z &= \frac{1}{13} \begin{bmatrix} 129 + 18i \\ 108 + 97i \\ 0 \end{bmatrix} + \frac{x_3}{13} \begin{bmatrix} -22 + 6i \\ -16 + 2i \\ 13 \end{bmatrix} + \frac{y_3}{13} \begin{bmatrix} -6 - 22i \\ -2 - 16i \\ 13i \end{bmatrix} \\
 &= \frac{1}{13} \begin{bmatrix} 129 + 18i \\ 108 + 97i \\ 0 \end{bmatrix} + \frac{x_3 + iy_3}{13} \begin{bmatrix} -22 + 6i \\ -16 + 2i \\ 13 \end{bmatrix}.
 \end{aligned}$$

6.5.13 Komplexe Inverse mit reellem GAUSS-Verfahren Die Berechnung einer inversen Matrix lässt sich natürlich auch mit dem reellen GAUSS-Verfahren durchführen (das ist ein Vorgriff auf das Thema *Inverse Matrix* (siehe 6.9), das an dieser Stelle der Vollständigkeit halber mitbehandelt wird): $C = A + iB$ habe die Inverse $C^{-1} = D + iE$. Das ist äquivalent zu $AD - BE + i(AE + BD) = \mathbb{1}$, oder $AD - BE = \mathbb{1}$ und $AE + BD = 0$, also zu

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} D \\ E \end{bmatrix} = \begin{bmatrix} \mathbb{1} \\ 0 \end{bmatrix}.$$

Diese Matrixgleichung lösen wir mit dem erweiterten GAUSS-Verfahren. Nehmen wir zur Demonstration das einfache Beispiel $C := \begin{bmatrix} 2-3i & 4+i \\ 4-i & i \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 0 \end{bmatrix} + i \begin{bmatrix} -3 & 1 \\ -1 & 1 \end{bmatrix}$:

I	2	4	3	-1	1	0	
II	4	0	1	-1	0	1	II - 2 I
III	-3	1	2	4	0	0	2 III + 3 I
IV	-1	1	4	0	0	0	2 IV + I
I	2	4	3	-1	1	0	2 I + II
II	0	-8	-5	1	-2	1	
III	0	14	13	5	3	0	4 III + 7 II
IV	0	6	11	-1	1	0	4 IV + 3 II
I	4	0	1	-1	0	1	(17 I - III)/2
II	0	-8	-5	1	-2	1	(17 II + 5 III)/4
III	0	0	17	27	-2	7	
IV	0	0	29	-1	-2	3	(17 IV - 29 III)/8
I	34	0	0	-22	1	5	(100 I - 22 IV)/34
II	0	-34	0	38	-11	13	-(100 II + 38 IV)/34
III	0	0	17	27	-2	7	(100 III + 27 IV)/17
IV	0	0	0	-100	3	-19	
I	100	0	0	0	1	27	
II	0	100	0	0	29	-17	
III	0	0	100	0	-7	11	
IV	0	0	0	100	-3	19	

Die zugehörige Matrixgleichung hat jetzt die Form

$$\begin{bmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{bmatrix} \begin{bmatrix} \mathbf{D} \\ \mathbf{E} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \\ \mathbf{E} \end{bmatrix} = \frac{1}{100} \begin{bmatrix} 1 & 27 \\ 29 & -17 \\ -7 & 11 \\ -3 & 19 \end{bmatrix}.$$

$$\text{Also ist } C^{-1} = \frac{1}{100} \begin{bmatrix} 1 & 27 \\ 29 & -17 \end{bmatrix} + \frac{i}{100} \begin{bmatrix} -7 & 11 \\ -3 & 19 \end{bmatrix} = \frac{1}{100} \begin{bmatrix} 1 - 7i & 27 + 11i \\ 29 - 3i & -17 + 19i \end{bmatrix}.$$

6.5.14 Orthonormalbasis Beispiel 6.5.10 zeigt, daß es mühsam sein kann, die Entwicklungskoeffizienten eines Vektors $\mathbf{x} \in V$ für eine Basis $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ von V zu bestimmen. Normalerweise läuft das auf ein inhomogenes Gleichungssystem hinaus, das mit dem GAUSS-Verfahren gelöst werden muß. Es gibt aber besonders gute Basen, für die das Auffinden der Koeffizienten sehr einfach wird. Wenn uns ein Skalarprodukt auf V zur Verfügung steht, wenn V also ein Hilbertraum ist, lassen sich Basen \mathcal{B} mit folgenden Eigenschaften bilden.

6.5.15 Definition Eine Basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ für einen Hilbertraum V über \mathbb{K} heißt orthonormal, falls sie folgende Eigenschaften aufweist:

- i) Die Basisvektoren sind paarweise orthogonal zueinander: $\mathbf{b}_i \perp \mathbf{b}_j$, $i \neq j = 1, \dots, n$.
- ii) Alle Basisvektoren sind auf die Länge 1 normiert: $\|\mathbf{b}_i\| = 1$, $i = 1, \dots, n$.

\mathcal{B} heißt Orthonormalbasis (ONB).

Die Orthonormalität einer Basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ läßt sich mit dem KRONECKER-Symbol δ_{ik} , das durch

$$\delta_{ik} := \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \quad (6.13)$$

definiert ist, folgendermaßen wiedergeben: Für alle $i, k = 1, \dots, n$ gilt

$$\langle \mathbf{b}_i | \mathbf{b}_k \rangle = \delta_{ik}. \quad (6.14)$$

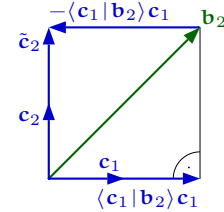
Die Entwicklung eines Vektors $\mathbf{x} = x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + \dots + x_n \mathbf{b}_n$ nach einer ONB $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ reduziert sich auf die Berechnung von Skalarprodukten:

$$\begin{aligned} \langle \mathbf{b}_1 | \mathbf{x} \rangle &= \langle \mathbf{b}_1 | x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + \dots + x_n \mathbf{b}_n \rangle \\ &= x_1 \langle \mathbf{b}_1 | \mathbf{b}_1 \rangle + x_2 \langle \mathbf{b}_1 | \mathbf{b}_2 \rangle + \dots + x_n \langle \mathbf{b}_1 | \mathbf{b}_n \rangle = \sum_{k=1}^n x_k \cdot \delta_{1k} = x_1. \end{aligned}$$

Genauso folgen $x_2 = \langle \mathbf{b}_2 | \mathbf{x} \rangle, \dots, x_n = \langle \mathbf{b}_n | \mathbf{x} \rangle$.

6.5.16 Das GRAM-SCHMIDT-Verfahren Das *GRAM-SCHMIDT-Verfahren*, erzeugt aus einer beliebigen Basis von $(V, \langle \cdot | \cdot \rangle)$ eine Orthonormalbasis.

Wir stellen die Idee an einer Basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ vor. Aus \mathcal{B} soll eine ONB $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ entstehen. Wir starten mit $\mathbf{c}_1 := \frac{1}{\|\mathbf{b}_1\|} \mathbf{b}_1$. Anschließend konstruieren wir aus \mathbf{c}_1 und \mathbf{b}_2 einen Vektor $\tilde{\mathbf{c}}_2$, der orthogonal zu \mathbf{c}_1 ist. Die Idee dazu lässt sich aus der Skizze leicht ablesen:



$$\tilde{\mathbf{c}}_2 := \mathbf{b}_2 - \langle \mathbf{c}_1 | \mathbf{b}_2 \rangle \mathbf{c}_1.$$

Dann gilt nämlich $\langle \mathbf{c}_1 | \tilde{\mathbf{c}}_2 \rangle = \langle \mathbf{c}_1 | \mathbf{b}_2 \rangle - \langle \mathbf{c}_1 | \mathbf{b}_2 \rangle \langle \mathbf{c}_1 | \mathbf{c}_1 \rangle = 0$. Den zweiten Basisvektor \mathbf{c}_2 erhalten wir durch Normierung von $\tilde{\mathbf{c}}_2$: $\mathbf{c}_2 := \frac{1}{\|\tilde{\mathbf{c}}_2\|} \tilde{\mathbf{c}}_2$. Damit sind die ersten beiden Basisvektoren bereits konstruiert. Sehen wir, ob wir das Verfahren richtig erweitern können:

$$\tilde{\mathbf{c}}_3 := \mathbf{b}_3 - \langle \mathbf{c}_1 | \mathbf{b}_3 \rangle \mathbf{c}_1 - \langle \mathbf{c}_2 | \mathbf{b}_3 \rangle \mathbf{c}_2.$$

Wir erhalten $\langle \mathbf{c}_1 | \tilde{\mathbf{c}}_3 \rangle = \langle \mathbf{c}_1 | \mathbf{b}_3 \rangle - \langle \mathbf{c}_1 | \mathbf{b}_3 \rangle \langle \mathbf{c}_1 | \mathbf{c}_1 \rangle - \langle \mathbf{c}_2 | \mathbf{b}_3 \rangle \langle \mathbf{c}_1 | \mathbf{c}_2 \rangle = 0$ und genauso $\langle \mathbf{c}_2 | \tilde{\mathbf{c}}_3 \rangle = \langle \mathbf{c}_2 | \mathbf{b}_3 \rangle - \langle \mathbf{c}_1 | \mathbf{b}_3 \rangle \langle \mathbf{c}_2 | \mathbf{c}_1 \rangle - \langle \mathbf{c}_2 | \mathbf{b}_3 \rangle \langle \mathbf{c}_2 | \mathbf{c}_2 \rangle = 0$. Damit ist $\mathbf{c}_1 \perp \tilde{\mathbf{c}}_3$ und $\mathbf{c}_2 \perp \tilde{\mathbf{c}}_3$ gewährleistet. Also ist $\mathbf{c}_3 := \frac{1}{\|\tilde{\mathbf{c}}_3\|} \tilde{\mathbf{c}}_3$ der letzte Vektor der ONB \mathcal{C} .

An diesem Beispiel ist klar zu erkennen, daß es sich um ein rekursives Verfahren handelt. Von $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ ausgehend werden nacheinander die Vektoren $\mathbf{c}_1, \mathbf{c}_2$ usw. berechnet. Ist \mathbf{c}_k für ein $k < n$ bereits konstruiert, dann wird zunächst

$$\tilde{\mathbf{c}}_{k+1} := \mathbf{b}_{k+1} - \sum_{i=1}^k \langle \mathbf{c}_i | \mathbf{b}_{k+1} \rangle \mathbf{c}_i \quad (6.15)$$

gebildet. Dieser Vektor ist zu allen $\mathbf{c}_1, \dots, \mathbf{c}_k$ orthogonal, denn für $1 \leq j \leq k$ gilt

$$\langle \mathbf{c}_j | \tilde{\mathbf{c}}_{k+1} \rangle := \langle \mathbf{c}_j | \mathbf{b}_{k+1} \rangle - \sum_{i=1}^k \langle \mathbf{c}_i | \mathbf{b}_{k+1} \rangle \langle \mathbf{c}_j | \mathbf{c}_i \rangle = \langle \mathbf{c}_j | \mathbf{b}_{k+1} \rangle - \sum_{i=1}^k \langle \mathbf{c}_i | \mathbf{b}_{k+1} \rangle \delta_{ji} = 0.$$

Anschließend wird dieser Vektor normiert: $\mathbf{c}_k := \frac{1}{\|\tilde{\mathbf{c}}_k\|} \tilde{\mathbf{c}}_k$. Auf diese Weise entsteht die ONB $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ aus $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$.

6.5.17 Beispiel Wir gehen von der Basis

$$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\} := \left\{ \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} i \\ i \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ i \end{bmatrix} \right\}$$

aus. Dann ist $\mathbf{c}_1 = \frac{1}{5} [3, 0, 4]^t$,

$$\tilde{\mathbf{c}}_2 = \begin{bmatrix} i \\ i \\ -1 \end{bmatrix} - \frac{1}{25} (3i - 4) \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 25i - 9i + 12 \\ 25i \\ -25 - 12i + 16 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 12 + 16i \\ 25i \\ -9 - 12i \end{bmatrix}$$

und $\mathbf{c}_2 = \frac{1}{25\sqrt{2}} [12 + 16i, 25i, -9 - 12i]^t$. Bevor wir $\tilde{\mathbf{c}}_3$ bestimmen, berechnen wir die beteiligten Skalarprodukte.

$$\langle \mathbf{c}_1 | \mathbf{b}_3 \rangle = \frac{1}{5}(3 + 4i), \quad \langle \mathbf{c}_2 | \mathbf{b}_3 \rangle = \frac{1}{25\sqrt{2}}(12 - 16i - 75i - 9i - 12) = -2\sqrt{2}i.$$

Damit erhalten wir

$$\begin{aligned} \tilde{\mathbf{c}}_3 &= \begin{bmatrix} 1 \\ 3 \\ i \end{bmatrix} - \frac{3+4i}{25} \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} + \frac{2i}{25} \begin{bmatrix} 12+16i \\ 25i \\ -9-12i \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 25-9-12i+24i-32 \\ 75-50 \\ 25i-12-16i-18i+24 \end{bmatrix} \\ &= \frac{1}{25} \begin{bmatrix} -16+12i \\ 25 \\ 12-9i \end{bmatrix}. \end{aligned}$$

Die ONB ist damit

$$\mathcal{C} = \{ \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3 \} = \left\{ \frac{1}{5} \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \frac{\sqrt{2}}{50} \begin{bmatrix} 12+16i \\ 25i \\ -9-12i \end{bmatrix}, \frac{\sqrt{2}}{50} \begin{bmatrix} -16+12i \\ 25 \\ 12-9i \end{bmatrix} \right\}.$$

Jetzt stellen wir noch den Vektor $\mathbf{x} := [1, 2, 3]^t$ in dieser Basis dar:

$$\langle \mathbf{c}_1 | \mathbf{x} \rangle = 3, \langle \mathbf{c}_2 | \mathbf{x} \rangle = -\frac{3\sqrt{2}}{10}(1+2i) \text{ und } \langle \mathbf{c}_3 | \mathbf{x} \rangle = \frac{\sqrt{2}}{10}(14+3i) \text{ ergibt}$$

$$\mathbf{x} = 3 \cdot \mathbf{c}_1 - \frac{3\sqrt{2}}{10}(1+2i) \cdot \mathbf{c}_2 + \frac{\sqrt{2}}{10}(14+3i) \cdot \mathbf{c}_3.$$

6.6 Matrizen als lineare Abbildungen

Im Zusammenhang mit linearen Gleichungssystemen sind Matrizen eine bequeme Methode das Wesentliche des Systems, nämlich die Koeffizienten a_{ij} , zusammenzufassen. Mit der maßgeschneiderten Multiplikation zwischen $m \times n$ -Matrix A und Spaltenvektor $\mathbf{x} \in \mathbb{K}^n$ ergibt sich die prägnante Form $A\mathbf{x} = \mathbf{b}$ für das Gleichungssystem. In diesem Zusammenhang hat A eine passive, sozusagen eine ordnende Aufgabe, die wir verwenden, um das GAUSS-Verfahren übersichtlich zu gestalten, oder um die Lösungsstruktur herauszuarbeiten.

Wir können der Matrix A in $A\mathbf{x} = \mathbf{b}$ aber auch eine aktive Rolle zuweisen, indem wir diese Gleichung als Abbildung lesen: Der Vektor $\mathbf{x} \in \mathbb{K}^n$ wird durch A auf den Vektor $\mathbf{b} \in \mathbb{K}^m$ abgebildet. Die Multiplikation zwischen A und \mathbf{x} bestimmt dabei die konkrete Abbildungsvorschrift. Der Einfachheit halber bezeichnen wir diese Abbildung wieder mit demselben Symbol A (wir werden aber Situationen kennenlernen, in denen wir zwischen solchen Abbildungen und ihren Realisierungen durch Matrizen unterscheiden müssen). A hat dabei eine entscheidende Eigenschaft, nämlich die *Linearität*

$$A(s\mathbf{x} + t\mathbf{y}) = sA\mathbf{x} + tA\mathbf{y} \quad \text{für alle } s, t \in \mathbb{K}, \mathbf{x}, \mathbf{y} \in \mathbb{K}^n,$$

die sie gegenüber einer beliebigen Abbildung zwischen \mathbb{K}^n und \mathbb{K}^m auszeichnet. Wir sprechen in diesem Zusammenhang von einer *linearen Abbildung*. Gelegentlich heben wir den Abbildungscharakter von A durch die Schreibweise

$$A : \mathbb{K}^n \rightarrow \mathbb{K}^m; \mathbf{x} \mapsto A\mathbf{x}$$

hervor. Der Linearität ist es zu verdanken, daß solche Abbildungen durch die Kenntnis weniger Daten bereits eindeutig festgelegt sind. Um zu verstehen, was genau damit gemeint ist, entwickeln wir \mathbf{x} in der kanonischen Basis \mathcal{E} : $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$. Wenden wir darauf die Abbildung A an, so folgt aus deren Linearität

$$A\mathbf{x} = A\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n x_i A\mathbf{e}_i.$$

Das zeigt, daß das Bild $A\mathbf{x}$ von \mathbf{x} als Linearkombination $\sum_{i=1}^n x_i A\mathbf{e}_i$ der Bilder $A\mathbf{e}_i$ der kanonischen Basisvektoren \mathbf{e}_i entsteht. Daher ist die Wirkung der linearen Abbildung A wirklich bereits durch wenige Daten festgelegt, nämlich durch die n Bilder $A\mathbf{e}_1, \dots, A\mathbf{e}_n$ der kanonischen Basisvektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$. Diese Bilder müßten im Zahlenschema der Matrix wiederzufinden sein:

$$A\mathbf{e}_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix} =: \mathbf{a}_1,$$

$$\begin{aligned}
A\mathbf{e}_2 &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{bmatrix} =: \mathbf{a}_2, \\
&\vdots \\
A\mathbf{e}_n &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{mn} \end{bmatrix} =: \mathbf{a}_n.
\end{aligned}$$

Und tatsächlich sind es einfach die Spaltenvektoren \mathbf{a}_i der Matrix. Um auf diesen Sachverhalt hinzuweisen, vereinbaren wir die folgende Schreibweise für eine Matrix:

$$A =: [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]. \quad (6.16)$$

D. h.:

In den Spalten einer Matrix stehen die Bilder der kanonischen Basisvektoren.

Daraus ist sofort ersichtlich, daß es sich bei $A\mathbf{x}$ einfach um eine Linearkombination der Spaltenvektoren \mathbf{a}_i handelt, mit den Koordinaten x_i als Linearfaktoren:

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \sum_{i=1}^n x_i\mathbf{a}_i. \quad (6.17)$$

Dieses Ergebnis läßt sich natürlich auch direkt aus der Definition (6.4) für die Multiplikation einer Matrix A mit einem Vektor \mathbf{x} gewinnen:

$$\begin{aligned}
A\mathbf{x} &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} x_1a_{11} \\ x_1a_{21} \\ x_1a_{31} \\ \vdots \\ x_1a_{m1} \end{bmatrix} + \begin{bmatrix} x_2a_{12} \\ x_2a_{22} \\ x_2a_{32} \\ \vdots \\ x_2a_{m2} \end{bmatrix} + \dots + \begin{bmatrix} x_na_{1n} \\ x_na_{2n} \\ x_na_{3n} \\ \vdots \\ x_na_{mn} \end{bmatrix} \\
&= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{mn} \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n.
\end{aligned}$$

Können wir also aus der Abbildungsvorschrift erkennen, was aus den Basisvektoren wird, so kennen wir die zugehörige Matrix. Als Beispiel dafür nehmen wir eine Drehung D_α im \mathbb{R}^2 um einen festen Winkel α . Aus der Skizze entnehmen wir die Bilder der Basisvektoren:

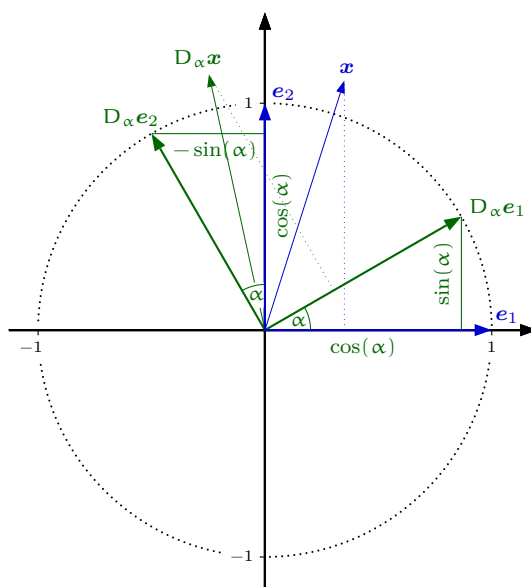


Abb. 6.1 Eine ebene Drehung

$$\mathbf{e}_1 \mapsto \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix}, \quad \mathbf{e}_2 \mapsto \begin{bmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{bmatrix}.$$

Daher ist

$$D(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}. \quad (6.18)$$

Als weiteres Beispiel für diese Regel verschaffen wir uns die Matrix, die zur Hintereinanderausführung zweier linearer Abbildungen $A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ und $B: \mathbb{K}^m \rightarrow \mathbb{K}^k$ gehört – diese Hintereinanderausführung ergibt eine lineare Abbildung von \mathbb{K}^n nach \mathbb{K}^k (wie man sich schnell überzeugt), die wir $B \cdot A$, meist aber einfach BA schreiben. Diese Matrix bezeichnen wir als *Matrixprodukt* von B mit A . Die Spaltenvektoren von BA erhalten wir durch Auswertung dieser Abbildung auf den kanonischen Basisvektoren \mathbf{e}_i :

$$BA\mathbf{e}_i = B(A\mathbf{e}_i) = B(\mathbf{a}_i) = B\mathbf{a}_i.$$

Die Darstellung gemäß (6.16) für BA durch die Spaltenvektoren lautet daher

$$BA = [B\mathbf{a}_1, B\mathbf{a}_2, \dots, B\mathbf{a}_n]. \quad (6.19)$$

Das ist eine konkrete Rechenvorschrift für die Matrixmultiplikation. Wir haben sie auf die Multiplikation Matrix \cdot Vektor (6.4) zurückgeführt: Die Spaltenvektoren von BA entstehen einfach durch Multiplikation der Spaltenvektoren \mathbf{a}_i von A mit B . Zum Beispiel für

$$A := \begin{bmatrix} -1 & 3 & 4 & 1 \\ 2 & 3 & 3 & -1 \end{bmatrix} \quad \text{und} \quad B := \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}$$

erhalten wir

$$BA = \left[\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right] = \begin{bmatrix} 3 & 9 & 10 & -1 \\ 0 & 9 & 11 & 1 \\ 4 & 6 & 6 & -2 \end{bmatrix}.$$

Ein interessanteres Beispiel ist die Hintereinanderausführung zweier Drehungen $D(\beta)$ und $D(\alpha)$. Sie muß die Drehung $D(\alpha + \beta)$ ergeben:

$$\begin{aligned} D(\alpha)D(\beta) &= \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{bmatrix} = D(\alpha + \beta). \end{aligned}$$

Vergleichen wir die Einträge, so finden wir die sogenannten *Additionssätze* für die trigonometrischen Funktionen \sin und \cos (vergl. auch Seite 79 und 276)

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \sin(\beta)\cos(\alpha), \quad (6.20)$$

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta). \quad (6.21)$$

Der Vollständigkeit halber seien noch zwei Abbildungen erwähnt, für die die zugehörigen Matrizen leicht anzugeben sind: Die erste ist die Abbildung $\mathcal{O}: \mathbb{K}^n \rightarrow \mathbb{K}^m$, die jedem Vektor $\mathbf{x} \in \mathbb{K}^n$ den Vektor $\mathbf{0} \in \mathbb{K}^m$ zuordnet: $\mathcal{O}(\mathbf{x}) := \mathbf{0}$. Das ist die sogenannte *Nullabbildung*. Die zweite ist die *identische Abbildung* $\text{id}: \mathbb{K}^n \rightarrow \mathbb{K}^n$, $\text{id}(\mathbf{x}) := \mathbf{x}$. Die zugehörigen Matrizen werden als *Nullmatrix* \mathbf{O}_n bzw. *Einheitsmatrix* $\mathbf{1}_n$ bezeichnet (wenn die Dimension n klar ist, meist einfach \mathbf{O} bzw. $\mathbf{1}$, wobei sich \mathbf{O} nicht durchgesetzt hat, weil es bequemer ist, dafür 0 zu schreiben und darauf zu vertrauen, daß aus dem Kontext klar wird, was gemeint ist). Sie sind offensichtlich durch

$$\mathbf{O}_n = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{1}_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad (6.22)$$

gegeben. Der Eintrag von $\mathbf{1}_n$ an der Position (i, k) ist δ_{ik} (vergl. (6.13)).

6.7 Direkte Zerlegung eines Vektorraums

6.7.1 Definition Für zwei Teilräume T_1 und T_2 eines Vektorraums V heißt

$$T_1 + T_2 := \{ \mathbf{x}_1 + \mathbf{x}_2 \mid \mathbf{x}_1 \in T_1, \mathbf{x}_2 \in T_2 \} \quad (6.23)$$

Summe der Teilräume T_1 und T_2 . Gilt $T_1 \cap T_2 = \{\mathbf{0}\}$, so heißt die Summe *direkt*. In diesem Fall schreibt man $T_1 \oplus T_2$.

6.7.2 Lemma $T_1 \oplus T_2$ ist ein Teilraum von V . Für einen Vektor $\mathbf{x} \in T_1 \oplus T_2$ ist die Zerlegung $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ in Vektoren $\mathbf{x}_1 \in T_1$ und $\mathbf{x}_2 \in T_2$ eindeutig. Es gilt

$$\dim T_1 \oplus T_2 = \dim T_1 + \dim T_2. \quad (6.24)$$

Beweis. Jeder Vektor $\mathbf{x} \in T_1 \oplus T_2$ hat per Definition eine Zerlegung $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, mit Vektoren $\mathbf{x}_1 \in T_1$ und $\mathbf{x}_2 \in T_2$. Gäbe es eine weitere Zerlegung $\mathbf{x} = \tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_2$ dieser Art, so müßte $\mathbf{x}_1 - \tilde{\mathbf{x}}_1 = \tilde{\mathbf{x}}_2 - \mathbf{x}_2$ gelten. Die linke Seite dieser Gleichung ist ein Vektor aus T_1 , die rechte einer aus T_2 . Dieser Vektor würde in $T_1 \cap T_2 = \{\mathbf{0}\}$ liegen, was $\mathbf{x}_1 = \tilde{\mathbf{x}}_1$ und $\tilde{\mathbf{x}}_2 = \mathbf{x}_2$ zur Folge hätte. Also gibt es keine von $\mathbf{x}_1 + \mathbf{x}_2$ verschiedene Zerlegung von \mathbf{x} .

Für $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \in T_1 \oplus T_2$, $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2 \in T_1 \oplus T_2$ und $t, s \in \mathbb{K}$ ist $t\mathbf{x} + s\mathbf{y} = t\mathbf{x}_1 + s\mathbf{y}_1 + t\mathbf{x}_2 + s\mathbf{y}_2$ wieder in $T_1 \oplus T_2$. Da es sich bei T_1 und T_2 nämlich um Teilräume handelt, ist $t\mathbf{x}_1 + s\mathbf{y}_1 \in T_1$ und $t\mathbf{x}_2 + s\mathbf{y}_2 \in T_2$.

Seien $B_1 := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ und $B_2 := \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ Basen von T_1 bzw. T_2 . B_1 und B_2 sind offensichtlich disjunkt, denn ein gemeinsames Element müßte in $T_1 \cap T_2$ liegen und daher $\mathbf{0}$ sein.

$B_1 \cup B_2$ ist linear unabhängig: Aus $\mathbf{0} = \sum_{k=1}^n \lambda_k \mathbf{b}_k + \sum_{k=1}^m \gamma_k \mathbf{c}_k$ folgt $\mathbf{0} = \sum_{k=1}^n \lambda_k \mathbf{b}_k$ und $\mathbf{0} = \sum_{k=1}^m \gamma_k \mathbf{c}_k$, denn die Zerlegung auch von $\mathbf{0}$ in $\mathbf{0} \in T_1$ und $\mathbf{0} \in T_2$ ist eindeutig. Die Basiseigenschaft von B_1 und B_2 ergibt $\lambda_1 = \dots = \lambda_n = 0$ und $\gamma_1 = \dots = \gamma_m = 0$.

$B_1 \cup B_2$ ist offensichtlich erzeugend und daher sogar eine Basis für $T_1 \oplus T_2$. Deshalb gilt $\dim T_1 \oplus T_2 = n + m = \dim T_1 + \dim T_2$. \square

6.7.3 Satz Für einen Teilraum T von V gibt es einen weiteren Teilraum S mit der Eigenschaft $V = T \oplus S$. Insbesondere gilt dann $\dim V = \dim T + \dim S$.

Beweis. Dieses Lemma ist im Wesentlichen ein Ergebnis des Basisergänzungssatzes 6.5.8. Der Teilraum T ist insbesondere ein Vektorraum und hat nach 6.5.5 eine Basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ mit einem m , nicht größer als $n := \dim V$. Ist $m = n$, so handelt es sich um den trivialen Fall $T = V$, für den $S := \{\mathbf{0}\}$ die formalen Kriterien des Lemmas erfüllt. Wir können also $m < n$ annehmen. Nach Satz 6.5.8 läßt sich \mathcal{B} zu einer Basis $\mathcal{B} \cup \{\mathbf{c}_1, \dots, \mathbf{c}_{n-m}\}$ von V ergänzen. Wir wählen $S := \text{lh}\{\mathbf{c}_1, \dots, \mathbf{c}_{n-m}\}$. Dann gilt $T \cap S = \{\mathbf{0}\}$, denn ein Vektor \mathbf{x} im Schnitt dieser Teilräume besäße Darstellungen $\mathbf{x} = \sum_{k=1}^m \lambda_k \mathbf{b}_k = \sum_{k=1}^{n-m} \gamma_k \mathbf{c}_k$. Ihre Differenz ergäbe eine nicht triviale Nullkombination der Basisvektoren, würden nicht alle Koeffizienten λ_k und γ_k einfach Null sein. Daher ist $\mathbf{x} = \mathbf{0}$. Bei $T + S$ handelt es sich damit sogar um die direkte Summe $T \oplus S$ von T und S . Die Vereinigung der Basen von T und S ist laut Konstruktion eine Basis für V , was sofort zu $V = T \oplus S$ und $\dim V = \dim T + \dim S$ führt. \square

Wie der Beweis zeigt, ist der Teilraum S durch T normalerweise nicht eindeutig festgelegt. Es gibt viele Möglichkeiten, eine Basis von T zu einer Basis von V zu ergänzen. Verfügen wir allerdings über ein Skalarprodukt auf V , dann kann S durch die Zusatzforderung, daß alle Vektoren aus S auf T senkrecht stehen sollen, eindeutig festgelegt werden.

6.7.4 Definition Sei U eine nicht leere Teilmenge des Vektorraums V . Dann heißt $U^\perp := \{\mathbf{x} \in V \mid \mathbf{y} \perp \mathbf{x} \text{ f. a. } \mathbf{y} \in U\}$ der zu U orthogonale Teilraum. Falls U nur einen Vektor \mathbf{x} enthält, schreiben wir \mathbf{x}^\perp statt $\{\mathbf{x}\}^\perp$.

Es ist eine gute Übung nachzurechnen, daß es sich bei U^\perp tatsächlich um einen Teilraum von V handelt und daß $\{0\}^\perp = V$, $V^\perp = \{0\}$ gilt.

6.7.5 Satz Für eine nicht leere Teilmenge $U \subseteq V$ ist $U^{\perp\perp}$ die lineare Hülle von U , also der kleinste Teilraum von V , der U umfaßt: $U^{\perp\perp} = \text{lh}(U)$. Wenn U sogar ein Teilraum von V ist, gilt also $U^{\perp\perp} = U$ und

$$V = U \oplus U^\perp. \quad (6.25)$$

Beweis. Ein Vektor $x \in \text{lh}(U)$ läßt sich als Linearkombination $x = \sum_{i=1}^p \lambda_i x_i$ von Vektoren $x_i \in U$ schreiben. Für jedes $y \in U^\perp$ gilt daher $\langle x | y \rangle = \sum_{i=1}^p \lambda_i \langle x_i | y \rangle = 0$. Das bedeutet $x \in U^{\perp\perp}$. Damit haben wir $\text{lh}(U) \subseteq U^{\perp\perp}$. Falls $\text{lh}(U) = V$ gelten sollte, ist nichts weiter zu zeigen. Wir können also von $\text{lh}(U) \subset V$ ausgehen. Für die umgekehrte Inklusion nehmen wir zunächst an, daß U nicht nur eine Teilmenge, sondern sogar ein (echter) Teilraum von V ist und zeigen erst einmal (6.25). Als endlich erzeugter Vektorraum hat U eine Basis $\mathcal{B} := \{b_1, \dots, b_k\}$, die mit dem GRAM-SCHMIDTSchen Orthonormalisierungsverfahren 6.5.16 zu einer ONB für U gemacht werden kann. Wir können also davon ausgehen, daß \mathcal{B} eine ONB ist. Nach dem Basisergänzungssatz 6.5.8 kann \mathcal{B} zu einer Basis $\{b_1, \dots, b_k, c_1, \dots, c_\ell\}$ von V ergänzt werden. Darauf wenden wir wieder das GRAM-SCHMIDTSche Orthonormalisierungsverfahren an. Dabei ändern sich die Vektoren b_i nicht. Deshalb können wir annehmen, daß die Vektoren c_j untereinander und auch zu allen b_i orthogonal sind. Das bedeutet, daß die Teilräume U und $W := \text{lh}(\{c_1, \dots, c_\ell\})$ orthogonal zueinander sind und zusammen eine direkte Zerlegung von V ergeben: $V = U \oplus W$. Offensichtlich gilt $W \subseteq U^\perp$. Jeder Vektor $x \in U^\perp$ läßt sich eindeutig in einen Anteil $x_1 \in U$ und einen Anteil $x_2 \in W$ zerlegen: $x = x_1 + x_2$. Das ergibt $\langle x | x_1 \rangle = 0 = \langle x_1 | x_1 \rangle + \langle x_2 | x_1 \rangle = \langle x_1 | x_1 \rangle$, also $x_1 = 0$ und damit $x \in W$. Daher folgt $U^\perp \subseteq W$ und daraus $U^\perp = W$. Damit ist (6.25) gezeigt. Für eine orthogonale direkte Zerlegung $V = U \oplus W$ ist also $W = U^\perp$. Wenden wir das auf die orthogonale Zerlegung $V = U^\perp \oplus U$ an, so erhalten wir $U = U^{\perp\perp}$.

Kommen wir auf den allgemeinen Fall zurück. Ein Vektor, der orthogonal zu U ist, ist auch orthogonal zu allen Linearkombinationen von Vektoren aus U , also orthogonal zu $\text{lh}(U)$. Das bedeutet $U^\perp \subseteq \text{lh}(U)^\perp$. Die umgekehrte Inklusion ist offensichtlich, so daß $U^\perp = \text{lh}(U)^\perp$ folgt. Jetzt ergibt sich $U^{\perp\perp} = \text{lh}(U)^{\perp\perp} = \text{lh}(U)$ aus den oben erhaltenen Ergebnissen, angewandt auf den Teilraum $\text{lh}(U)$. \square

6.7.6 Korollar Sei $0 \neq x \in V$. Dann gilt $\dim x^\perp = \dim V - 1$.

Beweis. Das folgt sofort aus $V = \text{lh}(x) \oplus x^\perp$, aus Lemma 6.7.2 und $\dim \text{lh}(x) = 1$. \square

6.8 Die Dimensionsformel

6.8.1 Definition Für eine lineare Abbildung $A : V \rightarrow W$ wird $\ker A := \{x \in V \mid Ax = 0\} \subseteq V$ als Kern von A bezeichnet. Mit $\operatorname{im} A := \{y \in W \mid \exists x \in V \ y = Ax\}$ wird, wie üblich, das Bild von A bezeichnet (vergl. 2.7).

Man überlege sich, daß $\ker A$ und $\operatorname{im} A$ Teilräume von V bzw. W sind.

6.8.2 Lemma Eine lineare Abbildung $A : V \rightarrow W$ ist genau dann injektiv, wenn $\ker A = \{0\}$ gilt. Für $A := [\alpha_1, \dots, \alpha_n]$ ist das genau dann der Fall, wenn die Menge $\{\alpha_1, \dots, \alpha_n\}$ der Spaltenvektoren von A linear unabhängig ist.

Beweis. Für eine lineare Abbildung A gilt immer $A0 = 0$ ($A0 = A00 = 0A0 = 0$). Ist A injektiv und $x \in \ker A$, so folgt $Ax = 0 = A0$, also $x = 0$. Daher muß $\ker A = \{0\}$ gelten.

Ist $\ker A = \{0\}$ und $Ax = Ay$, so folgt aus der Linearität $Ax - Ay = A(x - y) = 0$, also $x - y \in \ker A$ und daher $x = y$. Damit ist A eine injektive Abbildung.

Für einen Vektor $x = [x_1, \dots, x_n]^t \in V$ gilt $Ax = x_1\alpha_1 + \dots + x_n\alpha_n$. $Ax = 0$ für $x \neq 0$ ist daher äquivalent dazu, daß es eine nicht triviale Nullkombination $x_1\alpha_1 + \dots + x_n\alpha_n = 0$ der Spaltenvektoren von A gibt. Das ist äquivalent zur linearen Abhängigkeit von $\{\alpha_1, \dots, \alpha_n\}$. Damit haben wir gezeigt: A ist genau dann nicht injektiv, wenn $\{\alpha_1, \dots, \alpha_n\}$ linear abhängig ist. Damit ist alles gezeigt. \square

Injektive Abbildungen erhalten die lineare Unabhängigkeit von Mengen:

6.8.3 Proposition Für eine injektive lineare Abbildung $A : V \rightarrow W$ und eine linear unabhängige Menge $\mathcal{B} = \{b_1, \dots, b_m\} \subset V$ ist $A(\mathcal{B}) := \{Ab_1, \dots, Ab_m\} \subset W$ wieder linear unabhängig. Ist A sogar bijektiv, so gilt $\dim V = \dim W$ und $A(\mathcal{B})$ ist eine Basis für W .

Beweis. Aus $\sum_{k=1}^m \lambda_k Ab_k = 0$ folgt wegen der Linearität von A : $A \sum_{k=1}^m \lambda_k b_k = 0$, also $\sum_{k=1}^m \lambda_k b_k \in \ker A = \{0\}$. Da \mathcal{B} linear unabhängig ist, geht das nur für $\lambda_1 = \dots = \lambda_m = 0$. Daher enthält $A(\mathcal{B})$ nur triviale Nullkombinationen und ist somit ebenfalls linear unabhängig.

Für eine bijektive Abbildung A und eine Basis $\mathcal{B} = \{b_1, \dots, b_n\}$ von V ist $A(\mathcal{B})$ linear unabhängig. Für die Basiseigenschaft fehlt uns noch, daß $A(\mathcal{B})$ für W erzeugend ist:

Jedes $y \in W$ ist das Bild genau eines Vektors $x \in V$, mit einer Basisdarstellung $x = \sum_{k=1}^n x_k b_k$. Daher haben wir $y = Ax = \sum_{k=1}^n x_k Ab_k \in \operatorname{lh} A(\mathcal{B})$. Das zeigt, daß $A(\mathcal{B})$ erzeugend ist.

$A(\mathcal{B})$ hat wie \mathcal{B} genau n Elemente, so daß $\dim V = n = \dim W$ gelten muß. \square

6.8.4 Satz (Dimensionsformel) Für eine lineare Abbildung $A : V \rightarrow W$ zwischen den Vektorräumen V und W gilt

$$\dim V = \dim \ker A + \dim \operatorname{im} A. \quad (6.26)$$

Beweis. Die Beweisidee besteht darin, den Vektorraum V gemäß Lemma 6.7.3 in eine direkte Summe aus $\ker A$ und einem Teilraum V_1 aufzuspalten: $V = \ker A \oplus V_1$. $\ker A$ ist der Anteil von V , der auf $\mathbf{0}$ abgebildet wird. Dieser Teilraum trägt also zur Dimension des Bildes im A nichts bei. Dann muß V_1 der Teilraum sein, der den von $\mathbf{0}$ verschiedenen Anteil des Bildes erzeugt. Tatsächlich bildet A den Teilraum V_1 bijektiv auf $\operatorname{im} A$ ab, so daß nach Proposition 6.8.3 $\dim V_1 = \dim \operatorname{im} A$ gilt. Zusammen mit $\dim V = \dim \ker A + \dim V_1$ nach Lemma 6.7.2, ergibt das die Dimensionsformel.

Wir müssen somit nur noch zeigen, daß die Einschränkung $A|_{V_1} : V_1 \rightarrow \operatorname{im} A$ bijektiv ist.

Zunächst untersuchen wir den Kern von $A|_{V_1}$.

$\mathbf{x} \in \ker A|_{V_1}$ bedeutet $\mathbf{x} \in V_1$ und $A|_{V_1}\mathbf{x} = A\mathbf{x} = \mathbf{0}$. Das heißt $\mathbf{x} \in V_1 \cap \ker A = \{\mathbf{0}\}$, also $\mathbf{x} = \mathbf{0}$. Folglich ist $\ker A|_{V_1} = \{\mathbf{0}\}$ und daher $A|_{V_1}$ injektiv (Lemma 6.8.2).

Natürlich gilt $\operatorname{im} A|_{V_1} \subseteq \operatorname{im} A$, denn $A|_{V_1}$ ist ja die Einschränkung von A . Wir zeigen die umgekehrte Inklusion.

Zu jedem $\mathbf{y} \in \operatorname{im} A$ gibt es ein $\mathbf{x} \in V$ mit $\mathbf{y} = A\mathbf{x}$. \mathbf{x} läßt sich gemäß Lemma 6.7.2 eindeutig in seinen Anteil $\mathbf{x}_0 \in \ker A$ und $\mathbf{x}_1 \in V_1$ zerlegen: $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$. Dann gilt $\mathbf{y} = A\mathbf{x}_0 + A\mathbf{x}_1 = A\mathbf{x}_1 = A|_{V_1}\mathbf{x}_1 \in \operatorname{im} A|_{V_1}$. Da \mathbf{y} beliebig war, haben wir $\operatorname{im} A \subseteq \operatorname{im} A|_{V_1}$, insgesamt also $\operatorname{im} A|_{V_1} = \operatorname{im} A$ gezeigt. $A|_{V_1}$ ist daher surjektiv und somit bijektiv. \square

Eine der wichtigsten Folgerungen der Dimensionsformel ist, daß eine lineare Abbildung von V in denselben Vektorraum V schon dann bijektiv, also (auf ganz V) invertierbar ist, wenn sie injektiv, oder wenn sie surjektiv ist. Die jeweils zur Bijektivität fehlende Eigenschaft ist dann automatisch erfüllt. Für eine Abbildung $V \rightarrow W$ dagegen benötigt man normalerweise beide Eigenschaften.

6.8.5 Korollar *Eine lineare Abbildung $A : V \rightarrow V$ ist genau dann invertierbar, wenn sie injektiv, oder wenn sie surjektiv ist.*

Beweis. Ist A etwa injektiv, so gilt $\dim \ker A = 0$, denn $\ker A = \{\mathbf{0}\}$ nach Lemma 6.8.2. Nach der Dimensionsformel muß daher $\dim V = \dim \operatorname{im} A$ gelten, d. h. $\operatorname{im} A$ ist ein Teilraum von V mit derselben Dimension wie V . Das geht nur für $\operatorname{im} A = V$, also wenn A surjektiv und damit bijektiv ist.

Ist A surjektiv, so gilt $\operatorname{im} A = V$ und daher $\dim V = \dim \operatorname{im} A$. Laut Dimensionsformel folgt daraus $\dim \ker A = 0$, was nur für den 0-dimensionalen Teilraum $\{\mathbf{0}\} = \ker A$ möglich ist. Nach Lemma 6.8.2 ist A injektiv und daher auch bijektiv.

Die Umkehrung ist trivial. \square

6.8.6 A Zeigen Sie: Für eine lineare Abbildung $A : V \rightarrow W$ handelt es sich bei $\ker A$ und $\operatorname{im} A$ um einen Teilraum von V bzw. von W .

6.8.7 A Für eine bijektive lineare Abbildung $A : V \rightarrow W$ existiert die inverse Abbildung $A^{-1} : W \rightarrow V$. Zeigen Sie, daß diese ebenfalls linear ist.

6.9 Die inverse Matrix

6.9.1 Definition Die Menge der quadratischen $n \times n$ -Matrizen mit Einträgen aus einem Körper \mathbb{K} bezeichnen wir durch $M_n(\mathbb{K})$, oder, wenn der Körper aus dem Kontext heraus klar ist, auch einfach durch M_n .

6.9.2 Satz Für eine Matrix $A \in M_n(\mathbb{K})$, ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) folgt aus der Existenz einer Matrix $B \in M_n(\mathbb{K})$, die eine der beiden Bedingungen

$$BA = \mathbb{1} \quad \text{oder} \quad AB = \mathbb{1} \quad (6.27)$$

erfüllt, daß sie invertierbar ist und daß $A^{-1} = B$ gilt.

Beweis. Aus $BA = \mathbb{1}$ folgt für ein Element $\mathbf{x} \in \ker A$: $\mathbf{x} = BA\mathbf{x} = \mathbf{0}$. Das heißt, $\ker A = \{\mathbf{0}\}$. Daher ist A injektiv und nach Korollar 6.8.5 bijektiv, also invertierbar. Die Inverse ist bereits durch B gegeben, denn aus $BA = \mathbb{1}$ folgt durch Multiplikation mit A^{-1} von rechts: $B = BAA^{-1} = A^{-1}$.

Gilt $AB = \mathbb{1}$, dann liegt jedes $\mathbf{y} \in \mathbb{K}^n$ im Bild von A , denn $\mathbf{y} = AB\mathbf{y} \in \text{im } A$. Daher ist A surjektiv und nach Korollar 6.8.5 invertierbar. Wie oben zeigt man $B = A^{-1}$. \square

6.9.3 Korollar Für zwei Matrizen $A, B \in M_n(\mathbb{K})$ ist das Produkt AB genau dann invertierbar, wenn A und B invertierbar sind. In diesem Fall gilt $(AB)^{-1} = B^{-1}A^{-1}$.

Aus Satz 6.9.2 leiten wir uns ein Verfahren her, mit dem man gleichzeitig erfährt, ob A^{-1} existiert und wie A^{-1} gegebenenfalls aussieht. Dafür müssen wir nur die Gleichung $AB = \mathbb{1}$ nach B auflösen. Allerdings wissen wir im Augenblick noch nicht, wie das gehen könnte. Bisher haben wir nur lineare Vektorgleichungen $A\mathbf{b} = \mathbf{e}$ mit Hilfe des GAUSS-Verfahrens nach einem gesuchten Vektor \mathbf{b} aufgelöst, nicht eine Matrixgleichung $AB = \mathbb{1}$ nach einer Matrix B . Tatsächlich läßt sich letzteres aber auf das GAUSS-Verfahren zurückführen. Denn die Matrix $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$ ist bekannt, sobald wir ihre Spaltenvektoren \mathbf{b}_i kennen. Schreiben wir die Gleichung $AB = \mathbb{1}$ ausführlich auf und verwenden dabei $\mathbb{1} = [\mathbf{e}_1, \dots, \mathbf{e}_n]$, mit den kanonischen Basisvektoren \mathbf{e}_i :

$$AB = A[\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n] = [A\mathbf{b}_1, A\mathbf{b}_2, \dots, A\mathbf{b}_n] = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n] = \mathbb{1}.$$

Die Matrixgleichung $AB = \mathbb{1}$ ist also äquivalent zu den n Vektorgleichungen

$$A\mathbf{b}_1 = \mathbf{e}_1, \quad A\mathbf{b}_2 = \mathbf{e}_2, \quad \dots \quad A\mathbf{b}_n = \mathbf{e}_n.$$

Falls B existiert, müßten diese Gleichungen eindeutig nach den Vektoren \mathbf{b}_i auflösbar sein. Damit haben wir schon einen möglichen Lösungsweg gefunden: Wir müssen jede der n Gleichungen $A\mathbf{b}_i = \mathbf{e}_i$ mit dem GAUSS-Verfahren nach \mathbf{b}_i auflösen. Das hört sich allerdings nach viel Arbeit an. Überlegen wir uns doch einmal, was im Einzelnen dafür zu tun wäre. Zunächst würden wir die Gleichung $A\mathbf{b}_1 = \mathbf{e}_1$ in das zugehörige GAUSS-Schema verwandeln. Bei einer

inhomogenen Gleichung wie dieser, besteht das Schema aus der Matrix A , erweitert um die Inhomogenität \mathbf{e}_1 :

$$\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 1 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}$$

Jetzt verwenden wir das GAUSS-Verfahren. Da \mathbf{b}_1 eindeutig ist, muß es bis zur vollbesetzten Diagonalform durchführbar sein. Wir können sogar annehmen, daß die Diagonale nur die 1 als Einträge aufweist. Aus der rechten Seite \mathbf{e}_1 wird dabei der gesuchte Vektor $\mathbf{b}_1 = [b_{11}, b_{21}, \dots, b_{n1}]^t$:

$$\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 1 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array} \xrightarrow[\text{Verf.}]{\text{GAUSS}} \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_{11} \\ 0 & 1 & \cdots & 0 & b_{21} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{n1} \end{array}$$

Dasselbe machen wir zur Bestimmung von \mathbf{b}_2 aus $A\mathbf{b}_2 = \mathbf{e}_2$, von \mathbf{b}_3 usw., bis wir schließlich \mathbf{b}_n aus $A\mathbf{b}_n = \mathbf{e}_n$ bestimmt haben:

$$\begin{array}{c} \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 1 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array} \\ \downarrow \\ \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_{11} \\ 0 & 1 & \cdots & 0 & b_{21} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{n1} \end{array} \end{array} \quad \begin{array}{c} \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array} \\ \downarrow \\ \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_{12} \\ 0 & 1 & \cdots & 0 & b_{22} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{n2} \end{array} \end{array} \quad \cdots \quad \begin{array}{c} \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 1 \end{array} \\ \downarrow \\ \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_{1n} \\ 0 & 1 & \cdots & 0 & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{nn} \end{array} \end{array}$$

Für jeden der gesuchten Vektoren \mathbf{b}_1 bis \mathbf{b}_n ist ein GAUSS-Verfahren erforderlich. Die entscheidende Beobachtung ist aber, daß es sich dabei jedes mal um *dasselbe* Verfahren handelt, d. h., daß dabei jeweils dieselbe Abfolge von Zeilenumformungen durchgeführt werden. Die einzelnen Rechenschritte hängen nämlich nur von den Einträgen von A , nicht aber von den rechten Seiten $\mathbf{e}_1, \dots, \mathbf{e}_n$ ab. Daher können wir die n GAUSS-Verfahren zu einem einzigen *erweiterten GAUSS-Verfahren* bündeln, indem wir die rechten Seiten $\mathbf{e}_1, \dots, \mathbf{e}_n$ alle auf einmal berücksichtigen:

$$\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array} \xrightarrow[\text{Verf.}]{\text{GAUSS}} \begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & 1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{array}$$

In der praktischen Umsetzung geht man wie folgt vor. Anstatt die quadratische Matrix A daraufhin zu untersuchen, ob sie injektiv ist, was nach Lemma 6.8.2 die Bestimmung des Kerns von

A bedeuten würde, oder, was auf dasselbe hinausläuft, die Spaltenvektoren von A auf lineare Unabhängigkeit zu prüfen, versucht man einfach das erweiterte GAUSS-Verfahren durchzuführen. Gelingt das, dann wissen wir erstens, daß A invertierbar ist und wir kennen zweitens bereits A^{-1} . Gelingt es nicht, dann ist A nicht invertierbar. Für diese Erkenntnis haben wir dann nur etwa so viel Arbeit aufgewandt, wie nötig gewesen wäre, um zu zeigen, daß $Ax = \mathbf{0}$ eine Lösung $x \neq \mathbf{0}$ hat, oder daß die Spaltenvektoren von A linear abhängig sind. Dafür hätten wir nämlich dasselbe GAUSS-Verfahren durchführen müssen.

6.9.4 Beispiel Wir versuchen, die inverse Matrix der folgenden Matrix zu bestimmen:

$$A := \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{bmatrix}.$$

I	1	2	3	4	1	0	0	0	
II	4	1	2	3	0	1	0	0	II - 4 I
III	3	4	1	2	0	0	1	0	III - 3 I
IV	2	3	4	1	0	0	0	1	IV - 2 I
I	1	2	3	4	1	0	0	0	I · 7
II	0	-7	-10	-13	-4	1	0	0	II · (-1)
III	0	-2	-8	-10	-3	0	1	0	III · (-7)
IV	0	-1	-2	-7	-2	0	0	1	IV · (-7)
I	7	14	21	28	7	0	0	0	I - 2 II
II	0	7	10	13	4	-1	0	0	
III	0	14	56	70	21	0	-7	0	III - 2 II
IV	0	7	14	49	14	0	0	-7	IV - II
I	7	0	1	2	-1	2	0	0	36 I - III
II	0	7	10	13	4	-1	0	0	18 II - 5 III
III	0	0	36	44	13	2	-7	0	
IV	0	0	4	36	10	1	0	-7	9 IV - III
I	252	0	0	28	-49	70	7	0	(10 I - IV)/7
II	0	126	0	14	7	-28	35	0	(20 II - IV)/7
III	0	0	36	44	13	2	-7	0	(70 III - 11 IV)/7
IV	0	0	0	280	77	7	7	-63	9 IV/7
I	360	0	0	0	-81	99	9	9	
II	0	360	0	0	9	-81	99	9	
III	0	0	360	0	9	9	-81	99	
IV	0	0	0	360	99	9	9	-81	

Das zeigt bereits die Existenz der Matrix A^{-1} und daß sie durch

$$A^{-1} = \frac{1}{360} \begin{bmatrix} -81 & 99 & 9 & 9 \\ 9 & -81 & 99 & 9 \\ 9 & 9 & -81 & 99 \\ 99 & 9 & 9 & -81 \end{bmatrix} = \frac{1}{40} \begin{bmatrix} -9 & 11 & 1 & 1 \\ 1 & -9 & 11 & 1 \\ 1 & 1 & -9 & 11 \\ 11 & 1 & 1 & -9 \end{bmatrix}$$

gegeben ist. Es gibt eigentlich nie einen Grund, das Verfahren mit einer fehlerhaften inversen Matrix zu beenden. Wir können uns immer davon überzeugen, daß wir die richtige gefunden haben:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{bmatrix} \cdot \frac{1}{40} \begin{bmatrix} -9 & 11 & 1 & 1 \\ 1 & -9 & 11 & 1 \\ 1 & 1 & -9 & 11 \\ 11 & 1 & 1 & -9 \end{bmatrix} \\ = \frac{1}{40} \begin{bmatrix} -9+5+44 & 11-18+7 & 5+22-27 & 3+33-36 \\ -36+3+33 & 44-9+5 & 7+11-18 & 5+22-27 \\ -27+5+22 & 33-36+3 & 5+44-9 & 7+11-18 \\ -18+7+11 & 22-27+5 & 3+33-36 & 5+44-9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbb{1}.$$

Laut Satz 6.9.2 haben wir tatsächlich die Inverse von A berechnet.

Versuchen wir es mit der Matrix

$$B := \begin{bmatrix} 2 & 1 & 2 & 3 \\ -6 & -5 & 4 & 11 \\ 3 & 3 & 1 & -1 \\ -1 & 2 & 3 & 2 \end{bmatrix}.$$

I	2	1	2	3	1	0	0	0	
II	-6	-5	4	11	0	1	0	0	II + 3 I
III	3	3	1	-1	0	0	1	0	2 III - 3 I
IV	-1	2	3	2	0	0	0	1	2 IV + I
I	2	1	2	3	1	0	0	0	2 I + II
II	0	-2	10	20	3	1	0	0	
III	0	3	-4	-11	-3	0	2	0	2 III + 3 II
IV	0	5	8	7	1	0	0	2	2 IV + 5 II
I	4	0	14	26	5	1	0	0	
II	0	-2	10	20	3	1	0	0	
III	0	0	22	38	3	3	4	0	
IV	0	0	66	114	17	5	0	4	IV - 3 III
I	4	0	14	26	5	1	0	0	
II	0	-2	10	20	3	1	0	0	
III	0	0	22	38	3	3	4	0	
IV	0	0	0	0	8	-4	-12	4	

Die letzte Zeile zeigt, daß B^{-1} nicht existiert.

6.9.5 A Prüfen Sie, ob die folgenden Matrizen invertierbar sind und bestimmen Sie gegebenenfalls die Inversen.

$$\text{i) } A := \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \quad \text{ii) } B := \begin{bmatrix} 1 & 0 & 3 & 4 \\ 5 & 6 & 0 & 8 \\ 9 & 10 & 11 & 0 \\ 0 & 14 & 15 & 16 \end{bmatrix}$$

$$\begin{array}{ll}
 \text{iii)} & C(x) := \begin{bmatrix} 1 & x & x^2 & x^3 \\ x & 1 & x & x^2 \\ x^2 & x & 1 & x \\ x^3 & x^2 & x & 1 \end{bmatrix} & \text{iv)} & D := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\
 \text{v)} & E := \begin{bmatrix} 2 & -2 & 1 & 0 \\ 2i & i & -2i & 0 \\ i & 2i & 2i & i \\ 0 & 0 & 2-i & 0 \end{bmatrix} & \text{vi)} & F := \begin{bmatrix} 1 & 0 & 3 & 2 & 2 \\ 0 & 3 & 3 & 1 & 1 \\ 2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 4 & 0 & 0 & 4 & 2 \end{bmatrix}
 \end{array}$$

6.10 Die adjungierte Matrix

6.10.1 Definition Für eine $m \times n$ -Matrix A mit den Einträgen $a_{ij} \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$, oder $\mathbb{K} = \mathbb{C}$) an der Stelle (i, j) ist ihre adjungierte Matrix A^* die $n \times m$ -Matrix, mit den Einträgen $\overline{a_{ji}}$ an der Stelle (i, j) . A^* entsteht also aus A , indem man alle Einträge durch ihre konjugiert komplexen Werte ersetzt (natürlich nur für $\mathbb{K} = \mathbb{C}$) und diese Matrix anschließend transponiert:

$$A^* = \overline{A}^t \quad (6.28)$$

A^* wird auch die Adjungierte von A genannt.

Ausführlich aufgeschrieben bedeutet das

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{i1}} & \cdots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{i2}} & \cdots & \overline{a_{m2}} \\ \overline{a_{13}} & \overline{a_{23}} & \cdots & \overline{a_{i3}} & \cdots & \overline{a_{m3}} \\ \vdots & \vdots & & \vdots & & \vdots \\ \overline{a_{1j}} & \overline{a_{2j}} & \cdots & \overline{a_{ij}} & \cdots & \overline{a_{mj}} \\ \vdots & \vdots & & \vdots & & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{in}} & \cdots & \overline{a_{mn}} \end{bmatrix}.$$

In der knapperen Formulierung mit Hilfe der Spaltenvektoren \mathbf{a}_i :

$$A^* = [\mathbf{a}_1, \dots, \mathbf{a}_n]^* = \begin{bmatrix} \mathbf{a}_1^* \\ \vdots \\ \mathbf{a}_n^* \end{bmatrix},$$

mit den Zeilenvektoren $\mathbf{a}_j^* := [\overline{a_{1j}}, \overline{a_{2j}}, \dots, \overline{a_{mj}}]$.

Einen solchen Zeilenvektor sehen wir als eine $1 \times m$ -Matrix an.

Für $\mathbf{a} := [a_1, \dots, a_m]^t \in \mathbb{K}^m$ und $\mathbf{b} := [b_1, \dots, b_m]^t \in \mathbb{K}^m$ gilt

$$\langle \mathbf{a} | \mathbf{b} \rangle = \sum_{k=1}^m \overline{a_k} b_k = [\overline{a_1}, \overline{a_2}, \dots, \overline{a_m}] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \mathbf{a}^* \mathbf{b}. \quad (6.29)$$

Das läßt sich verallgemeinern: Für eine $m \times n$ -Matrix A und eine $m \times k$ -Matrix B gilt

$$\begin{aligned} A^*B &= \begin{bmatrix} \mathbf{a}_1^* \\ \mathbf{a}_2^* \\ \vdots \\ \mathbf{a}_n^* \end{bmatrix} [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k] = \begin{bmatrix} \mathbf{a}_1^* \mathbf{b}_1 & \mathbf{a}_1^* \mathbf{b}_2 & \cdots & \mathbf{a}_1^* \mathbf{b}_k \\ \mathbf{a}_2^* \mathbf{b}_1 & \mathbf{a}_2^* \mathbf{b}_2 & \cdots & \mathbf{a}_2^* \mathbf{b}_k \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_n^* \mathbf{b}_1 & \mathbf{a}_n^* \mathbf{b}_2 & \cdots & \mathbf{a}_n^* \mathbf{b}_k \end{bmatrix} \\ &= \begin{bmatrix} \langle \mathbf{a}_1 | \mathbf{b}_1 \rangle & \langle \mathbf{a}_1 | \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{a}_1 | \mathbf{b}_k \rangle \\ \langle \mathbf{a}_2 | \mathbf{b}_1 \rangle & \langle \mathbf{a}_2 | \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{a}_2 | \mathbf{b}_k \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{a}_n | \mathbf{b}_1 \rangle & \langle \mathbf{a}_n | \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{a}_n | \mathbf{b}_k \rangle \end{bmatrix}. \end{aligned} \quad (6.30)$$

6.10.2 Satz Die zu einer $m \times n$ -Matrix A adjungierte Matrix A^* erfüllt die Gleichung

$$\langle \mathbf{x} | A\mathbf{y} \rangle = \langle A^*\mathbf{x} | \mathbf{y} \rangle \quad (6.31)$$

f. a. $\mathbf{x} \in \mathbb{K}^m$, $\mathbf{y} \in \mathbb{K}^n$. A^* wird durch diese Beziehung eindeutig festgelegt.

Für eine $n \times k$ -Matrix B gilt

$$(AB)^* = B^*A^*. \quad (6.32)$$

Beweis. Für diese Gleichung gibt es mehrere Beweismöglichkeiten. Die kürzeste, möglicherweise aber nicht die anschaulichste, ist die folgende. Wir verwenden $\mathbf{x} := [x_1, \dots, x_m]^t$, $\mathbf{y} := [y_1, \dots, y_n]^t$ und $(A\mathbf{y})_i = \sum_{j=1}^n a_{ij}y_j$, sowie $(A^*\mathbf{x})_j = \sum_{i=1}^m \overline{a_{ij}}x_i (= \sum_{i=1}^m (A^*)_{ji}x_i)$:

$$\begin{aligned} \langle \mathbf{x} | A\mathbf{y} \rangle &= \sum_{i=1}^m \overline{x_i} (A\mathbf{y})_i = \sum_{i=1}^m \overline{x_i} \sum_{j=1}^n a_{ij}y_j = \sum_{j=1}^n \sum_{i=1}^m \overline{x_i} a_{ij} y_j \\ &= \sum_{j=1}^n \overline{\sum_{i=1}^m a_{ij}x_i} y_j = \sum_{j=1}^n \overline{(A^*\mathbf{x})_j} y_j = \langle A^*\mathbf{x} | \mathbf{y} \rangle. \end{aligned}$$

Man beachte, daß das erste Skalarprodukt auf \mathbb{K}^m und das letzte auf \mathbb{K}^n wirkt.

Zu Gleichung (6.32). Sei $\mathbf{z} := [z_1, \dots, z_k]^t \in \mathbb{K}^k$. Wir wenden (6.31) nacheinander auf A und dann auf B und ein Mal auf AB als Ganzes an:

$$\langle \mathbf{x} | AB\mathbf{z} \rangle = \langle A^*\mathbf{x} | B\mathbf{z} \rangle = \langle B^*A^*\mathbf{x} | \mathbf{z} \rangle = \langle (AB)^*\mathbf{x} | \mathbf{z} \rangle.$$

Das gilt für alle $\mathbf{x} \in \mathbb{K}^m$ und $\mathbf{z} \in \mathbb{K}^k$. Aus Übung 5.5.12 wissen wir, daß daraus $(AB)^*\mathbf{x} = B^*A^*\mathbf{x}$ für alle $\mathbf{x} \in \mathbb{K}^m$ folgt. Das ist äquivalent zu (6.32). Übrigens ergibt sich so auch die Eindeutigkeit von A^* : Aus $\langle A'\mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{x} | A\mathbf{y} \rangle$ für alle $\mathbf{x} \in \mathbb{K}^m$, $\mathbf{y} \in \mathbb{K}^n$ folgt $\langle A'\mathbf{x} | \mathbf{y} \rangle = \langle A^*\mathbf{x} | \mathbf{y} \rangle$ und daraus $A'\mathbf{x} = A^*\mathbf{x}$ für alle $\mathbf{x} \in \mathbb{K}^m$, also $A' = A^*$.

Eine anschaulichere Beweismöglichkeit für Gleichung (6.31) besteht darin, zuerst (6.32) und zwar zunächst für ein Matrixprodukt der Form C^*B zu zeigen (dafür muß C eine $n \times m$ -Matrix sein):

$$(C^*B)^* \stackrel{(6.30)}{=} \begin{bmatrix} \langle \mathbf{c}_1 | \mathbf{b}_1 \rangle & \cdots & \langle \mathbf{c}_1 | \mathbf{b}_k \rangle \\ \langle \mathbf{c}_2 | \mathbf{b}_1 \rangle & \cdots & \langle \mathbf{c}_2 | \mathbf{b}_k \rangle \\ \vdots & & \vdots \\ \langle \mathbf{c}_n | \mathbf{b}_1 \rangle & \cdots & \langle \mathbf{c}_n | \mathbf{b}_k \rangle \end{bmatrix}^* = \begin{bmatrix} \overline{\langle \mathbf{c}_1 | \mathbf{b}_1 \rangle} & \overline{\langle \mathbf{c}_2 | \mathbf{b}_1 \rangle} & \cdots & \overline{\langle \mathbf{c}_n | \mathbf{b}_1 \rangle} \\ \vdots & \vdots & & \vdots \\ \overline{\langle \mathbf{c}_1 | \mathbf{b}_k \rangle} & \overline{\langle \mathbf{c}_2 | \mathbf{b}_k \rangle} & \cdots & \overline{\langle \mathbf{c}_n | \mathbf{b}_k \rangle} \end{bmatrix}$$

$$= \begin{bmatrix} \langle \mathbf{b}_1 | \mathbf{c}_1 \rangle & \langle \mathbf{b}_1 | \mathbf{c}_2 \rangle & \cdots & \langle \mathbf{b}_1 | \mathbf{c}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{b}_k | \mathbf{c}_1 \rangle & \langle \mathbf{b}_k | \mathbf{c}_2 \rangle & \cdots & \langle \mathbf{b}_k | \mathbf{c}_n \rangle \end{bmatrix} = B^* C.$$

Jetzt ist (6.31) nur noch eine Frage des richtigen Einsetzens ($C := A^*$):

$$(AB)^* = ((A^*)^* B)^* = B^* A^*.$$

Durch Anwendung dieser Rechenregel auf den Spezialfall $(A^* \mathbf{x})^*$ erhalten wir (6.31)

$$\langle \mathbf{x} | A \mathbf{y} \rangle = \mathbf{x}^* (A \mathbf{y}) = (\mathbf{x}^* A) \mathbf{y} = (A^* \mathbf{x})^* \mathbf{y} = \langle A^* \mathbf{x} | \mathbf{y} \rangle. \quad \square$$

6.10.3 Korollar Eine Matrix $U \in M_n(\mathbb{K})$ vermittelt genau dann eine längentreue lineare Abbildung, also eine Abbildung, die für alle $\mathbf{x} \in \mathbb{K}^n$ die Gleichung $\|U\mathbf{x}\| = \|\mathbf{x}\|$ erfüllt, wenn sie winkeltreu ist, d. h., wenn sie $\langle U\mathbf{x} | U\mathbf{y} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ erfüllt. Das wiederum ist äquivalent zu $U^* = U^{-1}$.

Beweis. Wir zeigen hier nur den Fall $\mathbb{K} = \mathbb{C}$. Für $\mathbb{K} = \mathbb{R}$ verlaufen die Überlegungen analog. Die Längentreue ergibt mit der Polarisationsgleichung (5.81)

$$\begin{aligned} \langle U\mathbf{x} | U\mathbf{y} \rangle &= \frac{1}{4} [\|U\mathbf{x} + U\mathbf{y}\|^2 - \|U\mathbf{x} - U\mathbf{y}\|^2 - i(\|U\mathbf{x} + iU\mathbf{y}\|^2 - \|U\mathbf{x} - iU\mathbf{y}\|^2)] \\ &= \frac{1}{4} [\|U(\mathbf{x} + \mathbf{y})\|^2 - \|U(\mathbf{x} - \mathbf{y})\|^2 - i(\|U(\mathbf{x} + i\mathbf{y})\|^2 - \|U(\mathbf{x} - i\mathbf{y})\|^2)] \\ &= \frac{1}{4} [\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 - i(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2)] = \langle \mathbf{x} | \mathbf{y} \rangle \end{aligned}$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Daher ist U winkeltreu. Die Umkehrung folgt natürlich aus $\|U\mathbf{x}\|^2 = \langle U\mathbf{x} | U\mathbf{x} \rangle = \langle \mathbf{x} | \mathbf{x} \rangle = \|\mathbf{x}\|^2$. Mit unseren Überlegungen zur Adjungierten können wir aus der Winkeltreue

$$\langle \mathbf{x} | \mathbf{y} \rangle = \langle U\mathbf{x} | U\mathbf{y} \rangle = \langle U^* U\mathbf{x} | \mathbf{y} \rangle$$

die Gleichung $U^* U = \mathbb{1}$ folgern. Aus Satz 6.9.2 folgt, daß U invertierbar und daß U^{-1} durch U^* gegeben ist. Die Umkehrung sieht man genauso einfach: $\langle \mathbf{x} | \mathbf{y} \rangle = \langle U^{-1} U\mathbf{x} | \mathbf{y} \rangle = \langle U^* U\mathbf{x} | \mathbf{y} \rangle = \langle U\mathbf{x} | U\mathbf{y} \rangle. \quad \square$

6.10.4 Definition Eine lineare Abbildung $U \in M_n(\mathbb{K})$ mit der Eigenschaft $U^* U = \mathbb{1}$ heißt unitär. Im Fall $\mathbb{K} = \mathbb{R}$ wird sie auch oft orthogonal genannt.

6.10.5 A Bestimmen Sie die adjungierte Matrix von

$$A := \begin{bmatrix} 2i & (2-4i)e^{i\varphi} & \frac{2i}{1+2i} \\ 0 & 42i & -3e^{-4i} \\ \frac{-2i}{1-2i} & 5+2i & 9 \end{bmatrix}$$

und die Inverse von

$$U := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}.$$

6.10.6 Die duale Basis Um einen Vektor $\mathbf{x} \in V$ aus einem Vektorraum über \mathbb{K} nach einer Basis $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ zu entwickeln, müssen die Koeffizienten x_1, x_2, \dots, x_n in

$$\mathbf{x} = \sum_{k=1}^n x_k \mathbf{b}_k$$

bestimmt werden. Die Lösung fassen wir übersichtlich in einem Koeffizientenvektor $\mathbf{x}_{\mathcal{B}} := [x_1, \dots, x_n]_{\mathcal{B}}^t \in \mathbb{K}^n$ zusammen. Der Index \mathcal{B} bei $[x_1, \dots, x_n]_{\mathcal{B}}^t$ soll deutlich machen, auf welche Basis sich die Koeffizienten x_1, \dots, x_n beziehen. Das wird besonders dann wichtig, wenn mehrere Basen im Spiel sind, wie bei Basiswechseln (vergl. Abschnitt 6.11), da einem Vektor $[x_1, \dots, x_n]^t$ sonst nicht anzusehen ist, zu welcher Basis er gehört. Wir vereinbaren den Index generell wegzulassen, wenn es sich bei \mathcal{B} um die kanonische Basis (6.11) \mathcal{E} in \mathbb{K}^n handelt. Auch dann, wenn in einem Problemkreis nur eine einzige Basis relevant ist, werden wir den Index normalerweise ebenfalls einsparen.

Wir haben im Abschnitt 6.5.14 gesehen, wie leicht es ist, in einem Hilbertraum V einen Vektor \mathbf{x} nach einer ONB $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ zu entwickeln. Die Koeffizienten x_j sind nämlich einfach durch die Skalarprodukte $\langle \mathbf{b}_j | \mathbf{x} \rangle$ gegeben:

$$\langle \mathbf{b}_j | \mathbf{x} \rangle = \left\langle \mathbf{b}_j \left| \sum_{k=1}^n x_k \mathbf{b}_k \right. \right\rangle = \sum_{k=1}^n x_k \langle \mathbf{b}_j | \mathbf{b}_k \rangle = \sum_{k=1}^n \delta_{jk} x_k = x_j.$$

Wünschenswert wäre es, wenn uns eine solche Rechnung auch für eine beliebige Basis eines Hilbertraums möglich wäre. Hier kommt die sogenannte *duale Basis* $\mathcal{B}' := \{\mathbf{b}^1, \dots, \mathbf{b}^n\}$ von \mathcal{B} ins Spiel. Ihre Vektoren haben die maßgeschneiderte Eigenschaft

$$\langle \mathbf{b}^j | \mathbf{b}_k \rangle = \delta_{jk}. \quad (6.33)$$

Jetzt erhalten wir die Entwicklungskoeffizienten x_j von \mathbf{x} einfach durch $\langle \mathbf{b}^j | \mathbf{x} \rangle$:

$$\langle \mathbf{b}^j | \mathbf{x} \rangle = \left\langle \mathbf{b}^j \left| \sum_{k=1}^n x_k \mathbf{b}_k \right. \right\rangle = \sum_{k=1}^n x_k \langle \mathbf{b}^j | \mathbf{b}_k \rangle = \sum_{k=1}^n \delta_{jk} x_k = x_j.$$

Es stellt sich nur noch die Frage, wie man die duale Basis aus \mathcal{B} bestimmt. Eigentlich wissen wir das schon, wir müssen die Aufgabe nur von der richtigen Warte aus betrachten. Dabei hilft uns (6.30). In Verbindung mit (6.33) haben wir nämlich für die Matrizen $B := [\mathbf{b}_1, \dots, \mathbf{b}_n]$ und $B' := [\mathbf{b}^1, \dots, \mathbf{b}^n]$

$$B'^* B = \begin{bmatrix} \langle \mathbf{b}^1 | \mathbf{b}_1 \rangle & \langle \mathbf{b}^1 | \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{b}^1 | \mathbf{b}_n \rangle \\ \langle \mathbf{b}^2 | \mathbf{b}_1 \rangle & \langle \mathbf{b}^2 | \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{b}^2 | \mathbf{b}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{b}^n | \mathbf{b}_1 \rangle & \langle \mathbf{b}^n | \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{b}^n | \mathbf{b}_n \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbb{1}.$$

Das heißt: $B'^* = B^{-1}$, oder $B' = (B^{-1})^* = (B^*)^{-1}$. Die Inverse von B^* hat die duale Basis als Spaltenvektoren. Es ist klar, daß wir im Allgemeinen nicht billiger davon kommen können, als die Inverse der Matrix B zu berechnen. Denn bei Lichte besehen ist die duale Basis nur eine andere Lösung für die Aufgabe, für jedes $\mathbf{x} \in V$ die Gleichung $B\mathbf{x}_{\mathcal{B}} = \mathbf{x}$ nach $\mathbf{x}_{\mathcal{B}}$ aufzulösen. Und das läuft nun mal auf die Berechnung von B^{-1} hinaus.

6.10.7 Beispiel Wir wollen die duale Basis von

$$\mathcal{B} := \left\{ \begin{bmatrix} 2 \\ 2i \\ i \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ i \\ 2i \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2i \\ 2i \\ 2-i \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ i \\ 0 \end{bmatrix} \right\}$$

finden. In Übung 6.9.5 haben wir die Inverse der zugehörigen Matrix B bestimmt:

$$B^{-1} = \frac{1}{30} \begin{bmatrix} 5 & -10i & 0 & 3(2+i) \\ -10 & -10i & 0 & 6(2+i) \\ 0 & 0 & 0 & 6(2+i) \\ 15 & 30i & -30i & -27(2+i) \end{bmatrix}.$$

Die duale Basis ist demnach

$$\mathcal{B}' = \left\{ \frac{1}{30} \begin{bmatrix} 5 \\ 10i \\ 0 \\ 3(2-i) \end{bmatrix}, \frac{1}{15} \begin{bmatrix} -5 \\ 5i \\ 0 \\ 3(2-i) \end{bmatrix}, \frac{1}{5} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2-i \end{bmatrix}, \frac{1}{10} \begin{bmatrix} 5 \\ -10i \\ 10i \\ -9(2-i) \end{bmatrix} \right\}.$$

6.10.8 Beispiel Wir bestimmen die duale Basis $\mathcal{B}' := \{\mathbf{b}^1, \mathbf{b}^2\}$ einer beliebigen Basis $\mathcal{B} := \{\mathbf{b}_1, \mathbf{b}_2\}$ des \mathbb{R}^2 . Hier ist die Situation besonders übersichtlich. Das liegt daran, daß es (bis auf Streckung) nur einen Vektor gibt, der auf einem gegebenen senkrecht steht. Diesen erhalten wir einfach durch eine Drehung um $\frac{\pi}{2}$, also durch Anwendung der Matrix

$$J := \begin{bmatrix} \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (6.34)$$

Wie üblich sei $\mathbf{b}_1 = [b_{11}, b_{21}]^t$, $\mathbf{b}_2 = [b_{12}, b_{22}]^t$ und $B := [\mathbf{b}_1, \mathbf{b}_2]$.

$\mathbf{b}^1 \perp \mathbf{b}_2$ bedeutet $\mathbf{b}^1 \sim J\mathbf{b}_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} = \begin{bmatrix} -b_{22} \\ b_{12} \end{bmatrix}$. Für die Normierung berechnen

wir $\langle \mathbf{b}^1 | \mathbf{b}_1 \rangle \sim \left\langle \begin{bmatrix} -b_{22} \\ b_{12} \end{bmatrix} \middle| \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \right\rangle = -(b_{11}b_{22} - b_{12}b_{21}) = -\det(B)$. Dadurch ist \mathbf{b}^1 bestimmt: $\mathbf{b}^1 = \frac{-1}{\det(B)} J\mathbf{b}_2$. Genauso verfahren wir mit $\mathbf{b}^2 \sim J\mathbf{b}_1$ und erhalten $\mathbf{b}^2 = \frac{1}{\det(B)} J\mathbf{b}_1$. Zusammengefaßt:

$$\mathcal{B}' = \left\{ -\frac{1}{\det(B)} J\mathbf{b}_2, \frac{1}{\det(B)} J\mathbf{b}_1 \right\}, \quad B' = \frac{1}{\det(B)} \begin{bmatrix} b_{22} & -b_{21} \\ -b_{12} & b_{11} \end{bmatrix}. \quad (6.35)$$

Man rechnet leicht $B' = \frac{1}{\det(B)} JBJ^*$ nach und kann sich durch

$$B^*B' = \frac{1}{\det(B)} \begin{bmatrix} \mathbf{b}_1^* \\ \mathbf{b}_2^* \end{bmatrix} [J\mathbf{b}_1, J\mathbf{b}_2] J^* = \begin{bmatrix} \mathbf{b}_1^* \\ \mathbf{b}_2^* \end{bmatrix} [\mathbf{b}^2, -\mathbf{b}^1] J^* = \begin{bmatrix} \langle \mathbf{b}_1 | \mathbf{b}^2 \rangle & -\langle \mathbf{b}_1 | \mathbf{b}^1 \rangle \\ \langle \mathbf{b}_2 | \mathbf{b}^2 \rangle & -\langle \mathbf{b}_2 | \mathbf{b}^1 \rangle \end{bmatrix} J^* = JJ^* = \mathbb{1}$$

noch einmal davon überzeugen, daß B' die Inverse von B^* ist. Als kleines Nebenprodukt unserer Rechnung gewinnen wir eine (nicht sonderlich wichtige) Formel für die Inverse von B :

$$B^{-1} = \frac{1}{\det(B)} JB^*J^*. \quad (6.36)$$

6.10.9 Beispiel (Umkreis)

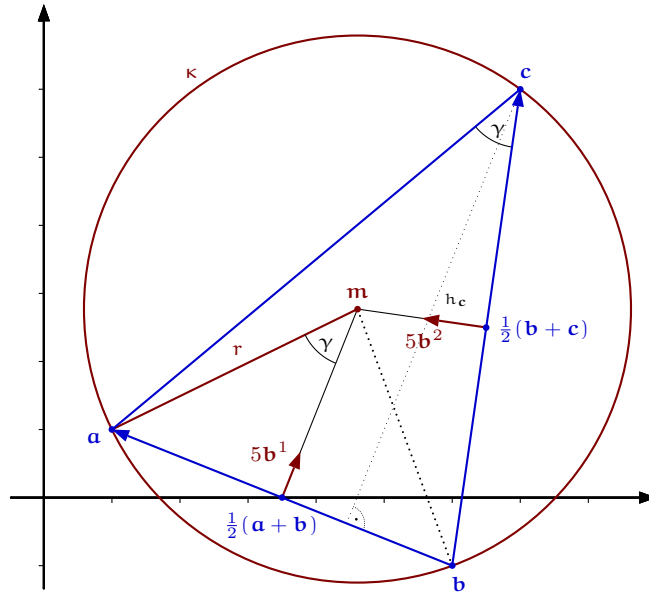


Abb. 6.2 Zum Umkreismittelpunkt und Umkreisradius

Drei Punkte \mathbf{a} , \mathbf{b} und \mathbf{c} liegen auf einem eindeutig bestimmten Kreis κ . Dabei vereinbaren wir, eine Gerade als einen Kreis mit unendlichem Radius anzusehen, denn nur so lässt sich die Aussage in ihrer Allgemeinheit aufrecht erhalten. Unsere Aufgabe besteht darin, den Mittelpunkt \mathbf{m} und den Radius r von κ aus \mathbf{a} , \mathbf{b} und \mathbf{c} zu berechnen.

$\mathcal{B} := \{\mathbf{b}_1, \mathbf{b}_2\}$, mit $\mathbf{b}_1 := \mathbf{c} - \mathbf{b}$ und $\mathbf{b}_2 := \mathbf{a} - \mathbf{b}$ ist eine dem Problem angepaßte Basis. $\mathcal{B}' := \{\mathbf{b}^1, \mathbf{b}^2\} = \{-\frac{1}{\lambda} \mathbf{J} \mathbf{b}_2, \frac{1}{\lambda} \mathbf{J} \mathbf{b}_1\}$ ist die zugehörige duale Basis (vergl. (6.35)) mit $A := \det(\mathbf{c} - \mathbf{b}, \mathbf{a} - \mathbf{b})$.

\mathbf{m} ist dadurch bestimmt, daß der Abstand von \mathbf{a} , \mathbf{b} und \mathbf{c} jeweils der gesuchte Radius r des Umkreises ist:

$$\begin{aligned} \|\mathbf{m} - \mathbf{a}\|^2 &= \|\mathbf{m}\|^2 + \|\mathbf{a}\|^2 - 2\langle \mathbf{m} | \mathbf{a} \rangle = r^2, \quad \|\mathbf{m} - \mathbf{b}\|^2 = \|\mathbf{m}\|^2 + \|\mathbf{b}\|^2 - 2\langle \mathbf{m} | \mathbf{b} \rangle = r^2, \\ \|\mathbf{m} - \mathbf{c}\|^2 &= \|\mathbf{m}\|^2 + \|\mathbf{c}\|^2 - 2\langle \mathbf{m} | \mathbf{c} \rangle = r^2. \end{aligned}$$

Nach Subtraktion der ersten beiden bzw. der letzten beiden Gleichungen erhalten wir:

$$\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2 = 2\langle \mathbf{m} | \mathbf{a} - \mathbf{b} \rangle = 2\langle \mathbf{m} | \mathbf{b}_2 \rangle, \quad \|\mathbf{c}\|^2 - \|\mathbf{b}\|^2 = 2\langle \mathbf{m} | \mathbf{c} - \mathbf{b} \rangle = 2\langle \mathbf{m} | \mathbf{b}_1 \rangle.$$

Damit gewinnen wir die Entwicklung $\mathbf{m} = \lambda_1 \mathbf{b}^1 + \lambda_2 \mathbf{b}^2$ von \mathbf{m} in der dualen Basis:

$$\langle \mathbf{m} | \mathbf{b}_1 \rangle = \lambda_1 \langle \mathbf{b}^1 | \mathbf{b}_1 \rangle + \lambda_2 \langle \mathbf{b}^2 | \mathbf{b}_1 \rangle = \lambda_1 = \frac{1}{2} (\|\mathbf{c}\|^2 - \|\mathbf{b}\|^2),$$

$$\langle \mathbf{m} | \mathbf{b}_2 \rangle = \lambda_1 \langle \mathbf{b}^1 | \mathbf{b}_2 \rangle + \lambda_2 \langle \mathbf{b}^2 | \mathbf{b}_2 \rangle = \lambda_2 = \frac{1}{2} (\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2).$$

$$\begin{aligned} \mathbf{m} &= \frac{\mathbf{J}}{2A} (-(\|\mathbf{c}\|^2 - \|\mathbf{b}\|^2)(\mathbf{a} - \mathbf{b}) + (\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2)(\mathbf{c} - \mathbf{b})) \\ &= \mathbf{J} \frac{\mathbf{a}(\|\mathbf{b}\|^2 - \|\mathbf{c}\|^2) + \mathbf{b}(\|\mathbf{c}\|^2 - \|\mathbf{a}\|^2) + \mathbf{c}(\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2)}{2 \det(\mathbf{c} - \mathbf{b}, \mathbf{a} - \mathbf{b})} \\ &= \mathbf{J} \frac{\mathbf{a}(\|\mathbf{b}\|^2 - \|\mathbf{c}\|^2) + \mathbf{b}(\|\mathbf{c}\|^2 - \|\mathbf{a}\|^2) + \mathbf{c}(\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2)}{2(\det(\mathbf{a}, \mathbf{b}) + \det(\mathbf{b}, \mathbf{c}) + \det(\mathbf{c}, \mathbf{a}))} \end{aligned} \quad (6.37)$$

Diese Formel zeigt sehr deutlich die vollkommene Symmetrie in den Punkten \mathbf{a} , \mathbf{b} und \mathbf{c} , die zu erwarten war, da keiner dieser Punkte vor den anderen ausgezeichnet ist.

Der Radius: Aus der Skizze ist $r = \frac{\|\mathbf{a} - \mathbf{b}\|}{2 \sin(\gamma)} = \frac{\|\mathbf{a} - \mathbf{b}\|}{2 \sqrt{1 - \cos^2(\gamma)}}$ abzulesen. Nach dem Satz vom

Faßkreisbogen (siehe Seite 14-L) ist γ auch bei \mathbf{c} zu finden: $\cos(\gamma) = \frac{\langle \mathbf{a} - \mathbf{c} | \mathbf{b} - \mathbf{c} \rangle}{\|\mathbf{a} - \mathbf{c}\| \|\mathbf{b} - \mathbf{c}\|}$. Daher ist

$$r = \frac{\|\mathbf{a} - \mathbf{b}\|}{2 \sqrt{1 - \cos^2(\gamma)}} = \frac{\|\mathbf{a} - \mathbf{b}\| \|\mathbf{a} - \mathbf{c}\| \|\mathbf{b} - \mathbf{c}\|}{2 \sqrt{\|\mathbf{a} - \mathbf{c}\|^2 \|\mathbf{b} - \mathbf{c}\|^2 - \langle \mathbf{a} - \mathbf{c} | \mathbf{b} - \mathbf{c} \rangle^2}}$$

$$\begin{aligned}
&= \frac{\|\mathbf{a} - \mathbf{b}\| \|\mathbf{b} - \mathbf{c}\| \|\mathbf{c} - \mathbf{a}\|}{2\sqrt{\det(\mathbf{a} - \mathbf{c}, \mathbf{b} - \mathbf{c})^2}} = \frac{\|\mathbf{a} - \mathbf{b}\| \|\mathbf{b} - \mathbf{c}\| \|\mathbf{c} - \mathbf{a}\|}{2|\det(\mathbf{a} - \mathbf{c}, \mathbf{b} - \mathbf{c})|}, \\
r &= \frac{\|\mathbf{a} - \mathbf{b}\| \|\mathbf{b} - \mathbf{c}\| \|\mathbf{c} - \mathbf{a}\|}{2|\det(\mathbf{a}, \mathbf{b}) + \det(\mathbf{b}, \mathbf{c}) + \det(\mathbf{c}, \mathbf{a})|} = \frac{\|\mathbf{a} - \mathbf{b}\| \|\mathbf{b} - \mathbf{c}\| \|\mathbf{c} - \mathbf{a}\|}{2|\det(\mathbf{a} - \mathbf{b}, \mathbf{c} - \mathbf{b})|}. \quad (6.38)
\end{aligned}$$

Dabei haben wir verwendet, daß für Vektoren \mathbf{x} und \mathbf{y} , die den Winkel α einschließen,

$$\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - \langle \mathbf{x} | \mathbf{y} \rangle^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 (1 - \cos^2(\alpha)) = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2(\alpha) = \det(\mathbf{x}, \mathbf{y})^2$$

gilt. Denn $\|\mathbf{x}\| \|\mathbf{y}\| \sin(\alpha)$ ist der Flächeninhalt des von \mathbf{x} und \mathbf{y} aufgespannten Parallelogramms.

Für $\mathbf{a} := [1, 1]^t$, $\mathbf{b} := [6, -1]^t$ und $\mathbf{c} := [7, 6]^t$ ist $A = 37$, $\|\mathbf{a}\|^2 = 2$, $\|\mathbf{b}\|^2 = 37$ und $\|\mathbf{c}\|^2 = 85$.

Damit erhalten wir

$$\mathbf{m} = \frac{1}{74} \left(-48 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 83 \begin{bmatrix} 1 \\ 6 \end{bmatrix} - 35 \begin{bmatrix} -6 \\ 7 \end{bmatrix} \right) = \frac{1}{74} \begin{bmatrix} 341 \\ 205 \end{bmatrix}, \quad r = \frac{\sqrt{29 \cdot 50 \cdot 61}}{74} = \frac{5\sqrt{3538}}{74}.$$

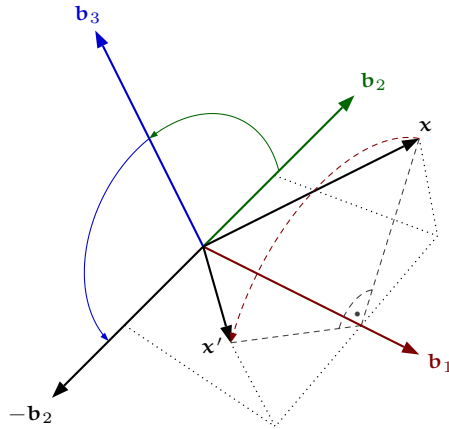
Eine kleine Zusatzüberlegung läßt uns auch eine Formel für die Höhen h_a , h_b , h_c durch \mathbf{a} , \mathbf{b} und \mathbf{c} finden: Der Flächeninhalt des Parallelogramms, zu dem man das Dreieck ergänzen kann, ist $A = \det(\mathbf{b} - \mathbf{a}, \mathbf{c} - \mathbf{a}) = \det(\mathbf{c} - \mathbf{b}, \mathbf{a} - \mathbf{b}) = \det(\mathbf{a} - \mathbf{c}, \mathbf{b} - \mathbf{c}) = h_a \|\mathbf{b} - \mathbf{c}\| = h_b \|\mathbf{c} - \mathbf{a}\| = h_c \|\mathbf{a} - \mathbf{b}\|$. Das ergibt

$$h_a = \frac{\det(\mathbf{b} - \mathbf{a}, \mathbf{c} - \mathbf{a})}{\|\mathbf{b} - \mathbf{c}\|}, \quad h_b = \frac{\det(\mathbf{c} - \mathbf{b}, \mathbf{a} - \mathbf{b})}{\|\mathbf{c} - \mathbf{a}\|}, \quad h_c = \frac{\det(\mathbf{a} - \mathbf{c}, \mathbf{b} - \mathbf{c})}{\|\mathbf{a} - \mathbf{b}\|}. \quad (6.39)$$

In unserem Beispiel: $h_a = \frac{37}{\sqrt{50}}$, $h_b = \frac{37}{\sqrt{61}}$ und $h_c = \frac{37}{\sqrt{29}}$.

6.11 Koordinatentransformation

Es gibt immer wieder Aufgaben, deren Lösung in der gegebenen Basisdarstellung schwer, in einer anderen, geeigneteren Basis aber leicht zu finden ist. Wünschenswert wäre dann ein Verfahren, mit dem die beteiligten Vektoren in dieser neuen Basis dargestellt werden. Dann löst man das Problem und transformiert das Ergebnis wieder zurück in die Ausgangsbasis.



Um ein konkretes Beispiel zu haben, stellen wir uns folgende Aufgabe: Wir wollen den Vektor $\mathbf{x} := [1, 3, 1]^t$ um 90° um die Drehachse $\mathbf{n} := [1, 2, 2]^t$ drehen. Da eine Drehung eine lineare Abbildung ist, könnten wir versuchen, die Drehmatrix zu finden. Wie bei jeder linearen Abbildung brauchen wir dafür nur die Bilder der kanonischen Basisvektoren anzugeben. Das ist zwar richtig, aber es gibt keine Garantie dafür, daß das immer leicht zu bewerkstelligen ist. Und in diesem Fall ist es tatsächlich nicht einfach. Eine dem Problem angepasste Basis $\mathcal{B} := \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ läßt sich folgendermaßen gewinnen: Wir wählen $\mathbf{b}_1 \sim \mathbf{n}$. \mathbf{b}_2 und \mathbf{b}_3 sind eine Ergänzung von \mathbf{b}_1 zu einer ONB von \mathbb{R}^3 . In dieser neuen Basis handelt es sich jetzt um die Aufgabe, eine 90° -Drehung um die \mathbf{b}_1 -Achse auszuführen. Dabei wird \mathbf{b}_1 in \mathbf{b}_1 , \mathbf{b}_2 in $\mathbf{b}'_2 := \mathbf{b}_3$ und \mathbf{b}_3 in $\mathbf{b}'_3 := -\mathbf{b}_2$ abgebildet. Die Darstellung von \mathbf{x} in dieser Basis ist leicht. Gemäß Abschnitt 6.5.14 haben wir nur die Skalarprodukte $\langle \mathbf{b}_1 | \mathbf{x} \rangle$, $\langle \mathbf{b}_2 | \mathbf{x} \rangle$ und $\langle \mathbf{b}_3 | \mathbf{x} \rangle$ auszurechnen, um

$$\mathbf{x} = \langle \mathbf{b}_1 | \mathbf{x} \rangle \mathbf{b}_1 + \langle \mathbf{b}_2 | \mathbf{x} \rangle \mathbf{b}_2 + \langle \mathbf{b}_3 | \mathbf{x} \rangle \mathbf{b}_3$$

und daher

$$\mathbf{x}' = \langle \mathbf{b}_1 | \mathbf{x} \rangle \mathbf{b}'_1 + \langle \mathbf{b}_2 | \mathbf{x} \rangle \mathbf{b}'_2 + \langle \mathbf{b}_3 | \mathbf{x} \rangle \mathbf{b}'_3 = \langle \mathbf{b}_1 | \mathbf{x} \rangle \mathbf{b}_1 + \langle \mathbf{b}_2 | \mathbf{x} \rangle \mathbf{b}_3 - \langle \mathbf{b}_3 | \mathbf{x} \rangle \mathbf{b}_2$$

zu erhalten. In unserem Beispiel wählen wir

$$\mathbf{b}_1 := \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{b}_2 := \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{b}_3 := \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ -1 \end{bmatrix}.$$

Dann ist $\tilde{x}_1 := \langle \mathbf{x} | \mathbf{b}_1 \rangle = 3$, $\tilde{x}_2 := \langle \mathbf{x} | \mathbf{b}_2 \rangle = 1$ und $\tilde{x}_3 := \langle \mathbf{x} | \mathbf{b}_3 \rangle = 1$. Das heißt $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$ und somit

$$\mathbf{x}' = 3\mathbf{b}_1 + \mathbf{b}_3 - \mathbf{b}_2 = \frac{1}{3} \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ -1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 \\ 7 \\ 7 \end{bmatrix}.$$

Übrigens darf das Ergebnis \mathbf{x}' der Drehung natürlich nicht von der Realisierung der Basisvektoren \mathbf{b}_1 , \mathbf{b}_2 und \mathbf{b}_3 abhängen. Die Wahl von \mathbf{b}_2 ist keineswegs eindeutig. Es muß nur ein Vektor

gesucht werden, der senkrecht auf \mathbf{b}_1 steht. Davon gibt es viele, etwa $\mathbf{b}_2 := \frac{1}{\sqrt{5}} [2, 0, -1]^t$. $\mathbf{b}_3 := \frac{1}{3\sqrt{5}} [-2, 5, -4]^t$ ist dann über das Kreuzprodukt festgelegt. Es ist eine gute Übung, die Rechnungen mit diesen Vektoren zu wiederholen.

Jetzt wollen wir unser Beispiel ausbauen und beliebige Drehwinkel α zulassen. \mathbf{b}_1 geht dabei weiterhin in \mathbf{b}_1 über, aber \mathbf{b}_2 in $\mathbf{b}'_2 := \cos(\alpha)\mathbf{b}_2 + \sin(\alpha)\mathbf{b}_3$ und \mathbf{b}_3 in $\mathbf{b}'_3 := -\sin(\alpha)\mathbf{b}_2 + \cos(\alpha)\mathbf{b}_3$ (vergl. Abb. 6.1). Die Basisdarstellung $\mathbf{x} = \tilde{x}_1\mathbf{b}_1 + \tilde{x}_2\mathbf{b}_2 + \tilde{x}_3\mathbf{b}_3$ hat sich nicht geändert. Daher wird

$$\begin{aligned}\mathbf{x}' &= \tilde{x}_1\mathbf{b}'_1 + \tilde{x}_2\mathbf{b}'_2 + \tilde{x}_3\mathbf{b}'_3 \\ &= \tilde{x}_1\mathbf{b}_1 + \tilde{x}_2(\cos(\alpha)\mathbf{b}_2 + \sin(\alpha)\mathbf{b}_3) + \tilde{x}_3(-\sin(\alpha)\mathbf{b}_2 + \cos(\alpha)\mathbf{b}_3) \\ &= \tilde{x}_1\mathbf{b}_1 + (\cos(\alpha)\tilde{x}_2 - \sin(\alpha)\tilde{x}_3)\mathbf{b}_2 + (\sin(\alpha)\tilde{x}_2 + \cos(\alpha)\tilde{x}_3)\mathbf{b}_3 \\ &= \tilde{x}'_1\mathbf{b}_1 + \tilde{x}'_2\mathbf{b}_2 + \tilde{x}'_3\mathbf{b}_3.\end{aligned}$$

Man sieht, daß die eigentliche Berechnung der Koordinaten \tilde{x}'_i von \mathbf{x}' bezüglich der Basis \mathcal{B} nur in den Koordinaten \tilde{x}_i von \mathbf{x} bezüglich dieser Basis stattfinden. Die Basisvektoren \mathbf{b}_i selbst nehmen daran gar nicht teil. Sie haben nur eine ordnende Rolle, indem sie sozusagen ihre Koeffizienten aufsammeln. Außerdem braucht man sie natürlich, um das Ergebnis wieder in der Ausgangsbasis angeben zu können, oder einfacher gesagt, um \mathbf{x}' letztendlich auszurechnen. Da uns \mathcal{B} ja bekannt ist, handelt es sich dabei nur noch um simples Einsetzen. Das gestellte Problem wurde im Wesentlichen mit dem Auffinden der Bildkoordinaten \tilde{x}'_i bezüglich \mathcal{B} gelöst. Es sollte daher möglich sein, ein Verfahren anzugeben, das allein mit den Koordinaten auskommt. Wir werden dabei mit Spaltenvektoren arbeiten, deren Einträge sich auf verschiedene Basen beziehen. Damit man dabei den Bezug nicht verliert, vereinbaren wir ein für alle Mal folgende *Regelung*:

Wenn nur eine Basis in Betracht kommt, meistens ist es die kanonische Basis $\mathcal{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, dann schreiben wir den Spaltenvektor, der die Koordinaten von \mathbf{x} bzgl. dieser Basis enthält, *ohne* Index

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Das ist die überwiegend anzutreffende Situation. Außerdem identifizieren wir meistens den Vektor \mathbf{x} mit seiner Koordinatenspalte. Wenn dagegen mehrere Basen nebeneinander Verwendung finden, machen wir durch einen Index am Spaltenvektor deutlich, auf welche Basis sich die Koordinaten beziehen. Für eine Basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ etwa fügen wir in der Kurznotation \mathbf{x} und in der ausführlichen Koordinatenschreibweise einen Index \mathcal{B} an:

$$\mathbf{x}_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{\mathcal{B}} = [x_1, x_2, \dots, x_n]_{\mathcal{B}}^t. \quad (6.40)$$

$\mathbf{x}_{\mathcal{B}}$ ist der Koordinatenvektor bzgl. der Basis \mathcal{B} für den Vektor $\mathbf{x} = x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + \cdots + x_n \mathbf{b}_n$. In unserem Beispiel sieht das folgendermaßen aus:

$$\mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \quad \mathbf{x}_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}}, \quad \mathbf{x}'_{\mathcal{B}} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}_{\mathcal{B}}, \quad \mathbf{x}' = \frac{1}{3} \begin{bmatrix} -1 \\ 7 \\ 7 \end{bmatrix}.$$

$\mathbf{x}_{\mathcal{B}}$ ist der *Koordinatenvektor von \mathbf{x} bzgl. der Basis \mathcal{B}* .

Kehren wir zu unserem erweiterten Beispiel zurück und arbeiten konsequent mit den Koordinatenvektoren: $\mathbf{x}_{\mathcal{B}} = [\tilde{x}_1, \tilde{x}_2, \tilde{x}_3]^t_{\mathcal{B}}$,

$$\mathbf{x}'_{\mathcal{B}} = \begin{bmatrix} \tilde{x}'_1 \\ \tilde{x}'_2 \\ \tilde{x}'_3 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} \tilde{x}_1 \\ \cos(\alpha)\tilde{x}_2 - \sin(\alpha)\tilde{x}_3 \\ \sin(\alpha)\tilde{x}_2 + \cos(\alpha)\tilde{x}_3 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}_{\mathcal{B}} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix}_{\mathcal{B}} =: D(\alpha)_{\mathcal{B}} \mathbf{x}_{\mathcal{B}}.$$

Die Matrix $D(\alpha)_{\mathcal{B}}$ ist im Wesentlichen eine ebene Drehung, wie in (6.18). Die 2- und die 3-Achse werden um den Winkel α gedreht, während die 1-Achse unverändert bleibt. Die günstige Wahl der Basisvektoren und das Rechnen mit den Koordinatenvektoren hat das deutlich hervortreten lassen.

Drehen wir den Vektor $\mathbf{x} = [1, 3, 1]^t$ um 45° . Zunächst in der Basis \mathcal{B} :

$$\mathbf{x}'_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}_{\mathcal{B}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ \sqrt{2} \end{bmatrix}_{\mathcal{B}}.$$

Daraus erhalten wir den Ergebnisvektor \mathbf{x}' durch

$$\mathbf{x}' = 3 \mathbf{b}_1 + \sqrt{2} \mathbf{b}_3 = \frac{1}{3} \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -2\sqrt{2} \\ 2\sqrt{2} \\ -\sqrt{2} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 3 - 2\sqrt{2} \\ 6 + 2\sqrt{2} \\ 6 - \sqrt{2} \end{bmatrix}.$$

Übrigens sollte die Länge von \mathbf{x}' mit der Länge $\sqrt{11}$ von \mathbf{x} übereinstimmen, denn eine Drehung ändert sie nicht: $\|\mathbf{x}'\|^2 = \frac{1}{9} (9 + 8 - 12\sqrt{2} + 36 + 8 + 24\sqrt{2} + 36 + 2 - 12\sqrt{2}) = \frac{99}{9} = 11$.

Die Rücktransformation des Vektors $\mathbf{x}'_{\mathcal{B}}$ aus der Basis \mathcal{B} in die Ausgangsbasis gestaltet sich sehr einfach nach folgendem Schema:

$$\mathbf{x}'_{\mathcal{B}} = \begin{bmatrix} \tilde{x}'_1 \\ \tilde{x}'_2 \\ \tilde{x}'_3 \end{bmatrix}_{\mathcal{B}} \Rightarrow \mathbf{x}' = \tilde{x}'_1 \mathbf{b}_1 + \tilde{x}'_2 \mathbf{b}_2 + \tilde{x}'_3 \mathbf{b}_3 = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3] \cdot \begin{bmatrix} \tilde{x}'_1 \\ \tilde{x}'_2 \\ \tilde{x}'_3 \end{bmatrix}_{\mathcal{B}} = \mathbf{B} \mathbf{x}'_{\mathcal{B}},$$

mit der Transformationsmatrix $\mathbf{B} := [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$, die einfach die Basisvektoren als Spaltenvektoren enthält. Ist die Basis \mathcal{B} also bekannt, so läßt sich \mathbf{B} sofort angeben. Wie oben zu sehen ist, transformiert sie die Koordinaten aus der neuen Basis \mathcal{B} in die Ausgangsbasis. \mathbf{B} ist seiner Natur nach eine Rücktransformation, denn meistens will man die Koordinaten eines Vektors

bzgl. der neuen Basis aus den bekannten der Ausgangsbasis berechnen, bei der es sich überwiegend um die kanonische handelt. In der Gleichung $\mathbf{x} = B\mathbf{x}_B$ ist also normalerweise der Vektor \mathbf{x}_B auf der rechten Seite gesucht, während \mathbf{x} bekannt ist. Das bedeutet, daß wir B auf die andere Seite bringen müssen – was kein Problem darstellt, wenn B invertierbar ist: $\mathbf{x}_B = B^{-1}\mathbf{x}$. Und natürlich ist B invertierbar, denn seine Spaltenvektoren bilden eine Basis und sind daher linear unabhängig (vgl. Lemma 6.8.2). Weil es sich bei B sogar um eine ONB handelt, ist die Inverse nach Korollar 6.10.3 einfach durch B^* gegeben. Damit können wir jetzt das Verfahren systematisch darstellen:

- i) Transformiere den Vektor \mathbf{x} in die neue Basis B , d. h., berechne den Koordinatenvektor \mathbf{x}_B von \mathbf{x} bzgl. B : $\mathbf{x}_B = B^{-1}\mathbf{x}$.
- ii) Auf \mathbf{x}_B wird die lineare Abbildung A angewandt. In der Basis B ist sie durch die Matrix A_B gegeben: $\mathbf{x}'_B = A_B\mathbf{x}_B$.
- iii) Durch die Matrix B wird das Ergebnis zurück in die Ausgangsbasis transformiert: $\mathbf{x}' = B\mathbf{x}'_B$.

Dieser Vorgang kann kompakt und übersichtlich durch $\mathbf{x}' = BA_B B^{-1}\mathbf{x}$ zusammengefaßt werden. Wenn wir uns erinnern, daß die Abbildung $A: \mathbf{x} \mapsto \mathbf{x}'$ die Drehung um den Vektor \mathbf{n} darstellt, (deren Matrix uns partout nicht einfallen wollte), dann zeigt obige Gleichung, daß wir A inzwischen kennen:

$$A = BA_B B^{-1}. \quad (6.41)$$

Sie ist auch leicht zu interpretieren: Mit B^{-1} wechselt man in die Basis B . Dort wird mit der Matrix A_B abgebildet und das Ergebnis anschließend mit B zurück in die Ausgangsbasis transformiert. In dieser Interpretation ist der Bezug zu unserem konkreten Beispiel in den Hintergrund getreten. Wir haben den allgemeinen Mechanismus einer Koordinatentransformation gefunden: *Man sucht eine geeignete Basis B (die durchaus nicht immer eine ONB sein muß), in der man die Matrix A_B einer Abbildung leicht angeben kann. Durch (6.41) berechnet man die Matrix A der Abbildung im Ausgangssystem.*

Das läßt sich übersichtlich durch ein sogenanntes *kommutierendes Diagramm* veranschaulichen:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^n \\ B^{-1} \downarrow & & \uparrow B \\ \mathbb{R}^n & \xrightarrow{A_B} & \mathbb{R}^n \end{array}$$

Statt den direkten Weg $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ zu gehen, beschreitet man den Umweg $B^{-1}A_B B$ über eine Basisdarstellung des \mathbb{R}^n , in der die Abbildungsmatrix A_B bekannt ist. Das läßt sich leicht auf allgemeine Vektorräume und auch auf lineare Abbildungen zwischen verschiedenen Räumen verallgemeinern:

$$\begin{array}{ccc} V & \xrightarrow{A} & W \\ B \downarrow & & \uparrow C \\ X & \xrightarrow{\tilde{A}} & Y \end{array}$$

Durch eine invertierbare lineare Abbildung B wird der Vektorraum V auf den Vektorraum X abgebildet. Von dort führt die bekannte lineare Abbildung \tilde{A} in einen passenden Vektorraum

Y , der mit einer invertierbaren Abbildung C in den Zielraum W transformiert wird. Der allgemeinste Vorgang, den wir als Koordinatentransformation deuten wollen, hat daher die Form

$$\mathcal{A} = C\tilde{A}B.$$

Für unser Beispiel bedeutet das

$$\begin{aligned} A &= \frac{1}{9} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 2 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix} \\ &= \frac{1}{9} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 2 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 2\cos(\alpha) + 2\sin(\alpha) & \cos(\alpha) - 2\sin(\alpha) & -2\cos(\alpha) + \sin(\alpha) \\ -2\cos(\alpha) + 2\sin(\alpha) & 2\cos(\alpha) + \sin(\alpha) & -\cos(\alpha) - 2\sin(\alpha) \end{bmatrix} \\ &= \frac{1}{9} \begin{bmatrix} 1 + 8\cos(\alpha) & 2 - 2\cos(\alpha) - 6\sin(\alpha) & 2 - 2\cos(\alpha) + 6\sin(\alpha) \\ 2 - 2\cos(\alpha) + 6\sin(\alpha) & 4 + 5\cos(\alpha) & 4 - 4\cos(\alpha) - 3\sin(\alpha) \\ 2 - 2\cos(\alpha) - 6\sin(\alpha) & 4 - 4\cos(\alpha) + 3\sin(\alpha) & 4 + 5\cos(\alpha) \end{bmatrix}. \end{aligned}$$

Für $\alpha = \frac{\pi}{2}$ ergibt sich wieder

$$A\mathbf{x} = \frac{1}{9} \begin{bmatrix} 1 & -4 & 8 \\ 8 & 4 & 1 \\ -4 & 7 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 \\ 7 \\ 7 \end{bmatrix}.$$

Natürlich läßt sich (6.41) auch nach $A_{\mathcal{B}}$ umstellen. Wir haben dann die Transformationsformel einer Abbildung A in die Basis \mathcal{B} :

$$A_{\mathcal{B}} = B^{-1}AB. \quad (6.42)$$

6.12 Determinanten

Wir haben im Abschnitt 5.1.13 das Spatprodukt $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ dreier Vektoren \mathbf{x} , \mathbf{y} und \mathbf{z} im \mathbb{R}^3 als Volumenfunktion kennengelernt. Es bestimmt, bis auf ein mögliches Vorzeichen, das Volumen des Spats, der von den drei Vektoren aufgespannt wird. Wir haben das Produkt mit dem systematischen Namen Determinante bezeichnet, der für die Volumenfunktion in höheren Dimensionen verwendet wird, mit der wir uns jetzt beschäftigen werden. Um die richtige Definition zu finden, rekapitulieren wir noch einmal die Eigenschaften, die die Determinante im \mathbb{R}^3 charakterisiert (vergl. Satz 5.1.14):

Die Determinante ist eine Abbildung von $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ nach \mathbb{R} , mit den Eigenschaften:

- i) $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \mapsto \det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ ist in jeder Komponente linear. D. h., für $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ und $t_1, t_2 \in \mathbb{R}$ gilt

$$\det(t_1 \mathbf{x}_1 + t_2 \mathbf{x}_2, \mathbf{y}, \mathbf{z}) = t_1 \det(\mathbf{x}_1, \mathbf{y}, \mathbf{z}) + t_2 \det(\mathbf{x}_2, \mathbf{y}, \mathbf{z}).$$

Entsprechendes gilt für die zweite und die dritte Komponente.

- ii) Die Vertauschung zweier Komponenten ergibt einen Vorzeichenwechsel der Determinante. Also etwa $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = -\det(\mathbf{y}, \mathbf{x}, \mathbf{z}) = \det(\mathbf{y}, \mathbf{z}, \mathbf{x}) \dots$

- iii) Für die kanonische Basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ des \mathbb{R}^3 gilt $\det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1$.

Genau genommen würde es ausreichen, unter i) die Linearität in der ersten Komponente zu fordern. Zusammen mit ii) läßt sie sich dann auf alle anderen ausdehnen.

Eine Abbildung mit der Eigenschaft i) heißt *multilinear*, mit der Eigenschaft ii) *alternierend* und *normiert*, wenn iii) erfüllt ist. Kurz gesagt ist eine Determinante daher eine *normierte, alternierende Multilinearform*. In dieser Formulierung ist der zugrunde liegende Raum \mathbb{R}^3 schon nicht mehr erkennbar. Deshalb werden wir ein solches Objekt im \mathbb{R}^n oder \mathbb{C}^n als Determinante bezeichnen und seine Eigenschaften in diesem allgemeinen Rahmen untersuchen. Dabei stellt sich zunächst die Frage nach der Realisierbarkeit, denn in diesen Räumen haben wir keine Hilfsmittel, wie etwa das Kreuzprodukt zu unserer Verfügung. Erfreulicherweise wird die Analyse der Eigenschaften einer Determinante in eine konkrete Rechenvorschrift münden und so die Frage nach der Existenz zur Zufriedenheit beantworten.

6.12.1 Definition Eine alternierende Multilinearform λ auf \mathbb{K}^n ($\mathbb{K} = \mathbb{R}$, oder $\mathbb{K} = \mathbb{C}$) ist eine Abbildung

$$\mathbb{K}^n \times \mathbb{K}^n \times \dots \times \mathbb{K}^n \ni (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \mapsto \lambda(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \in \mathbb{K}$$

mit folgenden Eigenschaften:

- i) Sie ist in jeder Komponente linear, d. h., für jedes $i = 1, \dots, n$ ist

$$\mathbb{K}^n \ni \mathbf{a}_i \mapsto \lambda(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)$$

eine lineare Abbildung von \mathbb{K}^n nach \mathbb{K} .

- ii) Sie ist alternierend, d. h., für zwei beliebige Positionen $i \neq j$ gilt

$$\lambda(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) = -\lambda(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n).$$

Ist λ darüber hinaus normiert, d. h. gilt für die kanonische Basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ von \mathbb{K}^n $\lambda(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$, so nennen wir diese Form Determinante \det auf \mathbb{K}^n . Wir verwenden den Begriff Determinante auch für Matrizen $A \in M_n(\mathbb{K})$, indem wir für eine solche Matrix $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ mit den Spaltenvektoren $\mathbf{a}_i \in \mathbb{K}^n$ die Determinante durch $\det(A) := \det(\mathbf{a}_1, \dots, \mathbf{a}_n)$ definieren.

In dieser Definition steckt auch schon eine Behauptung, weil wir von der Determinante auf \mathbb{K}^n sprechen. Wir werden gleich sehen, daß es tatsächlich nur eine Determinante auf \mathbb{K}^n gibt. Wir ziehen zunächst einige Folgerungen aus der Definition:

6.12.2 Satz Alternierende Multilinearformen $\lambda \neq 0$ auf \mathbb{K}^n haben folgende Eigenschaften:

- i) Ist einer der Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ der Nullvektor, oder sind zwei der Vektoren gleich, so gilt $\lambda(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$.
- ii) $\lambda(\mathbf{a}_1, \dots, t\mathbf{a}_j + \mathbf{a}_i, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) = \lambda(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n)$, $t \in \mathbb{K}$, d. h., man kann beliebige Vielfache eines Vektors zu einem anderen Vektor addieren, ohne daß sich der Wert der Form ändert.
- iii) Ist $\lambda(\mathbf{a}_1, \dots, \mathbf{a}_n) \neq 0$, so ist die Menge $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ linear unabhängig.
- iv) Zwei alternierende Multilinearformen λ und λ' auf \mathbb{K}^n unterscheiden sich nur durch einen gemeinsamen Faktor: $\lambda' = c \cdot \lambda$, $c \in \mathbb{K}$.
- v) Es gibt nur eine normierte alternierende Multilinearform auf \mathbb{K} , nämlich die Determinante \det . Insbesondere unterscheiden sich alle anderen alternierenden Multilinearformen von \det jeweils nur um einen festen Vorfaktor.

Beweis. Zu i): Wir können o. B. d. A. von $\mathbf{a}_1 = \mathbf{0}$ ausgehen. Die Linearität in der ersten Komponente zeigt dann $\det(\mathbf{0}, \mathbf{a}_2, \dots, \mathbf{a}_n) = \det(0 \cdot \mathbf{0}, \mathbf{a}_2, \dots, \mathbf{a}_n) = 0 \cdot \det(\mathbf{0}, \mathbf{a}_2, \dots, \mathbf{a}_n) = 0$. Gilt $\mathbf{a}_i = \mathbf{a}_j$ für $i \neq j$, dann ergibt die Vertauschung dieser Vektoren in der Determinante keine Veränderung, muß aber nach 6.12.1 ii) einen Vorzeichenwechsel verursachen. Das heißt: $\lambda(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n) = -\lambda(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n) = 0$.

Zu ii): Das folgt aus i) und der Multilinearität:

$$\lambda(\dots, t\mathbf{a}_j + \mathbf{a}_i, \dots, \mathbf{a}_j, \dots) = t\lambda(\dots, \mathbf{a}_j, \dots, \mathbf{a}_j, \dots) + \lambda(\dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots).$$

Nach i) verschwindet $t\lambda(\dots, \mathbf{a}_j, \dots, \mathbf{a}_j, \dots)$.

Diese Eigenschaft erinnert an die Umformungen beim GAUSS-Verfahren. Tatsächlich ist sie der Grund dafür, daß große Determinanten mit einer geringfügigen Variation des GAUSS-Verfahrens berechnet werden können.

Zu iii): Wir zeigen: $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ ist linear abhängig $\Rightarrow \lambda(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$.

Nach Lemma 6.4.7 ist $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ genau dann linear abhängig, wenn es einen Vektor aus dieser Menge gibt, o. B. d. A. dürfen wir von \mathbf{a}_1 ausgehen, der eine Linearkombination der übrigen ist: $\mathbf{a}_1 = t_2\mathbf{a}_2 + t_3\mathbf{a}_3 + \dots + t_n\mathbf{a}_n$. Die Linearität in der ersten Komponente und i) ergeben dann

$$\begin{aligned} \lambda(\mathbf{a}_1, \dots, \mathbf{a}_n) &= \lambda(t_2\mathbf{a}_2 + t_3\mathbf{a}_3 + \dots + t_n\mathbf{a}_n, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n) \\ &= t_2\lambda(\mathbf{a}_2, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n) + t_3\lambda(\mathbf{a}_3, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n) + \dots \end{aligned}$$

$$+ t_n \lambda(\mathbf{a}_n, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n) = 0.$$

Daher gilt: $\lambda(\mathbf{a}_1, \dots, \mathbf{a}_n) \neq 0 \Rightarrow \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ ist linear unabhängig (vergl. 1.9).

Zu iv): Im Laufe der Untersuchungen zu diesem Punkt werden wir eine konkrete Rechenvorschrift zur Berechnung von $\lambda(\mathbf{a}_1, \dots, \mathbf{a}_n)$ erarbeiten.

Zunächst untersuchen wir die Multilinearität von λ . Dazu entwickeln wir alle Vektoren \mathbf{a}_i in der kanonischen Basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$:

$$\mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} = a_{1i} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_{2i} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_{ni} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \sum_{k=1}^n a_{ki} \mathbf{e}_k.$$

Dann verwenden wir nacheinander die Linearität in jeder Komponente von λ :

$$\begin{aligned} \lambda(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) &= \lambda\left(\sum_{k_1=1}^n a_{k_1 1} \mathbf{e}_{k_1}, \mathbf{a}_2, \dots, \mathbf{a}_n\right) \\ &= \sum_{k_1=1}^n a_{k_1 1} \cdot \lambda\left(\mathbf{e}_{k_1}, \sum_{k_2=1}^n a_{k_2 2} \mathbf{e}_{k_2}, \dots, \mathbf{a}_n\right) \\ &= \sum_{k_1=1}^n \sum_{k_2=1}^n a_{k_1 1} a_{k_2 2} \cdot \lambda\left(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{a}_n\right) \\ &= \sum_{k_1=1}^n \sum_{k_2=1}^n \dots \sum_{k_n=1}^n a_{k_1 1} a_{k_2 2} \dots a_{k_n n} \cdot \lambda(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}). \end{aligned}$$

Dieser Ausdruck stellt bereits eine Rechenvorschrift dar, wenn wir die Ausdrücke $\lambda(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n})$ berechnen können. Alles andere hängt nur noch von den Koordinaten der Vektoren \mathbf{a}_i ab. Aus i) wissen wir, daß λ den Wert Null liefert, wenn zwei Vektoren \mathbf{e}_{k_i} und \mathbf{e}_{k_j} gleich sind. Das bedeutet, daß in der Vielfachsumme alle Summanden wegfallen, für die wenigstens zwei der Indizes k_1, k_2, \dots, k_n gleich sind. Oder anders ausgedrückt: Die Summe erstreckt sich nur über Indextupel $[k_1, k_2, \dots, k_n]$, in denen keine zwei Einträge k_i und k_j gleich sind. Diese Tupel stellen daher *Permutationen* π von $[1, 2, \dots, n]$ dar. Wir haben ihre Eigenschaften in Abschnitt 2.5.10 eingehend untersucht und verwenden jetzt die dort vereinbarten Notationen und die erzielten Ergebnisse – allerdings benötigen wir letztlich nur eine einzige Tatsache (s.u.). Die Vielfachsumme erstreckt sich also tatsächlich über alle Elemente π aus der Menge \mathcal{S}_n der Permutationen der Zahlen $1, \dots, n$. k_1, k_2, \dots sind dann die Bilder von $1, 2, \dots$ der Abbildung $\pi: k_i = \pi(i)$. Wir erhalten

$$\lambda(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \sum_{\pi \in \mathcal{S}_n} a_{\pi(1)1} a_{\pi(2)2} \dots a_{\pi(n)n} \cdot \lambda(\mathbf{e}_{\pi(1)}, \mathbf{e}_{\pi(2)}, \dots, \mathbf{e}_{\pi(n)}).$$

Der Ausdruck $\lambda(\mathbf{e}_{\pi(1)}, \mathbf{e}_{\pi(2)}, \dots, \mathbf{e}_{\pi(n)})$ unterscheidet sich von $\lambda(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ nur um ein Vorzeichen. Durch eine Folge von Transpositionen können die Vektoren

$\mathbf{e}_{\pi(1)}, \mathbf{e}_{\pi(2)}, \dots, \mathbf{e}_{\pi(n)}$ nämlich in die gewohnte Reihenfolge $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ gebracht werden. Deren Abfolge ist zwar nicht eindeutig, aber zwei verschiedene unterscheiden sich immer um eine gerade Anzahl an Transpositionen (mehr wird aus dem Abschnitt 2.5.10 eigentlich nicht gebraucht). Jede der k benötigten Transposition erzeugt nach 6.12.1 ii) ein Vorzeichenwechsel. Damit ist $(-1)^k$ der Faktor, der $\lambda(\mathbf{e}_{\pi(1)}, \mathbf{e}_{\pi(2)}, \dots, \mathbf{e}_{\pi(n)})$ von $\lambda(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ unterscheidet. Diesen Faktor bezeichnet man als Vorzeichen $\text{sgn}(\pi)$ von π (Definition 2.5.12).

Als Ergebnis unserer Untersuchung haben wir

$$\lambda(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \lambda(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \cdot \sum_{\pi \in \mathcal{S}_n} \text{sgn}(\pi) \cdot a_{\pi(1)1} a_{\pi(2)2} \dots a_{\pi(n)n}. \quad (6.43)$$

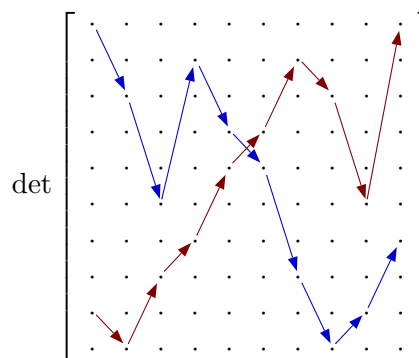
Die Summe ist eine konkrete Rechenvorschrift, die nur von den Koeffizienten der Vektoren \mathbf{a}_i abhängt, aber nicht mehr von λ . Zwei alternierende Multilinearformen $\lambda \neq 0$ und $\lambda' \neq 0$ unterscheiden sich also nur um einen gemeinsamen Faktor voneinander (nämlich um $\lambda(\mathbf{e}_1, \dots, \mathbf{e}_n)/\lambda'(\mathbf{e}_1, \dots, \mathbf{e}_n)$, wenn λ' nicht die triviale Form 0 ist, was von $\lambda'(\mathbf{e}_1, \dots, \mathbf{e}_n) \neq 0$ bereits garantiert wird).

Zu v): Gleichung 6.43 zeigt auch, daß es nur eine alternierende Multilinearform geben kann, die normiert ist. Diese ist nämlich durch die Summe in 6.43 bestimmt, in die nur die Koordinatendarstellungen der Vektoren \mathbf{a}_i eingehen. Der Vorfaktor $\lambda(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ ist wegen der Normierung 1. Gemäß Definition 6.12.1 gilt

$$\det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \sum_{\pi \in \mathcal{S}_n} \text{sgn}(\pi) \cdot a_{\pi(1)1} a_{\pi(2)2} \dots a_{\pi(n)n}. \quad (6.44)$$

Das ist die *Summendarstellung der Determinante*. □

Diese Darstellung ist für konkrete Rechnungen nicht sonderlich geeignet. Aber sie wird uns dabei helfen bessere Methoden zu finden. Obwohl sie so unhandlich ist, hat sie doch eine sehr anschauliche Deutung, die den ersten Schritt hin zu bequemeren Berechnungsmöglichkeiten darstellt. Es ist das sogenannte *Pfadbild* der Determinante. Dafür sehen wir für den Moment mal von den konkreten Einträgen der Matrix $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ ab und stellen sie uns als ein quadratisches Punkteschema vor:



Jeder Summand in (6.44) entspricht dann einem Pfad durch das Punktgitter, der jede Zeile und jede Spalte jeweils genau einmal trifft. Diese Eigenschaft einer Determinante bezeichnen

wir als *Pfadregel*.. Zwei Beispiele sind eingezeichnet, nämlich der Pfad, der zum Summanden $a_{11}a_{32}a_{63}a_{24}a_{45}a_{56}a_{87}a_{108}a_{99}a_{710}$ und der, der zu $-a_{91}a_{102}a_{83}a_{74}a_{55}a_{46}a_{27}a_{38}a_{69}a_{110}$ gehört.

Dieses Bild ist stark genug, um sofort einige praktische Folgerungen ziehen zu können, wenn die Matrix eine spezielle Form hat. Ist A etwa eine obere (oder untere) Dreiecksmatrix

$$\det \begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & a_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & a_{33} & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & a_{44} & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot \\ & 0 & & & & \cdot & \cdot & \cdot \\ & & & & & & \cdot & \cdot \\ & & & & & & & a_{nn} \end{bmatrix} = a_{11}a_{22}a_{33}a_{44} \cdots a_{nn},$$

dann ist die Determinante einfach das Produkt der Diagonalelemente, denn der einzige Pfad, der nicht wenigstens einmal in die Dreieckshälfte, die nur aus Nullen besteht, ausweichen muß, ist der entlang der Diagonalen. Die Permutation, die zu diesem Pfad gehört, ist natürlich id, für die $\text{sgn}(\text{id}) = 1$ gilt.

Diese Situation läßt sich noch verallgemeinern. Dafür benötigen wir aber eine wichtige Folgerung aus dem Eindeutigkeitsatz 6.12.2 iv).

6.12.3 Der Determinanten-Produktsatz

6.12.4 Satz Für zwei Matrizen $A, B \in M_n(\mathbb{K})$ gilt

$$\det(AB) = \det(A) \det(B). \quad (6.45)$$

Das bedeutet insbesondere $\det(AB) = \det(BA)$, obwohl im Allgemeinen $AB \neq BA$ gilt.

Beweis. Die Matrix B sei durch ihre Spaltenvektoren gegeben: $B := [\mathbf{b}_1, \dots, \mathbf{b}_n]$. Wir definieren die Abbildung $(\mathbf{b}_1, \dots, \mathbf{b}_n) \mapsto \lambda(\mathbf{b}_1, \dots, \mathbf{b}_n) := \det(AB) = \det(A[\mathbf{b}_1, \dots, \mathbf{b}_n]) = \det([A\mathbf{b}_1, \dots, A\mathbf{b}_n]) = \det(A\mathbf{b}_1, \dots, A\mathbf{b}_n)$. Sie ist multilinear, denn A ist eine lineare Abbildung:

$$\begin{aligned} \lambda(\mathbf{b}_1 + t \cdot \mathbf{c}_1, \dots, \mathbf{b}_n) &= \det(A\mathbf{b}_1 + t \cdot A\mathbf{c}_1, \dots, A\mathbf{b}_n) \\ &= \det(A\mathbf{b}_1, \dots, A\mathbf{b}_n) + t \cdot \det(A\mathbf{c}_1, \dots, A\mathbf{b}_n) \\ &= \lambda(\mathbf{b}_1, \dots, \mathbf{b}_n) + t \cdot \lambda(\mathbf{c}_1, \dots, \mathbf{b}_n). \end{aligned}$$

Das zeigt die Linearität in der ersten Komponente. λ ist auch alternierend, denn das erbt die Abbildung von \det . Daher ist λ eine alternierende Multilinearform. Sie unterscheidet sich von jeder anderen, insbesondere von \det , nur jeweils um eine Konstante. Das heißt, es gilt $\det(AB) = \lambda(\mathbf{b}_1, \dots, \mathbf{b}_n) = c \cdot \det(\mathbf{b}_1, \dots, \mathbf{b}_n) = c \cdot \det(B)$, mit einer Konstante c , die wir gleich bestimmen werden. Die Gleichung gilt für jede Matrix $B \in M_n(\mathbb{K})$ und insbesondere auch für $B = \mathbf{1}$: $\det(A) = \det(A\mathbf{1}) = c \cdot \det(\mathbf{1}) = c$. Damit haben wir $\det(AB) = \det(A) \det(B)$ gezeigt. \square

Der Determinanten-Produktsatz hat viele Anwendungen, von denen wir einige kennenlernen werden. Etwa für Matrizen der Form

$$\left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1k} & c_{11} & c_{12} & \cdots & c_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & c_{k1} & c_{k2} & \cdots & c_{kn} \\ \hline 0 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & 0 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{array} \right] = \left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array} \right].$$

Wir haben die $k \times k$ -Matrix in der linken oberen Ecke mit A , die anschließende $k \times n$ -Matrix mit C , die $n \times n$ -Matrix darunter mit B und die links anschließende Nullmatrix mit 0 bezeichnet. Die Matrix

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$$

ist eine *Blockmatrix*, da ihre Einträge selber Matrizen passender Abmessungen sind. Die Pfade, die einen Beitrag zur Determinante dieser Blockmatrix leisten, müssen sich die ersten k Schritte in der Matrix A aufhalten, da sie sonst unweigerlich in die Matrix 0 eintreten müßten. Dabei werden natürlich auch die ersten k Zeilen alle besucht. Daher kann keiner dieser Pfade in die Matrix C gelangen, denn dafür müßte eine der ersten k Zeilen ein weiteres Mal betreten werden, was nach der Pfadregel verboten ist. Die einzig mögliche Fortsetzung findet er also in der Matrix B . Da keiner dieser Pfade Einträge aus C aufweist, können wir C durch 0 ersetzen, ohne den Wert der Determinante zu ändern. Mit dem Determinanten-Produktsatz folgt

$$\begin{aligned} \det \left(\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \right) &= \det \left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right) = \det \left(\begin{bmatrix} A & 0 \\ 0 & \mathbb{1}_n \end{bmatrix} \begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & B \end{bmatrix} \right) \\ &= \det \left(\begin{bmatrix} A & 0 \\ 0 & \mathbb{1}_n \end{bmatrix} \right) \det \left(\begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & B \end{bmatrix} \right) = \det(A) \det(B). \end{aligned}$$

Die letzte Gleichheit bedarf noch einer Begründung. Wir zeigen nur, daß

$$\det \left(\begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & B \end{bmatrix} \right) = \det(B)$$

gilt. Dafür definieren wir die alternierende Multilinearform λ auf \mathbb{K}^n durch

$$\lambda(\mathbf{b}_1, \dots, \mathbf{b}_n) := \det \left(\begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & [\mathbf{b}_1, \dots, \mathbf{b}_n] \end{bmatrix} \right).$$

Die Eigenschaften einer alternierenden Multilinearform sind für λ schnell überprüft: Sie erbt sie alle von der Determinante auf \mathbb{K}^{k+n} . Wenn wir zeigen können, daß λ normiert ist, handelt es sich um die Determinante auf \mathbb{K}^n . Das ist einfach:

$$\lambda(\mathbf{e}_1, \dots, \mathbf{e}_n) = \det \left(\begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & [\mathbf{e}_1, \dots, \mathbf{e}_n] \end{bmatrix} \right) = \det \left(\begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & \mathbb{1}_n \end{bmatrix} \right) = 1.$$

Also ist $\lambda(\mathbf{b}_1, \dots, \mathbf{b}_n) = \det(B)$.

Halten wir das Ergebnis fest: Für eine obere (untere) Dreiecks-Blockmatrix gilt

$$\det \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{bmatrix} = \det(A_{11}) \det(A_{22}) \cdots \det(A_{nn}). \quad (6.46)$$

Eigentlich haben wir das nur für Blockmatrizen mit zwei Diagonaleinträgen gezeigt, aber das reicht, denn wir können dieses Ergebnis iterieren:

$$\det \left[\begin{array}{c|ccc} A_{11} & A_{12} & \cdots & A_{1n} \\ \hline 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{array} \right] = \det(A_{11}) \cdot \det \left[\begin{array}{c|ccc} A_{22} & \cdots & A_{2n} \\ \hline \vdots & \ddots & \vdots \\ 0 & \cdots & A_{nn} \end{array} \right].$$

6.12.5 Beispiel

$$\begin{aligned} \det \begin{bmatrix} 1 & 30 & 0 & 6 & 19 & 10 & 8 & i & 5 & -3 & 5 \\ 1 & 1 & 10 & 6 & 19 & 0 & 7 & 7i & 55 & -3 & 0 \\ 0 & 0 & 0 & 3 & 6 & 8 & 18 & 2i & 5 & 13 & 1 \\ 0 & 0 & 1 & 4 & 8 & 1 & 0 & 0 & 22 & 0 & 81 \\ 0 & 0 & 3 & 14 & 0 & 0 & 3 & 20 & 17 & 9 & i \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 5 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & 7 & 2 & 4 & 10 & 14 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 2 & 2 & 29 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 & 12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5 \end{bmatrix} \\ = \det \begin{bmatrix} 1 & 30 \\ 1 & 1 \end{bmatrix} \cdot \det \begin{bmatrix} 0 & 3 & 6 \\ 1 & 4 & 8 \\ 3 & 14 & 0 \end{bmatrix} \cdot \det \begin{bmatrix} 2 & 1 \\ 7 & 2 \end{bmatrix} \cdot 4 \cdot 5 \cdot \det \begin{bmatrix} 2 & 5 \\ 1 & 5 \end{bmatrix} \\ = -29 \cdot (3 \cdot 8 \cdot 3 + 6 \cdot 14 - 6 \cdot 4 \cdot 3) \cdot (-3) \cdot 20 \cdot 5 = 730800. \end{aligned}$$

Eine der wichtigsten Folgerungen aus dem Determinanten-Produktsatz betrifft die Charakterisierung der Invertierbarkeit einer Matrix. Falls A^{-1} für eine Matrix A existiert, kann $\det(A)$ nicht Null sein, denn $1 = \det(\mathbb{1}) = \det(AA^{-1}) = \det(A) \det(A^{-1})$. Nebenbei erhalten wir daraus $\det(A^{-1}) = \frac{1}{\det(A)}$.

Erfreulicherweise gilt auch die Umkehrung: Aus $\det(A) \neq 0$ folgt die Existenz von A^{-1} . Denn nach Satz 6.12.2 iii) sind die Spaltenvektoren \mathbf{a}_i von $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ linear unabhängig. Daher muß $\ker A = \{\mathbf{0}\}$ gelten, denn ein Vektor $\mathbf{0} \neq \mathbf{x} = [x_1, \dots, x_n] \in \ker A$ würde wegen $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$ eine nicht triviale Nullkombination der Spaltenvektoren ergeben, im Widerspruch zu deren linearen Unabhängigkeit. Nach Korollar 6.8.5 ist A invertierbar. Wir haben somit gezeigt:

6.12.6 Satz Für eine Matrix $A \in M_n(\mathbb{K})$ gilt:

$$A^{-1} \text{ existiert} \Leftrightarrow \det(A) \neq 0.$$

Wie verhält sich die Determinante einer Matrix unter äquivalenten Zeilenumformungen, wie sie für das GAUSS-Verfahren benötigt werden? Zur Beantwortung dieser Frage führen wir das GAUSS-Verfahren auf die Multiplikation von A mit geeigneten Matrizen D_{ij} und T_{ij} zurück und verwenden dann den Determinanten-Produktsatz. Zwei einfache Rechnungen zeigen die Idee:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \lambda \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} + \lambda a_{41} & a_{22} + \lambda a_{42} & a_{23} + \lambda a_{43} & a_{24} + \lambda a_{44} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix}.$$

Die Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \lambda \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4 + \lambda \mathbf{e}_2] =: D_{24}(\lambda)$$

bewirkt die Addition des λ -Fachen der vierten Zeile zur zweiten und die Matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = [\mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_3, \mathbf{e}_2] =: T_{24}$$

die Vertauschung der zweiten mit der vierten Zeile. Die Matrizen

$$D_{ij}(\lambda) := [\mathbf{e}_1, \dots, \mathbf{e}_i, \dots, \mathbf{e}_j + \lambda \mathbf{e}_i, \dots, \mathbf{e}_n] \quad (6.47)$$

$$T_{ij} := [\mathbf{e}_1, \dots, \mathbf{e}_j, \mathbf{e}_{i+1}, \dots, \mathbf{e}_{j-1}, \mathbf{e}_i, \dots, \mathbf{e}_n] \quad (6.48)$$

werden also das λ -Fache der j -ten Zeile zur i -ten Zeile addieren, bzw. die i -te mit der j -ten Zeile vertauschen. Nach der Pfadregel gilt $\det(D_{ij}(\lambda)) = 1$, denn es handelt sich um eine obere ($j > i$), oder um eine untere Dreiecksmatrix ($j < i$), mit einer Diagonalen, die nur aus Einsen besteht. Dagegen ist $\det(T_{ij}) = -1$, denn die Vertauschung der i -ten mit der j -ten Spalte ergibt die Einheitsmatrix. Die Anwendung des Determinanten-Produktsatzes,

$$\det(D_{ij}(\lambda)A) = \det(D_{ij}(\lambda)) \det(A) = \det(A),$$

$$\det(T_{ij}A) = \det(T_{ij}) \det(A) = -1 \det(A),$$

zeigt, daß sich $\det(A)$, was die Zeilenvektoren angeht, genauso verhält, wie für die Spaltenvektoren. Bis auf ein mögliches Vorzeichen ändert sich $\det(A)$ nicht, wenn wir in der Matrix A äquivalenten Zeilenumformungen (repräsentiert durch $D_{ij}(\lambda)$ und T_{ij}) und Vertauschungen von Spaltenvektoren vornehmen, bis A in eine obere bzw. unter Dreiecksmatrix D verwandelt wurde. Bis auf besagtes Vorzeichen stimmt $\det(A)$ mit $\det(D)$ überein.

Jetzt untersuchen wir die Situation für $\det(A^t)$. Führen wir hier, in genau derselben Reihenfolge wie für $\det(A)$, die entsprechenden äquivalenten Spaltenumformungen und Vertauschungen von Zeilen durch, so erhalten wir, bis auf dasselbe Vorzeichen, die Determinante von D^t . Wegen $\det(D) = \det(D^t)$ folgt $\det(A) = \det(A^t)$.

6.12.7 Satz Für jede Matrix $A \in M_n(\mathbb{K})$ gilt $\det(A) = \det(A^t)$.

Wir haben diesen Satz durch unsere Vorüberlegungen bereits bewiesen. Meist findet man für ihn aber einen anderen Beweis, der auf die Summendarstellung 6.44 zurückgreift. Es kann nicht schaden, für einen solchen Satz einen weiteren Beweis zu haben.

Beweis.

$$\begin{aligned} \det(A) &= \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi) \cdot a_{\pi(1)1} a_{\pi(2)2} \cdots a_{\pi(n)n} \\ &= \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi) \cdot a_{\pi(1)\pi^{-1}(\pi(1))} a_{\pi(2)\pi^{-1}(\pi(2))} \cdots a_{\pi(n)\pi^{-1}(\pi(n))} \\ &\stackrel{i)}{=} \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi^{-1}) \cdot a_{1\pi^{-1}(1)} a_{2\pi^{-1}(2)} \cdots a_{n\pi^{-1}(n)} \\ &\stackrel{ii)}{=} \sum_{\pi \in \mathcal{S}_n} \operatorname{sgn}(\pi) \cdot a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \stackrel{iii)}{=} \det(A^t). \end{aligned}$$

Dazu muß etwas gesagt werden.

Zu i): Im Produkt $a_{\pi(1)\pi^{-1}(\pi(1))} a_{\pi(2)\pi^{-1}(\pi(2))} \cdots a_{\pi(n)\pi^{-1}(\pi(n))}$ durchlaufen die Zahlen $\pi(1), \pi(2), \dots, \pi(n)$ die Zahlen von 1 bis n . Durch Umsortierung läßt es sich also in die Form $a_{1\pi^{-1}(1)} a_{2\pi^{-1}(2)} \cdots a_{n\pi^{-1}(n)}$ bringen. Außerdem gilt $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$, denn $\operatorname{sgn}(\pi) \operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi\pi^{-1}) = 1$ (vergl. (2.16)).

Zu ii): Wenn π alle Permutationen aus \mathcal{S}_n durchläuft, dann stimmt das auch für π^{-1} . Die Summe unter i) ist also nur die umsortierte Summe ii).

Zu iii): Die Summe, die in ii) erhalten wurde, hat wieder die Summenform der Determinante einer Matrix, nur daß gegenüber 6.44 in den Matrixeinträgen die Reihenfolge der Indizes vertauscht ist. Es handelt sich also um die Einträge der Matrix A^t und bei der Determinante daher um $\det(A^t)$. \square

Die Überlegungen, die uns zu Satz 6.12.7 geführt haben, zeigen schon, daß wir zur Berechnung von $\det(A)$ die Matrix A mit Hilfe des GAUSS-Verfahrens und eventuell durch Vertauschen geeigneter Spaltenvektoren in eine obere oder untere Dreiecksmatrix D transformieren können, deren Determinante als Produkt der Diagonalelemente leicht gebildet werden kann. Führt man dabei Buch über die Vorzeichenwechsel, die durch Zeilen- bzw. Spaltenvertauschungen hervorgerufen werden, dann ist $\det(D)$, multipliziert mit dem erhaltenen Vorzeichen gerade

$\det(A)$. Praktischerweise gestattet man es sich dabei auch, wenn es sich anbietet, Spalten oder Zeilen mit geeigneten Faktoren zu multiplizieren, um die Rechnungen zu vereinfachen. Wenn man auch über diese Faktoren Buch führt, kann man sie zuletzt aus dem Ergebnis wieder herausrechnen. Wie das aussehen kann zeigen wir an einem Beispiel.

6.12.8 Beispiel Wir wollen die Determinante folgender Matrix berechnen.

$$A := \begin{bmatrix} 2 & 5 & 0 & 0 & 2 & 1 \\ 1 & 2 & 6 & 0 & 5 & 2 \\ 0 & 5 & 3 & 5 & 4 & 3 \\ 8 & 5 & 3 & 0 & 1 & 2 \\ 3 & 1 & 1 & 0 & 0 & 3 \\ 0 & 2 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

	i	ii	iii	iv	v	vi				i	ii	iii	iv	v	vi			
I	2	5	0	0	2	1				I	1	2	0	0	1	0		
II	1	2	6	0	5	2				II	0	1	1	3	0	3		
III	0	5	3	5	4	3	III - 5 VI			III	0	0	8	15	-1	18	III - 2 V	
IV	8	5	3	0	1	2				IV	0	0	6	8	0	5	IV - 2 VI	
V	3	1	1	0	0	3				V	0	0	4	-5	5	-4		
VI	0	2	0	1	1	0				VI	0	0	3	2	1	4		
I	2	5	0	0	2	1	I ↔ VI			I	1	2	0	0	1	0		
II	1	2	6	0	5	2				II	0	1	1	3	0	3		
III	0	-5	3	0	-1	3				III	0	0	0	25-11	26			
IV	8	5	3	0	1	2				IV	0	0	0	4	-2	-3		
V	3	1	1	0	0	3				V	0	0	4	-5	5	-4	3 V - 4 VI	
VI	0	2	0	1	1	0				VI	0	0	3	2	1	4		
I	0	2	0	1	1	0	i ↔ iv	·(-1)		I	1	2	0	0	1	0		
II	1	2	6	0	5	2				II	0	1	1	3	0	3		
III	0	-5	3	0	-1	3				III	0	0	0	25-11	26		III ↔ VI	
IV	8	5	3	0	1	2				IV	0	0	0	4	-2	-3		
V	3	1	1	0	0	3				V	0	0	0	-23	11	-28		· 1/3
VI	2	5	0	0	2	1				VI	0	0	3	2	1	4		
I	1	2	0	0	1	0				I	1	2	0	0	1	0		
II	0	2	6	1	5	2	II ↔ V	·(-1)		II	0	1	1	3	0	3		
III	0	-5	3	0	-1	3				III	0	0	3	2	1	4		
IV	0	5	3	8	1	2	IV + III			IV	0	0	0	4	-2	-3		
V	0	1	1	3	0	3				V	0	0	0	-23	11	-28	V + VI	·(-1)
VI	0	5	0	2	2	1	VI + III			VI	0	0	0	25-11	26			
I	1	2	0	0	1	0				I	1	2	0	0	1	0		
II	0	1	1	3	0	3				II	0	1	1	3	0	3		
III	0	-5	3	0	-1	3	III + 5 II	·(-1)		III	0	0	3	2	1	4		
IV	0	0	6	8	0	5				IV	0	0	0	4	-2	-3		
V	0	2	6	1	5	2	V - 2 II			V	0	0	0	2	0	-2		
VI	0	0	3	2	1	4				VI	0	0	0	25-11	26			

An dieser Stelle brechen wir das GAUSS-Verfahren ab. Wie man sieht, haben wir eine weitere Spalte dem Schema angefügt, in der wir über Korrekturfaktoren Buch führen, die durch Vertauschungen oder Multiplikation einzelner Zeilen mit geeigneten Zahlen entstanden sind. Da unsere Matrix weitgehend dreieckig geworden ist, erledigen wir den Rest mit Hilfe von 6.46 und erhalten

$$\det(A) = (-1)^4 \cdot \frac{1}{3} \cdot 3 \det \begin{bmatrix} 4 & -2 & -3 \\ 2 & 0 & -2 \\ 25 & -11 & 26 \end{bmatrix} = \det \begin{bmatrix} 4 & -2 & -3 \\ 2 & 0 & -2 \\ 1 & 1 & 44 \end{bmatrix} = -\det \begin{bmatrix} -2 & 4 & -3 \\ 0 & 2 & -2 \\ 1 & 1 & 44 \end{bmatrix}$$

$$= -\frac{1}{2} \det \begin{bmatrix} -2 & 4 & -3 \\ 0 & 2 & -2 \\ 0 & 6 & 85 \end{bmatrix} = -\frac{1}{2} \det \begin{bmatrix} -2 & 4 & -3 \\ 0 & 2 & -2 \\ 0 & 0 & 91 \end{bmatrix} = 182,$$

oder kürzer, unter Verwendung der Merkregel 5.32:

$$\det(A) = \det \begin{bmatrix} 4 & -2 & -3 \\ 2 & 0 & -2 \\ 25 & -11 & 26 \end{bmatrix} = 4 \cdot 25 + 6 \cdot 11 - 8 \cdot 11 + 4 \cdot 26 = 182.$$

6.12.9 Der LAPLACESche Entwicklungssatz Neben dem GAUSS-Verfahren gibt es eine rekursive Methode, um Determinanten zu berechnen. Dabei wird die Determinante einer $n \times n$ -Matrix auf eine Summe über n Determinanten von $(n-1) \times (n-1)$ -Matrizen zurückgeführt. Der eigentliche Grund für diese Darstellung liegt in der Multilinearität der Determinante, die wir nach Satz 6.12.7 jetzt auch für die Zeilenvektoren zur Verfügung haben. Denken wir uns dazu die Matrix A durch Zeilenvektoren $\mathbf{a}_i^t = [a_{i1}, a_{i2}, \dots, a_{in}] = \sum_{j=1}^n a_{ij} \mathbf{e}_j^t$ gegeben. Benutzen wir die Linearität der Determinante für die i -te Zeile:

$$\det(A) = \det \begin{bmatrix} \mathbf{a}_1^t \\ \vdots \\ \mathbf{a}_i^t \\ \vdots \\ \mathbf{a}_n^t \end{bmatrix} = \sum_{j=1}^n a_{ij} \det \begin{bmatrix} \mathbf{a}_1^t \\ \vdots \\ \mathbf{e}_j^t \\ \vdots \\ \mathbf{a}_n^t \end{bmatrix}.$$

Das ist eigentlich schon der Entwicklungssatz, nur nicht in einer sonderlich anwenderfreundlichen Form. Wir müssen noch etwas nacharbeiten. Dazu schauen wir uns die Determinanten in der Summe genauer an:

$$\det \begin{bmatrix} \mathbf{a}_1^t \\ \vdots \\ \mathbf{e}_j^t \\ \vdots \\ \mathbf{a}_n^t \end{bmatrix} = \det \begin{bmatrix} a_{11} & \cdots & a_{1j-1} & a_{1j} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j} & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj} & a_{nj+1} & \cdots & a_{nn} \end{bmatrix}$$

$$\stackrel{\text{i)}}{=} \det \begin{bmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn} \end{bmatrix}$$

$$\begin{aligned}
& \stackrel{\text{ii)}}{=} \det \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn} \end{bmatrix} \cdot (-1)^{i-1} \\
& \stackrel{\text{iii)}}{=} \det \left[\begin{array}{c|cccccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline 0 & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ \hline 0 & a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{array} \right] \cdot (-1)^{i+j} \\
& \stackrel{\text{iv)}}{=} \det \left[\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \hline a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \hline a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{array} \right] \cdot (-1)^{i+j} \\
& =: (-1)^{i+j} M_{ij}.
\end{aligned}$$

M_{ij} sind die sogenannten *Minoren*, oder *Unterdeterminanten* von A . Es sind die Determinanten der $(n-1) \times (n-1)$ -Matrizen, die aus A durch Streichen der i -ten Zeile und der j -ten Spalte entstehen.

Überlegen wir uns die einzelnen Schritte, die zu diesem Ergebnis geführt haben.

Zu i): Alle Pfade, die zur Determinante beitragen können, müssen die 1 an der Position (i, j) passieren. In keinem Summanden der Summendarstellung 6.44 kann also ein Faktor vorkommen, der oberhalb oder unterhalb der 1 in der j -ten Spalte steht. Wir ändern an dem Wert der Determinante nichts, wenn wir diese Einträge einfach durch 0 ersetzen.

Zu ii): Durch $i-1$ Vertauschungen benachbarter Zeilen bringen wir die i -te Zeile e_j^t in die erste Zeile.

Zu iii): Weitere $j-1$ Spaltenvertauschungen transportieren die j -te Spalte in die erste. Das Vorzeichen ändert sich dabei zu $(-1)^{j-1}(-1)^{i-1} = (-1)^{i+j}$.

Zu iv): Das ist nur noch die Anwendung von (6.46).

Die Entwicklung von $\det(A)$ nach der i -ten Zeile haben wir damit begründet. Man kann die Determinante aber auch nach der j -ten Spalte entwickeln. Es ist jedem nahegelegt, die geringfügigen Abänderungen unserer Überlegungen, die dafür nötig sind, selbst nachzuvollziehen.

6.12.10 Satz (LAPLACEscher Entwicklungssatz) Die Determinante einer $n \times n$ -Matrix A lässt sich nach einer beliebigen Zeile oder Spalte entwickeln. Die Entwicklung nach der i -ten Zeile bzw. der

j -ten Spalte ist respektive durch

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} M_{ij}, \quad (6.49)$$

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij} \quad (6.50)$$

gegeben. Dabei sind die Unterdeterminanten M_{ij} die Determinanten derjenigen Matrizen, die aus A durch Streichen der i -ten Zeile und der j -ten Spalte entstehen.

6.12.11 Beispiel Wir entwickeln nach der dritten Zeile:

$$\begin{aligned} \det \begin{bmatrix} 3^{+2-1+4-} \\ 3^{-3+2-0+} \\ 3^{+2-0+0-} \\ 2^{-3+7-4+} \end{bmatrix} &= +3 \cdot \det \begin{bmatrix} 3 & 2 & 1 & 4 \\ 3 & 3 & 2 & 0 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 7 & 4 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 3 & 2 & 1 & 4 \\ 3 & 3 & 2 & 0 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 7 & 4 \end{bmatrix} \\ &= 3 \cdot \det \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 0 \\ 3 & 7 & 4 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 3 & 1 & 4 \\ 3 & 2 & 0 \\ 2 & 7 & 4 \end{bmatrix} \\ &= 3 \cdot \left(4 \cdot \det \begin{bmatrix} 3 & 2 \\ 3 & 7 \end{bmatrix} + 4 \cdot \det \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \right) - 2 \cdot \left(4 \cdot \det \begin{bmatrix} 3 & 2 \\ 2 & 7 \end{bmatrix} + 4 \cdot \det \begin{bmatrix} 3 & 1 \\ 3 & 2 \end{bmatrix} \right) \\ &= 3 \cdot (4 \cdot (21 - 6) + 4 \cdot (4 - 3)) - 2 \cdot (4 \cdot (21 - 4) + 4 \cdot (6 - 3)) = 32. \end{aligned}$$

Das Vorzeichen $(-1)^{i+j}$ der Unterdeterminante M_{ij} findet man, in dem man in der Matrix A die linke obere Position in Gedanken mit einem $+$ markiert und dann die restlichen Felder des Quadrats schachbrettartig abwechselnd mit $-$ und $+$ versieht:

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

An der Position (i, j) dieser Vorzeichenmatrix steht das Vorzeichen von M_{ij} .

Das Beispiel zeigt es schon. Man wird den Entwicklungssatz bei großen Matrizen nur dann anwenden wollen, wenn die Matrix *schwach besetzt* ist, d. h., wenn sie nicht allzu viele Einträge hat, die von 0 verschieden sind. Ansonsten wird die Berechnung schnell sehr umfangreich: Die Determinante einer $n \times n$ -Matrix ergibt n Unterdeterminanten von $(n-1) \times (n-1)$ -Matrizen, die ihrerseits jeweils $n-1$ Unterdeterminanten von $(n-2) \times (n-2)$ -Matrizen erfordern, usw. Insgesamt sind das $n \cdot (n-1) \cdot (n-2) \cdots 2 = n!$ Summanden. Für eine 10×10 -Matrix könnte das bis auf 3 628 800 Summanden anwachsen.

6.12.12 Die Adjunkte und die inverse Matrix Wir wenden den LAPLACESchen Entwicklungssatz auf die Matrix an, die aus A entsteht, wenn wir die i -te Zeile \mathbf{a}_i^t durch die k -te Zeile \mathbf{a}_k^t

ersetzen, mit einem k , das von i verschieden ist. Wir wissen, daß die Determinante dieser Matrix dann Null ergibt, da ja die i -te und die k -te Zeile gleich sind. Wir entwickeln nach der i -ten Zeile und erhalten gemäß (6.49):

$$0 = \sum_{j=1}^n (-1)^{i+j} a_{kj} M_{ij}.$$

Gleichung (6.49) und dieses Ergebnis läßt sich in einer Formel zusammenfassen:

$$\sum_{j=1}^n (-1)^{i+j} a_{kj} M_{ij} = \det(A) \cdot \delta_{ki}. \quad (6.51)$$

Die rechte Seite gibt die Einträge der mit $\det(A)$ multiplizierten Einheitsmatrix wieder. Erinnern wir uns daran, wie der (k, i) -Eintrag eines Matrixprodukts aussieht, nämlich $(AB)_{ki} = \sum_{j=1}^n a_{kj} b_{ji}$, dann erkennen wir, daß die linke Seite von (6.51), bis auf die Reihenfolge der Indizes im zweiten Faktor, diese Form hat. Wir definieren die sogenannte *Adjunkte* \mathcal{A} von A durch die Einträge

$$\mathcal{A}_{ij} := (-1)^{i+j} M_{ij}. \quad (6.52)$$

Es ist die Matrix, die alle Unterdeterminanten, multipliziert mit dem zugehörigen Vorzeichen aus der Vorzeichenmatrix enthält. Gleichung (6.51) gibt dann $(A\mathcal{A}^t)_{ki}$ wieder. Als Ergebnis haben wir

$$A\mathcal{A}^t = \det(A) \cdot \mathbb{1}. \quad (6.53)$$

Sollte also $\det(A) \neq 0$ sein, dann ist nach Satz 6.9.2 die inverse Matrix von A durch

$$A^{-1} = \frac{1}{\det(A)} \cdot \mathcal{A}^t \quad (6.54)$$

gegeben.

6.12.13 Beispiel Beginnen wir mit $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Dann ist $\mathcal{A}_{11} = d$, $\mathcal{A}_{22} = a$, $\mathcal{A}_{12} = -c$ und $\mathcal{A}_{21} = -b$. Unter der Voraussetzung $\det(A) = ad - bc \neq 0$ ist

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \quad (6.55)$$

Zugegeben, dafür hätte man nicht (6.54) bemühen müssen, durch ein bisschen Nachdenken wäre man auch so darauf gekommen. Anders sieht es für eine 3×3 -Matrix aus. Hier kann es durchaus von Vorteil sein, für die inverse Matrix eine Formel zu haben, um etwa in zeitkritischen Rechenprozessen nicht mit dem GAUSS-Verfahren arbeiten zu müssen.

$$A := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

hat die Adjunkte

$$A = \begin{bmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{23}a_{31} - a_{21}a_{33} & a_{21}a_{32} - a_{22}a_{31} \\ a_{13}a_{32} - a_{12}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{12}a_{31} - a_{11}a_{32} \\ a_{12}a_{23} - a_{13}a_{22} & a_{13}a_{21} - a_{11}a_{23} & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix}$$

und, falls $\det(A) \neq 0$ ist, die Inverse

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix}. \quad (6.56)$$

Dann ist

$$\begin{bmatrix} 2 & 3 & -1 \\ 5 & 10 & 3 \\ 0 & 4 & 9 \end{bmatrix}^{-1} = \begin{bmatrix} 78 & -31 & 19 \\ -45 & 18 & -11 \\ 20 & -8 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3i & -1 \\ 5 & 10 & 3 \\ 2 & 4 & 9 \end{bmatrix}^{-1} = \frac{4+3i}{975} \begin{bmatrix} 78 & -4-27i & 10+9i \\ -39 & 20 & -11 \\ 0 & -8+6i & 20-15i \end{bmatrix}.$$

An Gleichung (6.56) sieht man auch, daß die Inverse einer 4×4 -Matrix wohl die Grenze für eine programmierbare Formel darstellt.

6.12.14 Der Rang einer Matrix Die Überlegungen, die uns zum GAUSS-Verfahren zur Berechnung von Determinanten geführt haben, können wir verwenden, um den Begriff *Rang einer Matrix* zu klären. Darunter wollen wir die maximale Anzahl linear unabhängiger Spaltenvektoren einer beliebigen $m \times n$ -Matrix A verstehen. Vorläufig sprechen wir vorsichtshalber von dem Spaltenrang von A . Denn es gibt auch eine maximale Anzahl linear unabhängiger Zeilenvektoren, und wir wissen vorerst nicht, ob beide übereinstimmen. Da das Bild einer Matrix die lineare Hülle ihrer Spaltenvektoren ist, handelt es sich beim Spaltenrang einer Matrix um die Dimension ihres Bildes. Mit der Dimensionsformel haben wir ein einfaches Mittel, um über die Bestimmung von $\ker A$ mit dem GAUSS-Verfahren auf die Dimension des Bildes schließen zu können. Das läßt vermuten, daß dieses Verfahren den Spaltenrang einer Matrix nicht ändert. Sicherlich ändert er sich nicht bei Vertauschung von Spaltenvektoren, aber der wesentliche Bestandteil des GAUSS-Verfahrens besteht in elementaren *Zeilenumformungen*. Hier ist es nicht mehr so offensichtlich, daß dabei der Spaltenrang, der ja die Spaltenvektoren betrifft, erhalten bleibt. Daß das doch der Fall ist, liegt an der Invertierbarkeit der quadratischen $m \times m$ -Matrizen $D_{ij}(\lambda)$ und T_{ij} in (6.47) bzw. (6.48), mit denen wir durch $D_{ij}(\lambda)A$ bzw. $T_{ij}A$ die Addition des λ -fachen der j -ten Zeile zur i -ten bzw. die Vertauschung der j -ten mit der i -ten Zeile bewirken können. Um das einzusehen, definieren wir für eine endliche Menge $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ von Vektoren den Begriff *Rang* als die größtmögliche Anzahl von Elementen einer linear unabhängigen Teilmenge $\{\mathbf{b}_{\ell_1}, \dots, \mathbf{b}_{\ell_k}\} \subseteq \mathcal{B}$.

6.12.15 Lemma Es sei $T: V \rightarrow W$ eine invertierbare lineare Abbildung zwischen den Vektorräumen V und W .

- i) Die Mengen $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subset V$ und $T(\mathcal{B}) := \{T\mathbf{b}_1, \dots, T\mathbf{b}_n\} \subset W$ haben den gleichen Rang.

- ii) Der Rang von \mathcal{B} ist $\dim \text{lh}(\mathcal{B})$.
- iii) Addiert man zu einem Vektor \mathbf{b}_i von \mathcal{B} eine Linearkombination der übrigen und ersetzt den Vektor durch das Ergebnis, so ändert sich der Rang der Menge dabei nicht.

Beweis. i): Der Rang der Menge $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ sei k und $\mathcal{C} := \{\mathbf{b}_{\ell_1}, \dots, \mathbf{b}_{\ell_k}\}$ eine linear unabhängige Teilmenge von \mathcal{B} mit k Elementen. Nach Proposition 6.8.3 ist $T(\mathcal{C}) = \{T\mathbf{b}_{\ell_1}, \dots, T\mathbf{b}_{\ell_k}\}$ wieder linear unabhängig. Also ist der Rang von $T(\mathcal{B})$ sicher nicht kleiner als k . Wäre er größer, so gäbe es eine linear unabhängige Teilmenge $\mathcal{D} := \{T\mathbf{b}_{t_1}, \dots, T\mathbf{b}_{t_{k+1}}\}$ von $T(\mathcal{B})$. Nach Proposition 6.8.3 wäre dann aber $T^{-1}(\mathcal{D}) = \{\mathbf{b}_{t_1}, \dots, \mathbf{b}_{t_{k+1}}\}$ eine linear unabhängige Teilmenge von \mathcal{B} mit $k+1$ Elementen, im Widerspruch dazu, daß solche Teilmengen höchstens k Elemente haben können.

ii): Die Teilmenge \mathcal{C} von \mathcal{B} ist eine Basis für $\text{lh}(\mathcal{B})$, denn aus der Maximalität von \mathcal{C} als linear unabhängige Teilmenge von \mathcal{B} folgt, daß alle Vektoren in $\mathcal{B} \setminus \mathcal{C}$ Linearkombinationen von Vektoren aus \mathcal{C} sein müssen. Daher ist $\text{lh}(\mathcal{B}) \subseteq \text{lh}(\mathcal{C}) \subseteq \text{lh}(\mathcal{B})$, also $\text{lh}(\mathcal{B}) = \text{lh}(\mathcal{C})$. \mathcal{C} ist erzeugend und natürlich linear unabhängig, also eine Basis. Insbesondere ist der Rang von \mathcal{B} einfach die Dimension von $\text{lh}(\mathcal{B})$.

iii): Wenn wir (o. B. d. A.) \mathbf{b}_1 durch $\tilde{\mathbf{b}}_1 := \mathbf{b}_1 + \sum_{k=2}^n \lambda_k \mathbf{b}_k$ ersetzen, dann hat die Menge $\tilde{\mathcal{B}} := \{\tilde{\mathbf{b}}_1, \dots, \mathbf{b}_n\} \subset V$ jedenfalls keinen größeren Rang als \mathcal{B} , denn $\text{lh}(\tilde{\mathcal{B}}) \subseteq \text{lh}(\mathcal{B})$. War \mathbf{b}_1 kein Element der Basis \mathcal{C} , dann ist $\text{lh}(\mathcal{B}) = \text{lh}(\tilde{\mathcal{B}})$, denn beide Räume enthalten die Basis \mathcal{C} . Für $\mathbf{b}_1 \in \mathcal{C}$ dagegen ist $\tilde{\mathcal{C}} := \{\tilde{\mathbf{b}}_1, \dots, \mathbf{b}_{\ell_k}\}$ nach dem Basis-Austausch-Satz 6.5.3 wieder eine Basis für $\text{lh}(\mathcal{B})$, so daß auch jetzt $\text{lh}(\mathcal{B}) = \text{lh}(\tilde{\mathcal{B}})$ gilt. In jedem Fall stimmt der Rang von $\tilde{\mathcal{B}}$ und \mathcal{B} überein. \square

Dieses Lemma zeigt, daß sich der Spaltenrang k einer Matrix, also der Rang der Menge ihrer Spaltenvektoren, nicht ändert, wenn wir sie durch das GAUSS-Verfahren in eine Matrix mit maximaler Diagonalenlänge k verwandeln:

$$\begin{bmatrix} a_1 & & & * & \cdots & * \\ & a_2 & & * & \cdots & * \\ & & \ddots & \vdots & \cdots & \vdots \\ & & & a_k & * & \cdots & * \\ 0 & & & & 0 & & \end{bmatrix} \quad (6.57)$$

Wie sieht das für den Zeilenrang von A aus? Das GAUSS-Verfahren vertauscht Vektoren, was den Zeilenrang natürlich nicht ändert. Außerdem wird zu einem Zeilenvektor das Vielfache anderer Zeilenvektoren addiert. Auch das ändert den Zeilenrang nicht, wie Lemma 6.12.15 iii) lehrt. Bleibt noch die Auswirkung der Vertauschung von Spaltenvektoren auf den Zeilenrang. Wir gehen von A zu A^t über, um zu erkennen, daß wir diesen Fall schon behandelt haben. Jetzt geht es um die Spaltenvektoren von A^t bei Vertauschung zweier Zeilen. Das wird durch Multiplikation von A^t mit einer passenden Matrix T_{ij} erreicht, die den Spaltenrang nicht ändert. Im Ergebnis bleibt auch der Zeilenrang von A unter dem GAUSS-Verfahren konstant. Das bedeutet, daß wir den Spalten- und den Zeilenrang einer Matrix am Endergebnis (6.57) ablesen

können. Und hier ist klar zu erkennen, daß es genauso viel linear unabhängige Spaltenvektoren wie Zeilenvektoren gibt.

6.12.16 Satz Für eine beliebige $m \times n$ -Matrix A stimmt die maximale Anzahl linear unabhängiger Zeilenvektoren mit der maximalen Anzahl linear unabhängiger Spaltenvektoren überein. Diese Zahl wird Rang der Matrix, $\text{ran } A$, genannt.

6.12.17 A Überlegen Sie sich, daß das Ergebnis (6.56) für die Inverse einer 3×3 -Matrix auch direkt aus ein paar Grundprinzipien der Matrizenrechnung gewonnen werden kann. Dafür beschreiben Sie die Matrix A über ihre Zeilenvektoren \mathbf{a}_1 , \mathbf{a}_2 und \mathbf{a}_3 :

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix}.$$

Für die inverse Matrix $A^{-1} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$ muß dann

$$\begin{bmatrix} \mathbf{a}_1 \mathbf{b}_1 & \mathbf{a}_1 \mathbf{b}_2 & \mathbf{a}_1 \mathbf{b}_3 \\ \mathbf{a}_2 \mathbf{b}_1 & \mathbf{a}_2 \mathbf{b}_2 & \mathbf{a}_2 \mathbf{b}_3 \\ \mathbf{a}_3 \mathbf{b}_1 & \mathbf{a}_3 \mathbf{b}_2 & \mathbf{a}_3 \mathbf{b}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

gelten. Zeigen Sie, daß A^{-1} folgendermaßen berechnet werden kann:

$$A^{-1} = \frac{1}{\langle \mathbf{a}_1^t \times \mathbf{a}_2^t | \mathbf{a}_3^t \rangle} [\mathbf{a}_2^t \times \mathbf{a}_3^t, \mathbf{a}_3^t \times \mathbf{a}_1^t, \mathbf{a}_1^t \times \mathbf{a}_2^t]. \quad (6.58)$$

7 Eigenwerttheorie

7.1 Spektrum und Eigenvektoren

Eine lineare Abbildung A von einem Vektorraum V nach V ist meistens als quadratisches Zahlenschema gegeben, dem man im Allgemeinen nicht direkt ansehen kann, wie sie denn genau wirkt. Die Bilder von Vektoren lassen sich leicht berechnen, doch kommt man auf diese Weise einem Verständnis der Abbildung normalerweise nicht näher: Vektoren werden eben gestreckt und gedreht und manche werden auf Null abgebildet und bilden den Kern der Abbildung. Man könnte also auf die Idee kommen, nach Richtungen zu suchen, die durch die Abbildung nicht verändert werden. Diese müssen durch Vektoren $x \neq 0$ beschrieben werden, die von A nur gestreckt werden, sagen wir um den Faktor λ :

$$Ax = \lambda x. \quad (7.1)$$

Einen solchen Vektor x nennt man *Eigenvektor* der Abbildung A . Der zugehörige Streckungsfaktor λ heißt *Eigenwert* von A . Gleichung (7.1) heißt *Eigenwertgleichung*. Die Menge $\sigma(A)$ aller Eigenwerte von A wird *Spektrum* genannt.

Es stellt sich die Frage, ob es überhaupt Eigenvektoren gibt und wenn ja, wie man sie findet. Einer bestimmten Sorte von Eigenvektoren sind wir schon begegnet, nämlich den Vektoren aus $\ker A$. Sie sind Eigenvektoren zum Eigenwert 0: $Ax = 0 = 0 \cdot x$. Denken wir aber an eine ebene Drehung um einen festen Winkel, so können wir uns beim besten Willen keinen Vektor vorstellen, der von 0 verschieden ist und nur gestreckt wird. Der Grund ist, daß es auch keinen reellen Vektor mit dieser Eigenschaft gibt. Wir werden aber bald sehen, daß es für komplexe Vektorräume immer wenigstens einen Eigenvektor geben muß (manchmal aber auch nicht mehr).

Das Problem bei der Lösung von (7.1) ist, daß uns *zwei* Dinge fehlen, nämlich der Eigenvektor und der Eigenwert. Wäre uns der Eigenwert bekannt, dann ließe sich (7.1) in das Gleichungssystem $Ax - \lambda x = (A - \lambda \mathbb{1})x = 0$ umwandeln, das mit dem GAUSS-Verfahren zu lösen wäre. Hätten wir dagegen x , dann wäre λ leicht aus dem Vergleich von Ax mit x zu bestimmen. Um zu einem Lösungsverfahren zu kommen, müssen wir das Problem so umformen, daß nur noch der Eigenvektor oder der Eigenwert vorkommt. Dafür stehen uns aus der bisher entwickelten Theorie linearer Abbildungen alle Werkzeuge zur Verfügung:

$$\begin{aligned} & \exists_{0 \neq x} Ax = \lambda x \\ \Leftrightarrow & \exists_{0 \neq x} (A - \lambda \mathbb{1})x = 0 \\ \Leftrightarrow & \ker(A - \lambda \mathbb{1}) \neq \{0\} \\ \Leftrightarrow & A - \lambda \mathbb{1} \text{ ist nicht invertierbar} \end{aligned}$$

$$\Leftrightarrow \det(A - \lambda \mathbb{1}) = 0.$$

Die letzte Bedingung folgt aus Satz 6.12.6. Sie hängt nur noch von den möglichen Eigenwerten λ ab und eignet sich daher dazu, erst einmal alle Eigenwerte zu bestimmen. $\det(A - \lambda \mathbb{1})$ ist ein Polynom der Variablen λ . Es wird *charakteristisches Polynom* von A genannt und mit $\chi(\lambda)$ bezeichnet. Die Eigenwerte von A sind die Nullstellen des charakteristischen Polynoms. Hat man diese bestimmt, so lassen sich die zugehörigen Eigenvektoren \mathbf{x} mit Hilfe des GAUSS-Verfahrens aus der homogenen Gleichung $(A - \lambda \mathbb{1})\mathbf{x} = \mathbf{0}$ bestimmen. Da nach dem Fundamentalsatz der Algebra 11.8.23 jedes Polynom wenigstens eine Nullstelle haben muß (die komplex sein könnte), gibt es für jede lineare Abbildung wenigstens einen Eigenvektor.

7.1.1 Beispiel

Für $A := \begin{bmatrix} 11 & -2 \\ 3 & 4 \end{bmatrix}$ ist $A - \lambda \mathbb{1} = \begin{bmatrix} 11 & -2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 11-\lambda & -2 \\ 3 & 4-\lambda \end{bmatrix}$ und $\chi(\lambda) = (11-\lambda)(4-\lambda) + 6 = \lambda^2 - 15\lambda + 50 = (\lambda-10)(\lambda-5)$. Daher ist $\sigma(A) = \{5, 10\}$. Der Eigenvektor \mathbf{x} zum Eigenwert 5 bestimmt sich aus der Gleichung $(A-5\mathbb{1})\mathbf{x} = \begin{bmatrix} 6 & -2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Das zugehörige GAUSS-Verfahren führt auf das Schema $\begin{array}{cc|c} 6 & -2 & 0 \\ 3 & -1 & 0 \end{array}$. Da wir $\lambda = 5$ gerade so bestimmt haben, daß für die Gleichung $(A - 5\mathbb{1})\mathbf{x} = \mathbf{0}$ eine nicht triviale Lösung \mathbf{x} existiert, muß während des GAUSS-Verfahrens wenigstens eine der Zeilen verschwinden. In diesem einfachen Fall haben wir aber nur zwei. Wir können also einfach die zweite Zeile streichen und müssen nur noch $6x_1 - 2x_2 = 0$ lösen. Wir sind nur an einer Lösung interessiert. Daher wählen wir z. B. $x_2 = 3$ und finden damit $x_1 = 1$. Der Eigenvektor zum Eigenwert 5 ist $[1, 3]^t$. Genauso findet man den Eigenvektor $[2, 1]^t$ zum Eigenwert 10. Daß die beiden Eigenvektoren linear unabhängig sind ist hier kein Zufall, wie der folgende Satz beweist. \mathbb{R}^2 hat eine Basis aus Eigenvektoren von A . Daß das aber nicht immer so ist, zeigt die Abbildung $B := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Man überzeugt sich leicht von $\sigma(B) = \{0\}$. Ein Eigenvektor ist offensichtlich $\mathbf{x}_1 = [1, 0]^t$. Gäbe es noch einen weiteren \mathbf{x}_2 , so daß $\{\mathbf{x}_1, \mathbf{x}_2\}$ eine Basis bildet, dann würde jeder Vektor aus \mathbb{R}^2 auf Null abgebildet. B müßte die Nullabbildung sein, was offensichtlich nicht der Fall ist. Also hat B (bis auf Vielfache) nur den einen Eigenvektor \mathbf{x}_1 .

7.1.2 Satz Die Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Beweis. Wir führen den Beweis erst einmal nur für zwei verschiedene Eigenwerte $\lambda_1 \neq \lambda_2$ einer linearen Abbildung A . Die zugehörigen Eigenvektoren seien \mathbf{x}_1 und \mathbf{x}_2 . Für eine Nullkombination $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 = \mathbf{0}$ folgt $(A - \lambda_1\mathbb{1})(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1(A\mathbf{x}_1 - \lambda_1\mathbf{x}_1) + \alpha_2(A\mathbf{x}_2 - \lambda_1\mathbf{x}_2) = \alpha_2(\lambda_2 - \lambda_1)\mathbf{x}_2 = \mathbf{0}$. Wegen $\mathbf{x}_2 \neq \mathbf{0}$ und $\lambda_1 \neq \lambda_2$ muß $\alpha_2 = 0$ gelten. Das bedeutet $\alpha_1\mathbf{x}_1 = \mathbf{0}$. Wegen $\mathbf{x}_1 \neq \mathbf{0}$ folgt auch $\alpha_1 = 0$, d. h. die Nullkombination ist trivial. $\{\mathbf{x}_1, \mathbf{x}_2\}$ ist daher linear unabhängig.

Jetzt der allgemeine Fall: $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ und $\lambda_i \neq \lambda_j$ für $i \neq j$, $i, j \in \{1, \dots, n\}$. $\sum_{i=1}^n \alpha_i\mathbf{x}_i = \mathbf{0}$

sei eine Nullkombination der Eigenvektoren. Dann folgt für ein beliebiges $k \in \{1, \dots, n\}$

$$\prod_{j \neq k} (A - \lambda_j \mathbb{1}) \sum_{i=1}^n \alpha_i \mathbf{x}_i = \sum_{i=1}^n \alpha_i \prod_{j \neq k} (A - \lambda_j \mathbb{1}) \mathbf{x}_i = \alpha_k \prod_{j \neq k} (\lambda_k - \lambda_j) \mathbf{x}_k = \mathbf{0}.$$

Das bedeutet $\alpha_k = 0$ für jedes $k \in \{1, \dots, n\}$, d. h. die Nullkombination ist trivial. Somit ist $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ linear unabhängig. \square

7.2 Selbstadjungierte lineare Abbildungen

7.2.1 Definition Sei V ein endlich dimensionaler Vektorraum über \mathbb{K} , versehen mit einem Skalarprodukt $\langle \cdot | \cdot \rangle$ und A eine lineare Abbildung von V nach V mit der Eigenschaft

$$\langle \mathbf{x} | A\mathbf{y} \rangle = \langle A\mathbf{x} | \mathbf{y} \rangle \quad (7.2)$$

für alle $\mathbf{x}, \mathbf{y} \in V$. Dann heißt A selbstadjungiert.

Ist die lineare Abbildung A durch eine quadratische Matrix auf $V = \mathbb{K}^n$ gegeben, so ist A genau dann selbstadjungiert, wenn

$$A = A^* \quad (7.3)$$

gilt (vergl. Satz 6.10.2). Insbesondere sind also die Diagonalelemente a_{ii} von A reell.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \overline{a_{12}} & \overline{a_{22}} & \overline{a_{23}} & \cdots & \overline{a_{2n}} \\ \overline{a_{13}} & \overline{a_{23}} & \overline{a_{33}} & \cdots & \overline{a_{3n}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \overline{a_{3n}} & \cdots & \overline{a_{nn}} \end{bmatrix}.$$

Für die folgenden Untersuchungen gehen wir generell von einem Vektorraum V mit Skalarprodukt $\langle \cdot | \cdot \rangle$ aus.

7.2.2 Lemma Für eine selbstadjungierte lineare Abbildung ist das Spektrum reell.

Beweis. Es sei $\mathbf{x} \neq \mathbf{0}$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$. Dann gilt

$$\begin{aligned} \langle \mathbf{x} | A\mathbf{x} \rangle &= \langle \mathbf{x} | \lambda \mathbf{x} \rangle = \lambda \langle \mathbf{x} | \mathbf{x} \rangle = \\ \langle A\mathbf{x} | \mathbf{x} \rangle &= \langle \lambda \mathbf{x} | \mathbf{x} \rangle = \bar{\lambda} \langle \mathbf{x} | \mathbf{x} \rangle. \end{aligned}$$

Das bedeutet $(\lambda - \bar{\lambda}) \langle \mathbf{x} | \mathbf{x} \rangle = 0$. Wegen der positiven Definitheit des Skalarprodukts ist $\langle \mathbf{x} | \mathbf{x} \rangle \neq 0$, so daß nur $\lambda - \bar{\lambda} = 0$ bleibt. Das heißt aber $\lambda \in \mathbb{R}$. Das Spektrum $\text{sp}(A)$ ist also in \mathbb{R} enthalten. \square

7.2.3 Lemma Für eine selbstadjungierte lineare Abbildung stehen die Eigenvektoren zu verschiedenen Eigenwerten senkrecht aufeinander.

Beweis. Es seien $\lambda \neq \gamma$ zwei Eigenwerte zu den Eigenvektoren \mathbf{x} bzw. \mathbf{y} . Dann gilt

$$\begin{aligned}\langle \mathbf{x} | A\mathbf{y} \rangle &= \langle \mathbf{x} | \gamma \mathbf{y} \rangle = \gamma \langle \mathbf{x} | \mathbf{y} \rangle = \\ \langle A\mathbf{x} | \mathbf{y} \rangle &= \langle \lambda \mathbf{x} | \mathbf{y} \rangle = \lambda \langle \mathbf{x} | \mathbf{y} \rangle.\end{aligned}$$

Daraus folgt $(\lambda - \gamma) \langle \mathbf{x} | \mathbf{y} \rangle = 0$. Da $\lambda - \gamma \neq 0$ gilt, muß $\langle \mathbf{x} | \mathbf{y} \rangle = 0$ erfüllt sein, d. h. \mathbf{x} und \mathbf{y} sind zueinander orthogonal. \square

7.2.4 Theorem Für jede selbstadjungierte lineare Abbildung A auf einem endlichdimensionalen Vektorraum V gibt es eine ONB aus Eigenvektoren von A .

Beweis. Wir führen den Beweis durch Induktion nach der Dimension n von V . Der Fall $n = 1$ ist trivial.

$n \rightarrow n + 1$: Es gilt $\dim V = n + 1$. Ist $A = 0$, so leistet jede ONB von V das Gewünschte. Sei also $A \neq 0$. Dann gibt es wenigstens einen Eigenwert λ von A , denn das charakteristische Polynom von A hat in \mathbb{C} wenigstens eine Nullstelle (Fundamentalsatz der Algebra 11.8.23). Also gibt es einen Eigenvektor \mathbf{x} ($\|\mathbf{x}\| = 1$) zu λ . Wir definieren den Vektorraum $V_1 := \mathbf{x}^\perp$ als den zu \mathbf{x} orthogonalen Teilraum von V . Nach Proposition 6.7.6 hat V_1 die Dimension n . Die Beweisidee besteht nun darin, nachzuweisen, daß A den Vektorraum V_1 in sich abbildet und daß A auf V_1 wieder selbstadjungiert ist. Dann besagt die Induktionsvoraussetzung, daß wir eine ONB $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ von V_1 aus Eigenvektoren von A haben, die zusammen mit \mathbf{x} eine ONB aus Eigenvektoren für ganz V ergibt.

Sei also $\mathbf{y} \in V_1$, dann ist $A\mathbf{y} \in V_1$, also $A\mathbf{y} \perp \mathbf{x}$ nachzuweisen. Das folgt aber sofort aus

$$\langle \mathbf{x} | A\mathbf{y} \rangle = \langle A\mathbf{x} | \mathbf{y} \rangle = \lambda \langle \mathbf{x} | \mathbf{y} \rangle = 0.$$

Die Selbstadjungiertheit von A auf V_1 ist nun klar, da $\langle \mathbf{y} | A\mathbf{z} \rangle = \langle A\mathbf{y} | \mathbf{z} \rangle$ für alle $\mathbf{y}, \mathbf{z} \in V$, erst recht also auch für alle $\mathbf{y}, \mathbf{z} \in V_1$ gilt. \square

7.2.5 A Bestimmen Sie die Spektren und die zugehörigen Eigenvektoren der sogenannten PAULI-Matrizen

$$\sigma_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (7.4)$$

7.2.6 Definition Eine selbstadjungierte lineare Abbildung P mit der Eigenschaft

$$P^2 = P \quad (7.5)$$

heißt (orthogonale) Projektion oder Projektor. $[P] := \text{im } P$ bezeichnet ihren Projektionsraum.

Offensichtlich ist mit P auch $\mathbb{1} - P$ eine Projektion, denn $\mathbb{1} - P$ ist selbstadjungiert und es gilt $(\mathbb{1} - P)^2 = \mathbb{1} - 2P + P^2 = \mathbb{1} - P$. Es gibt allgemeinere Projektionen, nämlich lineare Abbildungen P , die zwar die Projektionseigenschaft $P^2 = P$ haben, aber nicht selbstadjungiert sind. Da

wir uns hier nur für orthogonale Projektionen interessieren, sprechen wir einfach von Projektionen und meinen damit orthogonale Projektionen. Wegen

$$(\mathbb{1} - P)P = P - P^2 = \mathbb{0} = P(\mathbb{1} - P)$$

wird $\mathbb{1} - P$ als die zu P *orthogonale Projektion* bezeichnet und mitunter auch P^\perp geschrieben. Die Projektionsräume $[P]$ und $[\mathbb{1} - P]$ sind nämlich orthogonal zueinander: Für ein $\mathbf{x} \in [P]$ und ein $\mathbf{y} \in [\mathbb{1} - P]$ gilt $\mathbf{x} = P\mathbf{x}$ und $\mathbf{y} = (\mathbb{1} - P)\mathbf{y}$ (warum?). Das hat

$$\langle \mathbf{x} | \mathbf{y} \rangle = \langle P\mathbf{x} | (\mathbb{1} - P)\mathbf{y} \rangle = \langle \mathbf{x} | P(\mathbb{1} - P)\mathbf{y} \rangle = 0,$$

also $P\mathbf{x} \perp P\mathbf{y}$ zur Folge. Das zeigt bereits $[P] \perp [\mathbb{1} - P]$. Tatsächlich gilt sogar noch mehr:

7.2.7 Lemma Für eine Projektion P gilt $[P]^\perp = [\mathbb{1} - P]$.

$[P]$ und $[\mathbb{1} - P]$ sind also nicht nur orthogonal zueinander, sondern $[\mathbb{1} - P]$ ist bereits der *ganze* zu $[P]$ orthogonale Teilraum.

Beweis. Da wir $[P] \perp [\mathbb{1} - P]$ bereits wissen, haben wir schon $[\mathbb{1} - P] \subseteq [P]^\perp$. Um die umgekehrte Inklusion zu zeigen, sei $\mathbf{y} \in [P]^\perp$. Dann gilt $\langle \mathbf{y} | P\mathbf{x} \rangle = \langle P\mathbf{y} | \mathbf{x} \rangle = 0$ für alle $\mathbf{x} \in V$. Nach Übung 5.5.12 ist $P\mathbf{y} = \mathbf{0}$, also $(\mathbb{1} - P)\mathbf{y} = \mathbf{y}$. Daher liegt \mathbf{y} in $[\mathbb{1} - P]$. Das zeigt $[P]^\perp \subseteq [\mathbb{1} - P]$. \square

Die Aussage dieses Lemmas läßt sich auch so formulieren:

$$[P]^\perp = [P^\perp]. \quad (7.6)$$

Der zum Projektionsraum $[P]$ orthogonale Teilraum $[P]^\perp$ ist der Projektionsraum $[\mathbb{1} - P] = [P^\perp]$ der zu P orthogonalen Projektion P^\perp . Dieser Sachverhalt rechtfertigt zur Genüge die Namensgebung für $\mathbb{1} - P$.

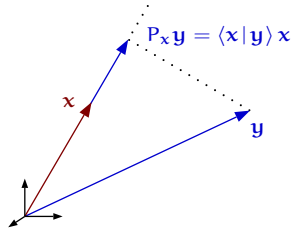
Offensichtlich ist auch $\mathbb{1}$ und $\mathbb{0}$ eine Projektion. Projektionen, die von diesen verschieden sind, heißen *nicht trivial*.

7.2.8 Lemma Das Spektrum einer nicht trivialen Projektion ist immer die Menge $\{0, 1\}$. Der Eigenraum zum Eigenwert 1 ist der Projektionsraum von P und der zum Eigenwert 0 ist der Projektionsraum von $\mathbb{1} - P$.

Beweis. Ist $P \neq 0$ und $P \neq \mathbb{1}$, so gibt es immer einen Vektor $\mathbf{x} \neq \mathbf{0}$ im Projektionsraum und einen dazu orthogonalen $\mathbf{y} \neq \mathbf{0}$ (warum?). Damit folgt $P\mathbf{x} = \mathbf{x}$ und $(\mathbb{1} - P)\mathbf{y} = \mathbf{y}$, also $P\mathbf{y} = \mathbf{0}$. Folglich liegen schon 1 und 0 im Spektrum von P : $\{0, 1\} \subseteq \text{sp}(P)$. Bleibt zu zeigen, daß es keine weiteren Elemente haben kann. Aus $P\mathbf{z} = \lambda\mathbf{z}$ für ein $\mathbf{z} \neq \mathbf{0}$ folgt nämlich $\lambda\mathbf{z} = P\mathbf{z} = PP\mathbf{z} = \lambda P\mathbf{z} = \lambda^2\mathbf{z}$, d. h. $\lambda^2 = \lambda$. Da λ als Eigenwert einer selbstadjungierten Abbildung reell sein muß, bleiben nur die Möglichkeiten $\lambda = 0$, oder $\lambda = \pm 1$. Aber $P\mathbf{z} = -\mathbf{z}$ führt auf den Widerspruch

$$0 < \langle \mathbf{z} | \mathbf{z} \rangle = \langle P\mathbf{z} | P\mathbf{z} \rangle = \langle \mathbf{z} | P\mathbf{z} \rangle = -\langle \mathbf{z} | \mathbf{z} \rangle < 0.$$

Damit haben wir auch $\text{sp}(P) \subseteq \{0, 1\}$. Der Rest ist klar. \square



Die Dimension des Projektionsraumes $[P]$ nennt man auch *Dimension der Projektion*. Am leichtesten sind eindimensionale Projektionen zu bestimmen. Ihr Projektionsraum wird durch einen normierten Vektor x aufgespannt. Die zugehörige Projektion bezeichnen wir mit P_x . Wenn wir uns daran erinnern, daß das Skalarprodukt $\langle x|y \rangle$ gerade die Länge der Projektion von y auf die Richtung x beschreibt, ist die Wirkung von P_x leicht anzugeben:

$$P_x y = \langle x|y \rangle x. \quad (7.7)$$

Um die Matrix für P_x zu finden, gehen wir von der Koordinatendarstellung $x = [x_1, \dots, x_n]^t \in \mathbb{C}^n$ aus:

$$\begin{aligned} P_x y &= \begin{bmatrix} x_1 \langle x|y \rangle \\ x_2 \langle x|y \rangle \\ \vdots \\ x_n \langle x|y \rangle \end{bmatrix} = \begin{bmatrix} x_1 \overline{x_1} y_1 + x_1 \overline{x_2} y_2 + x_1 \overline{x_3} y_3 + \dots + x_1 \overline{x_n} y_n \\ x_2 \overline{x_1} y_1 + x_2 \overline{x_2} y_2 + x_2 \overline{x_3} y_3 + \dots + x_2 \overline{x_n} y_n \\ \vdots \\ x_n \overline{x_1} y_1 + x_n \overline{x_2} y_2 + x_n \overline{x_3} y_3 + \dots + x_n \overline{x_n} y_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 \overline{x_1} & x_1 \overline{x_2} & x_1 \overline{x_3} & \dots & x_1 \overline{x_n} \\ x_2 \overline{x_1} & x_2 \overline{x_2} & x_2 \overline{x_3} & \dots & x_2 \overline{x_n} \\ \vdots & \vdots & \vdots & & \vdots \\ x_n \overline{x_1} & x_n \overline{x_2} & x_n \overline{x_3} & \dots & x_n \overline{x_n} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}. \end{aligned}$$

Damit ist die Matrix bestimmt:

$$P_x = \begin{bmatrix} x_1 \overline{x_1} & x_1 \overline{x_2} & x_1 \overline{x_3} & \dots & x_1 \overline{x_n} \\ x_2 \overline{x_1} & x_2 \overline{x_2} & x_2 \overline{x_3} & \dots & x_2 \overline{x_n} \\ \vdots & \vdots & \vdots & & \vdots \\ x_n \overline{x_1} & x_n \overline{x_2} & x_n \overline{x_3} & \dots & x_n \overline{x_n} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^* = [x_i \overline{x_j}]_{i,j=1,\dots,n}. \quad (7.8)$$

7.2.9 A Zeigen Sie, daß durch Gleichung (7.7) tatsächlich eine Projektion definiert wird.

7.2.10 Lemma Zu jedem Teilraum T von V gibt es genau eine Projektion P , die T als Projektionsraum hat: $T = [P]$. Wir nennen P den Projektor auf den Teilraum T .

Die Projektoren auf die trivialen Teilräume $T = \mathbf{0}$ und $T = V$ sind natürlich $P = 0$ bzw. $P = \mathbb{1}$.

Beweis. T ist Teilvektorraum von V und hat daher nach Satz 6.5.5 eine Basis, die nach dem GRAM-SCHMIDTSchen Orthonormalisierungsverfahren 6.5.16 zu einer ONB $\{b_1, \dots, b_k\}$ gemacht werden kann. Jetzt läßt sich P durch

$$Px := \sum_{i=1}^k \langle x|b_i \rangle b_i, \quad (7.9)$$

$\mathbf{x} \in V$, definieren. Wir rechnen die Projektionseigenschaft nach:

$$\begin{aligned} P^2\mathbf{x} &= P(P\mathbf{x}) = \sum_{i=1}^k \langle P\mathbf{x} | \mathbf{b}_i \rangle \mathbf{b}_i = \sum_{i=1}^k \sum_{j=1}^k \langle \langle \mathbf{x} | \mathbf{b}_j \rangle \mathbf{b}_j | \mathbf{b}_i \rangle \mathbf{b}_i = \sum_{i=1}^k \sum_{j=1}^k \langle \mathbf{x} | \mathbf{b}_j \rangle \langle \mathbf{b}_j | \mathbf{b}_i \rangle \mathbf{b}_i \\ &= \sum_{i=1}^k \sum_{j=1}^k \langle \mathbf{x} | \mathbf{b}_j \rangle \delta_{ji} \mathbf{b}_i = \sum_{i=1}^k \langle \mathbf{x} | \mathbf{b}_i \rangle \mathbf{b}_i = P\mathbf{x}. \end{aligned}$$

Das zeigt $P^2 = P$. Bleibt noch $P = P^*$ zu zeigen: Für alle $\mathbf{x}, \mathbf{y} \in V$ gilt

$$\begin{aligned} \langle \mathbf{x} | P\mathbf{y} \rangle &= \left\langle \mathbf{x} \left| \sum_{i=1}^k \langle \mathbf{y} | \mathbf{b}_i \rangle \mathbf{b}_i \right. \right\rangle = \sum_{i=1}^k \overline{\langle \mathbf{y} | \mathbf{b}_i \rangle} \langle \mathbf{x} | \mathbf{b}_i \rangle = \sum_{i=1}^k \langle \mathbf{x} | \mathbf{b}_i \rangle \langle \mathbf{b}_i | \mathbf{y} \rangle = \left\langle \sum_{i=1}^k \langle \mathbf{x} | \mathbf{b}_i \rangle \mathbf{b}_i \left| \mathbf{y} \right. \right\rangle \\ &= \langle P\mathbf{x} | \mathbf{y} \rangle. \end{aligned}$$

Für den Projektor P gilt laut Konstruktion $P\mathbf{x} \in T$, d. h., wir haben bereits $[P] \subseteq T$. Jeder Vektor $\mathbf{x} \in T$ hat eine Basisdarstellung $\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{b}_i$. Wegen $P\mathbf{b}_j = \sum_{i=1}^k \langle \mathbf{b}_j | \mathbf{b}_i \rangle \mathbf{b}_i = \sum_{i=1}^k \delta_{ji} \mathbf{b}_i = \mathbf{b}_j$ folgt sofort $P\mathbf{x} = \mathbf{x} \in [P]$. Das zeigt die andere Inklusion $T \subseteq [P]$ und damit $T = [P]$. \square

7.2.11 Definition Eine selbstadjungierte Abbildung A heißt positiv, wenn für alle $\mathbf{x} \in V$

$$\langle \mathbf{x} | A\mathbf{x} \rangle \geq 0 \quad (7.10)$$

gilt. Dafür schreibt man auch $A \geq 0$.

7.2.12 Satz Eine selbstadjungierte Abbildung A ist genau dann positiv, wenn $\sigma(A) \subseteq \mathbb{R}_0^+$ gilt.

Beweis. Zunächst sei $A \geq 0$ und \mathbf{x} ein normierter Eigenvektor zum Eigenwert λ . Dann gilt $0 \leq \langle \mathbf{x} | A\mathbf{x} \rangle = \langle \mathbf{x} | \lambda \mathbf{x} \rangle = \lambda \langle \mathbf{x} | \mathbf{x} \rangle = \lambda$. Das zeigt $\text{sp}(A) \subset \mathbb{R}_0^+$. Gehen wir umgekehrt von dieser Eigenschaft aus, dann sind also alle Eigenwerte $\lambda_1, \dots, \lambda_n$ nicht negativ. $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ sei eine ONB aus Eigenvektoren. Dann läßt sich jeder Vektor \mathbf{x} nach dieser Basis entwickeln: $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{b}_k$. Es folgt

$$\begin{aligned} \langle \mathbf{x} | A\mathbf{x} \rangle &= \left\langle \sum_{k=1}^n x_k \mathbf{b}_k \left| \sum_{l=1}^n x_l A\mathbf{b}_l \right. \right\rangle = \left\langle \sum_{k=1}^n x_k \mathbf{b}_k \left| \sum_{l=1}^n x_l \lambda_l \mathbf{b}_l \right. \right\rangle \\ &= \sum_{k=1}^n \sum_{l=1}^n \overline{x_k} x_l \lambda_l \langle \mathbf{b}_k | \mathbf{b}_l \rangle = \sum_{k=1}^n |x_k|^2 \lambda_k \geq 0. \end{aligned} \quad \square$$

Dieses Ergebnis wird zur Identifikation lokaler Extrema in der mehrdimensionalen Analysis gebraucht. Insbesondere ist jede Projektion eine positive Abbildung, denn ihr Spektrum ist in $\{0, 1\} \subset \mathbb{R}_0^+$ enthalten.

7.3 Funktionalkalkül

7.3.1 A Zeigen Sie:

- i) Für jede lineare Abbildung A ist die Abbildung A^*A positiv.
- ii) A ist genau dann positiv, wenn es eine positive Abbildung W mit der Eigenschaft $A = W^2$ gibt. W wird als *Wurzel* von A bezeichnet.

7.3.2 A (*) Zeigen Sie, daß auf Vektorräumen V über \mathbb{C} eine lineare Abbildung A genau dann selbstadjungiert ist, wenn $\langle x | Ax \rangle$ für alle $x \in V$ reell ist. Gehen Sie dabei folgendermaßen vor:

- i) Zeigen Sie, daß für jede lineare Abbildung A die *sesquilineare Abbildung* $V \times V \ni (x, y) \mapsto \langle x | Ay \rangle$ eine *Polarisationsgleichung* erfüllt:

$$\begin{aligned} \langle x | Ay \rangle = \frac{1}{4} [& \langle x + y | A(x + y) \rangle - \langle x - y | A(x - y) \rangle \\ & - i(\langle x + iy | A(x + iy) \rangle - \langle x - iy | A(x - iy) \rangle)]. \end{aligned}$$

- ii) Verwenden Sie Identitäten der Art $\langle x + iy | A(x + iy) \rangle = \langle y - ix | A(y - ix) \rangle$ etc., um aus i) auf $\langle x | Ay \rangle = \overline{\langle y | Ax \rangle}$ zu schließen.

7.3.3 Theorem Sind A und B selbstadjungierte Abbildungen auf dem Vektorraum V , die miteinander vertauschen ($AB = BA$), so gibt es eine gemeinsame ONB aus Eigenvektoren von A und B für V .

Beweis. $V_{\lambda_1} := \{ y \in V \mid Ay = \lambda_1 y \}$ sei der Eigenraum von A zum Eigenwert λ_1 (also die Menge aller Vektoren, die Eigenvektor zum selben Eigenwert λ_1 sind). Ist $\lambda_2 \neq \lambda_1$ ein weiterer Eigenwert von A , so sind die zugehörigen Eigenräume V_{λ_1} und V_{λ_2} nach Lemma 7.2.3 orthogonal zueinander (d. h. jeder Vektor aus V_{λ_1} ist zu jedem Vektor aus V_{λ_2} orthogonal). Die Abbildung B läßt den Teilraum V_{λ_1} invariant, d. h. es gilt $BV_{\lambda_1} \subseteq V_{\lambda_1}$: Für alle $y \in V_{\lambda_1}$ gilt

$$AB y = BA y = B \lambda_1 y = \lambda_1 B y.$$

$B y$ ist also ebenfalls ein Eigenvektor von A zum Eigenwert λ_1 und liegt somit wieder in V_{λ_1} . Damit ist B eine selbstadjungierte Abbildung auf dem Vektorraum V_{λ_1} und hat nach Theorem 7.2.4 eine ONB aus Eigenvektoren für V_{λ_1} . Auf V_{λ_1} gibt es demnach eine *gemeinsame* ONB aus Eigenvektoren von A und B . Diese Überlegungen gelten natürlich für alle Eigenräume $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ zu den verschiedenen Eigenwerten $\lambda_1, \lambda_2, \dots, \lambda_k$ von A . Auf jedem dieser paarweise orthogonalen Räumen gibt es eine gemeinsame ONB aus Eigenvektoren von A und B . Zusammengenommen erhalten wir ein Orthonormalsystem aus gemeinsamen Eigenvektoren. Wir müssen uns nur noch überlegen, daß dieses System für V erzeugend ist, um die behauptete gemeinsame ONB nachgewiesen zu haben. Das liegt an Theorem 7.2.4, nach dem es eine ONB aus Eigenvektoren von A für V gibt. Gruppieren wir diese nach den verschiedenen Eigenwerten $\lambda_1, \lambda_2, \dots, \lambda_k$, so erzeugen sie jeweils die Räume $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$. Daher ist der Aufspann dieser Räume und damit auch der Aufspann des gemeinsamen Orthonormalsystems erzeugend für V . \square

7.4 Normale Abbildungen

7.4.1 Definition Eine lineare Abbildung T auf einem Vektorraum heißt *normal*, wenn T mit T^* vertauscht.

Natürlich ist jede selbstadjungierte Abbildung normal, die Umkehrung gilt aber nicht. Die Abbildung

$$U := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}.$$

z. B. ist normal, denn $UU^* = \mathbb{1} = U^*U$, aber $U \neq U^*$.

7.4.2 Korollar Für jede normale Abbildung auf einem komplexen Vektorraum V gibt es eine ONB aus Eigenvektoren von T .

Beweis. Da T mit T^* vertauscht, gilt das auch für die selbstadjungierten Abbildungen

$$R := \frac{1}{2}(T + T^*) \quad \text{und} \quad I := \frac{1}{2i}(T - T^*).$$

Nach Theorem 7.3.3 gibt es für V eine gemeinsame Eigenvektorbasis $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ für R und I zu den resp. Eigenwerten $\lambda_1, \dots, \lambda_n$ bzw. $\gamma_1, \dots, \gamma_n$. Aus

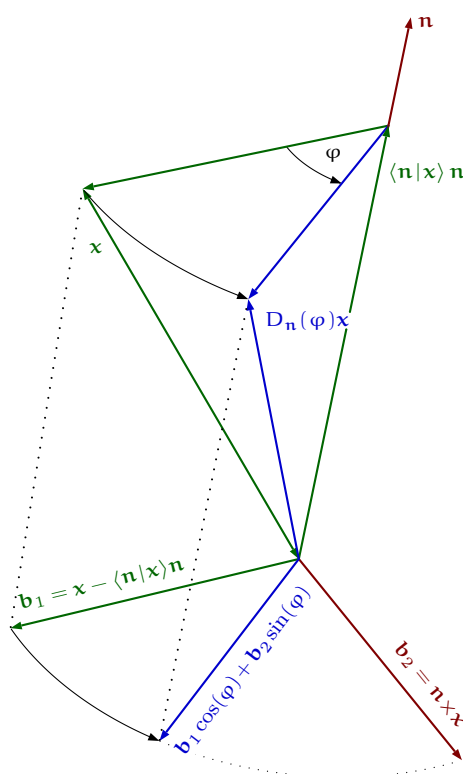
$$T = R + iI$$

folgt $T\mathbf{b}_k = R\mathbf{b}_k + iI\mathbf{b}_k = (\lambda_k + i\gamma_k)\mathbf{b}_k$. \mathcal{B} ist daher auch eine ONB aus Eigenvektoren für T (zu den komplexen Eigenwerten $\lambda_1 + i\gamma_1, \dots, \lambda_n + i\gamma_n$). \square

Die wichtigsten normalen Abbildungen sind neben den selbstadjungierten die *unitären* Abbildungen, auf reellen Vektorräumen auch als *orthogonale* Abbildungen bezeichnet (Definition 6.10.4). Für unitäre Abbildungen U existiert die Inverse und diese ist durch die Adjungierte U^* gegeben (vergl. Korollar 6.10.3). Als normale Abbildung hat U eine ONB aus Eigenvektoren. Ist U durch eine Matrix gegeben, dann stehen in ihren Spalten die Bilder $\mathbf{u}_1, \dots, \mathbf{u}_n$ der kanonischen Basisvektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$. Im Hinblick auf die Winkeltreue von U ist klar, dass $\mathcal{B} := \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ wieder eine ONB für V bestimmt. Das lässt zwei Interpretationen für die Wirkungsweise von U zu. Zunächst die *passive*, die U als eine Koordinatentransformation von Darstellungen bzgl. der Basis \mathcal{B} in die kanonische Basis ansieht. Oder die *aktive*, für die wir allerdings reelle Vektorräume voraussetzen, in der U eine winkel- und längentreue Abbildung ist, die auf den Basisvektoren \mathbf{e}_k wie $U\mathbf{e}_k = \mathbf{u}_k$ wirkt, die also ein Achsendreibein $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ (im Falle $\dim V = 3$) wieder in ein solches $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ überführt. Wenn letzteres ein Rechtssystem ist ($\det(U) = 1$), handelt es sich um eine *Drehung*. Ist es ein Linkssystem ($\det(U) = -1$), dann ist die Drehung mit einer *Spiegelung* (einer reellen selbstadjungierten linearen Abbildung S mit der Eigenschaft $S^2 = \mathbb{1}$) kombiniert.

7.4.3 A Zeigen Sie mit Hilfe der folgenden Skizze, daß sich eine Drehung um den Winkel φ , mit einer Drehachse, die durch den normierten Richtungsvektor $\mathbf{n} \in \mathbb{R}^3$ gegeben ist, durch die Drehmatrix $D_{\mathbf{n}}(\varphi)$ beschreiben läßt:

$$D_{\mathbf{n}}(\varphi) = P_{\mathbf{n}} + (\mathbb{1} - P_{\mathbf{n}}) \cos(\varphi) + R_{\mathbf{n}} \sin(\varphi). \quad (7.11)$$



Dabei ist $P_{\mathbf{n}}$ die eindimensionale Projektion auf die Richtung \mathbf{n} gemäß (7.8) und $R_{\mathbf{n}}$ die Matrix, die zum Kreuzprodukt mit \mathbf{n} gehört: $R_{\mathbf{n}} \mathbf{x} := \mathbf{n} \times \mathbf{x}$.

i) Zeigen Sie dafür zunächst

$$R_{\mathbf{n}} = \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}. \quad (7.12)$$

ii) Zeigen Sie $\|\mathbf{b}_1\| = \|\mathbf{b}_2\|$, $\mathbf{b}_1 \perp \mathbf{b}_2$, $\mathbf{b}_1 \perp \mathbf{n}$ und $\mathbf{b}_2 \perp \mathbf{n}$. Folgern Sie daraus, daß $\mathbf{b}_1 \cos(\varphi) + \mathbf{b}_2 \sin(\varphi)$ der um den Winkel φ in Richtung \mathbf{b}_2 gedrehte Vektor \mathbf{b}_1 ist.

iii) Entnehmen Sie der Skizze

$$D_{\mathbf{n}}(\varphi) \mathbf{x} = P_{\mathbf{n}} \mathbf{x} + \mathbf{b}_1 \cos(\varphi) + \mathbf{b}_2 \sin(\varphi)$$

und folgern Sie daraus (7.11).

iv) Bestimmen Sie für die kanonischen Basisvektoren \mathbf{e}_1 , \mathbf{e}_2 und \mathbf{e}_3 die Drehungen $D_{\mathbf{e}_1}(\varphi)$, $D_{\mathbf{e}_2}(\varphi)$ und $D_{\mathbf{e}_3}(\varphi)$.

- v) Bestimmen Sie $D_{\mathbf{n}}(\frac{\pi}{3})$ für $\mathbf{n} := \frac{1}{3} [1, 2, 2]^t$. Drehen Sie damit den Vektor $\mathbf{x} := [2, 3, 6]^t$. Welche Länge hat $D_{\mathbf{n}}(\frac{\pi}{3})\mathbf{x}$?

7.4.4 A Zeigen Sie: Das Spektrum einer unitären Matrix besteht aus Zahlen des komplexen Einheitskreises (daher auch der Name *unitär*).

7.4.5 A Wir haben oben eine reelle unitäre Matrix U mit positiver Determinante als Drehung bezeichnet. Im \mathbb{R}^3 sollte U also eine Drehung mit Drehachse \mathbf{n} und Drehwinkel α sein.

Machen Sie sich dafür folgende Punkte klar:

- i) Das Spektrum von U ist durch $\{1, \lambda, \bar{\lambda}\}$ gegeben, mit einem λ auf dem Einheitskreis von \mathbb{C} . Die zugehörige Eigenvektorbasis sei $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$. Dann ist \mathbf{b}_1 eine natürliche Wahl für die Richtung der Drehachse. Dafür mache man sich klar, daß $\mathbf{b}_1 \in \mathbb{R}^3$ gewählt werden kann (während \mathbf{b}_2 und \mathbf{b}_3 normalerweise in \mathbb{C}^3 liegen). Zeigen Sie, daß immer $\det(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = i$ erreicht werden kann, falls $U \neq U^*$ gilt.
- ii) Sei $\mathbf{b}_2 = \mathbf{e} + i\mathbf{f}$, mit Vektoren \mathbf{e} und \mathbf{f} aus \mathbb{R}^3 . Zeigen Sie, daß dann $\mathbf{b}_3 = \mathbf{e} - i\mathbf{f}$ wählbar ist. Verwenden Sie $\mathbf{b}_2 \perp \mathbf{b}_3$ um $\mathbf{e} \perp \mathbf{f}$ und $\|\mathbf{e}\| = \|\mathbf{f}\| = \frac{1}{2}\sqrt{2}$ nachzuweisen. Zeigen Sie, daß $\{\mathbf{b}_1, \mathbf{f}, \mathbf{e}\}$ positiv orientiert ist.
- iii) Zeigen Sie schließlich

$$U\mathbf{e} = \cos(\alpha)\mathbf{f} + \sin(\alpha)\mathbf{e}, \quad U\mathbf{f} = -\sin(\alpha)\mathbf{f} + \cos(\alpha)\mathbf{e}.$$

Dabei ist α durch $\lambda = e^{i\alpha} = \cos(\alpha) + i\sin(\alpha)$ definiert (natürlich nur bis auf uninteressante Vielfache von 2π).

Bestimmen Sie $U_{\mathcal{B}}$ für $\mathcal{B} := \{\mathbf{b}_1, \mathbf{f}, \mathbf{e}\}$, nachdem Sie \mathbf{e} und \mathbf{f} normiert haben.

- iv) Führen Sie das Verfahren für die folgenden beiden unitären Matrizen durch.

$$U := \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{bmatrix}, \quad V := \frac{1}{3} \begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ -2 & 2 & 1 \end{bmatrix}.$$

- v) Zeigen Sie, daß es sich bei den sechs Permutationsmatrizen

$$P_0 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_1 := \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad P_2 := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$

$$P_3 := \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_4 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad P_5 := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

um Drehungen bzw. um Spiegelungen ($P = P^* = P^{-1}$ und $\det(P) = -1$) handelt. Bestimmen Sie alle relevanten Größen, also Eigenvektorbasen, Drehachsen und Drehwinkel, bzw. Spiegelebenen.

7.4.6 A

- i) Zeigen Sie, daß R_n eine *schiefsymmetrische Abbildung* ist, daß also $R_n^* = -R_n$ gilt.
 ii) Zeigen Sie $R_{a \times b} = ba^t - ab^t$ und damit die GRASSMANN-Identität

$$a \times (b \times c) = \langle a | c \rangle b - \langle a | b \rangle c.$$

- iii) Folgern Sie aus ii)

$$(a \times b) \times (c \times d) = \det(a, b, d)c - \det(a, b, c)d.$$

- iv) Folgern Sie aus ii) die LAGRANGE-Identität

$$\langle a \times b | c \times d \rangle = \langle a | c \rangle \langle b | d \rangle - \langle a | d \rangle \langle b | c \rangle.$$

- v) Zeigen Sie $R_a^* R_a = \|a\|^2 \mathbb{1} - aa^t$ und damit $\|a \times b\| = \|a\| \|b\| \sin(\varphi)$. Dabei ist φ der von a und b eingeschlossene Winkel (vergl. auch (5.23)).
 vi) Folgern Sie aus iv) die GRASSMANN-PLÜCKER-Identität für 2×2 -Matrizen

$$\det(a, b) \det(c, d) + \det(b, c) \det(a, d) + \det(c, a) \det(b, d) = 0.$$

Verwenden Sie dabei die Einbettung $\varepsilon : a = [a_1, a_2]^t \mapsto \varepsilon(a) := [a_1, a_2, 0]^t$ des \mathbb{R}^2 in den \mathbb{R}^3 und die lineare Abbildung $J : [x_1, x_2, x_3]^t \mapsto [x_2, -x_1, x_3]$. Zeigen Sie zunächst $\varepsilon(a) \times \varepsilon(b) = \det(a, b) e_3$ und $\langle \varepsilon(a) | J\varepsilon(c) \rangle = \det(a, c)$. Benutzen Sie anschließend die LAGRANGE-Identität für die Vektoren $\varepsilon(a)$, $\varepsilon(b)$, $J\varepsilon(c)$ und $J\varepsilon(d)$.

Alternativ könnten Sie auch folgendermaßen vorgehen:

Zeigen Sie, daß $\lambda(a, b) := \det(a, b) \det(c, d) + \det(b, c) \det(a, d) + \det(c, a) \det(b, d)$ bei festgehaltenen Vektoren c und d eine alternierende Multilinearform des \mathbb{R}^2 definiert, daß also für eine geeignete Konstante k die Gleichung $\lambda(a, b) = k \cdot \det(a, b)$ gilt. Bestimmen Sie k .

8 Hauptachsentransformation

In diesem Kapitel soll die Eigenwerttheorie an dem geometrischen Beispiel der quadratischen Formen und der zugehörigen Quadriken angewandt werden. Im \mathbb{R}^2 handelt es sich dabei, wenn man von Sonderfällen absieht, im Wesentlichen darum, die genaue Lage von verschobenen und gedrehten Ellipsen, Hyperbeln, oder Parabeln zu bestimmen. Die Hauptarbeit wird darin bestehen, die Richtungen der Hauptachsen dieser Figuren zu finden. Sie werden sich als die Eigenvektoren einer selbstadjungierten Matrix herausstellen, die der Quadrik zugeordnet werden kann. Über die zugehörigen Eigenwerte läßt sich auf die Länge dieser Achsen schließen. Die Ideen, die wir hier entwickeln, lassen sich auch auf höhere Dimensionen ohne Änderung anwenden, allerdings erhöht sich dabei der Rechenaufwand erheblich. Wir streben hier jedoch keine systematische Klassifikation quadratischer Formen an, weshalb wir die Situation im \mathbb{R}^3 nur durch eine Übungsaufgabe streifen werden. Zunächst aber stellen wir die Quadriken in ihrer einfachsten Form als klassische Kegelschnitte vor, da erfahrungsgemäß nicht davon ausgegangen werden kann, daß diese aus der Schule bekannt sind.

8.1 Kegelschnitte

8.1.1 Ellipse Eine Ellipse ist der geometrische Ort aller Punkte \mathbf{p} , deren Abstände d_1 und d_2 von zwei Punkten \mathbf{f}_1 und \mathbf{f}_2 , den sogenannten *Brennpunkten*, die konstante Summe $d_1 + d_2 = 2a$ haben.

8.1.2 Satz Eine Ellipse E ist durch

$$E = \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\} \quad (8.1)$$

gegeben. Die Tangente \mathbf{t} bzw. die Normale \mathbf{n} im Punkt $\mathbf{p} := [x_0, y_0]^t \in E$ sind die Geraden

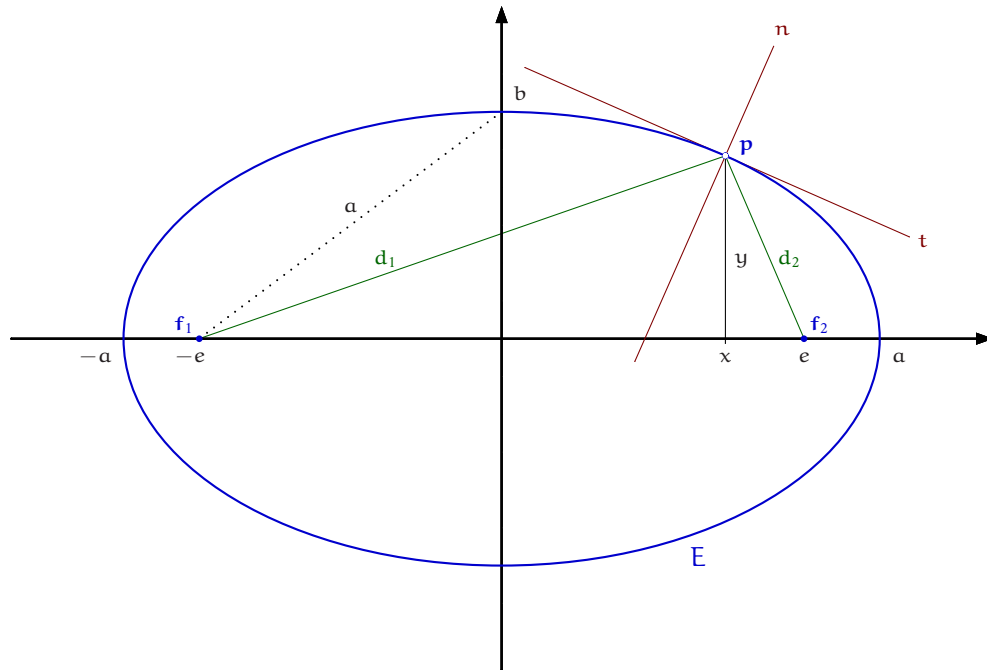
$$\begin{aligned} \mathbf{t} &= \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1 \right\} = \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{(x - x_0)x_0}{a^2} + \frac{(y - y_0)y_0}{b^2} = 0 \right\}, \\ \mathbf{n} &= \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{(x - x_0)y_0}{b^2} - \frac{(y - y_0)x_0}{a^2} = 0 \right\}. \end{aligned}$$

\mathbf{n} halbiert den Winkel $\angle \mathbf{f}_1 \mathbf{p} \mathbf{f}_2$. Dabei ist $b := \sqrt{a^2 - e^2}$ und $e := \frac{1}{2} \|\mathbf{f}_1 - \mathbf{f}_2\|$.

Beweis: Aus der Skizze entnimmt man

$$d_{1/2} = \sqrt{(e \pm x)^2 + y^2} = \sqrt{x^2 + y^2 + e^2 \pm 2xe}.$$

Es folgt $4xe = d_1^2 - d_2^2 = d_1^2 - (2a - d_1)^2 = 4ad_1 - 4a^2$, also $d_1 = a + \frac{e}{a}x$. Damit erhalten wir nacheinander



$$d_1^2 = x^2 + y^2 + e^2 + 2ex = \frac{e^2}{a^2}x^2 + 2ex + a^2,$$

$$x^2\left(1 - \frac{e^2}{a^2}\right) + y^2 = x^2\frac{b^2}{a^2} + y^2 = a^2 - e^2 = b^2.$$

Teilen wir diese Gleichung durch b^2 , so erhalten wir die Ellipsengleichung (8.1).

Umgekehrt ergibt sich aus (8.1) auch die definierende Eigenschaft $d_1 + d_2 = 2a$ der Ellipse. Zunächst haben wir $y^2 = b^2 - \frac{b^2}{a^2}x^2$, woraus bereits $|x| \leq a$ zu erkennen ist. Damit erhalten wir

$$d_{1/2}^2 = \left(1 - \frac{b^2}{a^2}\right)x^2 + b^2 + e^2 \pm 2xe = \frac{e^2}{a^2}x^2 + a^2 \pm 2xe = \left(a \pm \frac{e}{a}x\right)^2.$$

Aus $|x| \leq a$ folgt $\frac{e}{a}|x| \leq e \leq a$, so daß $d_{1/2} = a \pm \frac{e}{a}x$ und daher $d_1 + d_2 = 2a$ gilt.

Um zu zeigen, daß t wirklich die Tangente in $\mathbf{p} = [x_0, y_0]^t \in E$ ist, nehmen wir an, es gäbe einen weiteren Ellipsenpunkt $\mathbf{p}_1 := [x_1, y_1]^t$ auf t (in der Skizze sind die Koordinaten x und y von \mathbf{p} durch x_0 und y_0 zu ersetzen, um eine Unterscheidung zu den Punkten $[x, y]^t$ auf t zu erhalten). Dann haben wir

$$\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1, \quad \frac{x_0x_1}{a^2} + \frac{y_0y_1}{b^2} = 1, \quad \frac{x_1^2}{a^2} + \frac{y_1^2}{b^2} = 1.$$

Das hat $\frac{(x_0-x_1)^2}{a^2} + \frac{(y_0-y_1)^2}{b^2} = 0$ zur Folge, wie man leicht durch Ausmultiplizieren bestätigt. Da das nur für $x_0 = x_1$ und $y_0 = y_1$ zu erfüllen ist, kann es über \mathbf{p}_0 hinaus keinen weiteren

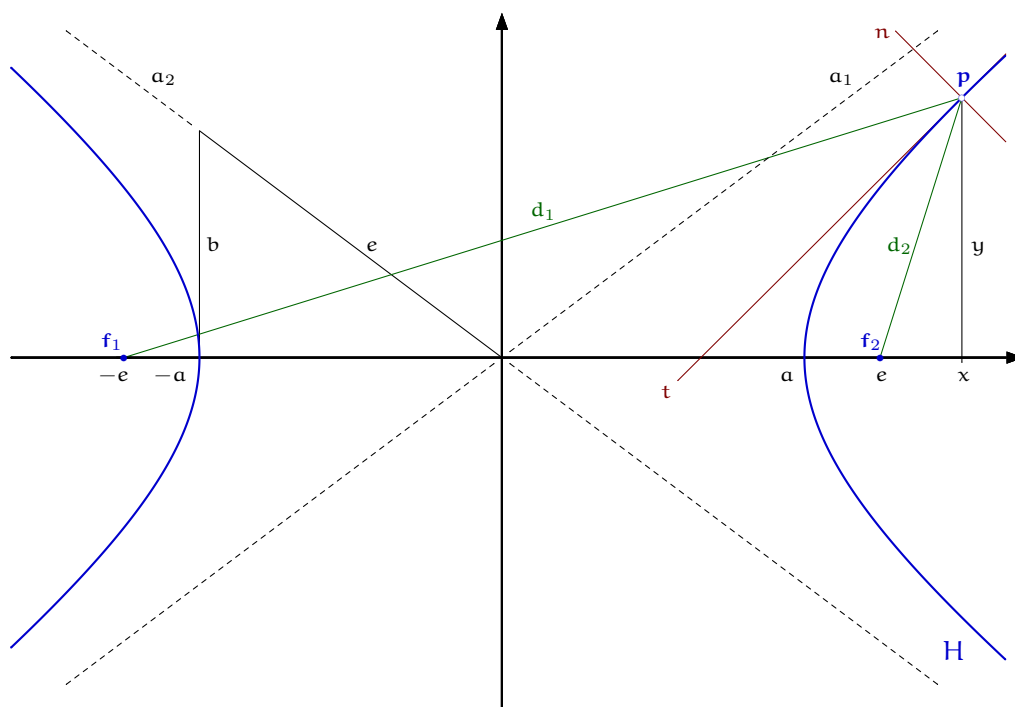
gemeinsamen Punkt von E und t geben. Das ist genau die Eigenschaft, die die Tangente in $p_0 \in E$ auszeichnet.

Die Gerade n schneidet die x -Achse offensichtlich bei $x_n = \frac{e^2}{a^2} x_0$. Sie teilt die Strecke $\|f_1 - f_2\|$ im Verhältnis

$$\frac{e - \frac{e^2}{a^2} x_0}{e + \frac{e^2}{a^2} x_0} = \frac{a - \frac{e}{a} x_0}{a + \frac{e}{a} x_0} = \frac{d_2}{d_1},$$

d. h., im Verhältnis der beim Winkel $\angle f_1 p f_2$ anliegenden Seiten d_2 und d_1 . Daher halbiert n diesen Winkel (vergl. Übung 8.1.6). Daß es sich bei n um die Normale im Punkt p handelt, ist eine einfache Übungsaufgabe. \square

8.1.3 Hyperbel Eine Hyperbel ist der geometrische Ort aller Punkte p , deren Abstände d_1 und d_2 von zwei Punkten f_1 und f_2 , den *Brennpunkten*, die konstante Differenz $|d_1 - d_2| = 2a$ haben.



8.1.4 Satz Eine Hyperbel H ist durch

$$H = \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \right\} \quad (8.2)$$

gegeben. Sie hat die Asymptoten $\alpha_{1/2}(x) := \pm \frac{b}{a} x$. Die Tangente t und die Normale n im Punkt $p := [x_0, y_0]^t \in H$ sind die Geraden

$$t = \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{xx_0}{a^2} - \frac{yy_0}{b^2} = 1 \right\} = \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{(x - x_0)x_0}{a^2} - \frac{(y - y_0)y_0}{b^2} = 0 \right\},$$

$$n = \left\{ [x, y]^t \in \mathbb{R}^2 \mid \frac{(x - x_0)y_0}{b^2} + \frac{(y - y_0)x_0}{a^2} = 0 \right\}.$$

t halbiert den Winkel $\angle \mathbf{f}_1 \mathbf{p} \mathbf{f}_2$. Dabei ist $b := \sqrt{e^2 - a^2}$ und $e := \frac{1}{2} \|\mathbf{f}_1 - \mathbf{f}_2\|$.

Der Beweis verläuft weitgehend wie der für die Ellipse.

Beweis. Aus der Skizze entnimmt man

$$d_{1/2} = \sqrt{(x \pm e)^2 + y^2} = \sqrt{x^2 + y^2 + e^2 \pm 2xe}.$$

Es folgt $4ex = d_1^2 - d_2^2 = d_1^2 - (d_1 - 2a)^2 = 4ad_1 - 4a^2$, also $d_1 = \frac{e}{a}x + a$ und $d_2 = \frac{e}{a}x - a$ ($d_2 \geq \frac{e}{a}a - a = e - a \geq 0$ für $x \geq a$). Für $x \leq -a$ ist $d_1 = -\frac{e}{a}x - a$ und $d_2 = -\frac{e}{a}x + a$.

Damit erhalten wir nacheinander

$$\begin{aligned} d_1^2 &= x^2 + y^2 + e^2 + 2ex = a^2 + 2ex + \frac{e^2}{a^2}x^2, \\ x^2 \left(1 - \frac{e^2}{a^2}\right) + y^2 &= -\frac{b^2}{a^2}x^2 + y^2 = a^2 - e^2 = -b^2. \end{aligned}$$

Division durch $-b^2$ ergibt die Hyperbelgleichung (8.2).

Umgekehrt ergibt sich aus (8.2) auch die definierende Eigenschaft $|d_1 - d_2| = 2a$ der Hyperbel. Zunächst haben wir $y^2 = \frac{b^2}{a^2}x^2 - b^2$, woraus bereits $|x| \geq a$ abzulesen ist. Damit erhalten wir

$$d_{1/2}^2 = \left(1 + \frac{b^2}{a^2}\right)x^2 - b^2 + e^2 \pm 2xe = \frac{e^2}{a^2}x^2 + a^2 \pm 2xe = \left(\frac{e}{a}x \pm a\right)^2.$$

Es folgt $d_{1/2} = \left|\frac{e}{a}x \pm a\right|$ und $d_1 d_2 = \left|\frac{e^2}{a^2}x^2 - a^2\right| = \frac{e^2}{a^2}x^2 - a^2$, denn $x^2 \geq a^2$. Das bedeutet

$$(d_1 - d_2)^2 = d_1^2 + d_2^2 - 2d_1 d_2 = 2\frac{e^2}{a^2}x^2 + 2a^2 - 2\left(\frac{e^2}{a^2}x^2 - a^2\right) = 4a^2.$$

Die Wurzel aus beiden Seiten der Gleichung zeigt $|d_1 - d_2| = 2a$.

Die Tangenteneigenschaft von t : Wir nehmen an, daß $\mathbf{p}_1 := [x_1, y_1]^t$ ein gemeinsamer Punkt von t und H ist, verschieden vom Tangentenpunkt $\mathbf{p} = [x_0, y_0]^t \in H \cap t$ (in der Skizze sind die Koordinaten x und y von \mathbf{p} wieder durch x_0 und y_0 zu ersetzen). Dann haben wir

$$\frac{x_0^2}{a^2} - \frac{y_0^2}{b^2} = 1, \quad \frac{x_0 x_1}{a^2} - \frac{y_0 y_1}{b^2} = 1, \quad \frac{x_1^2}{a^2} - \frac{y_1^2}{b^2} = 1.$$

Daraus erhalten wir

$$\frac{(x_0 - x_1)^2}{a^2} - \frac{(y_0 - y_1)^2}{b^2} = 0 \quad \text{und} \quad \frac{(x_0 - x_1)x_0}{a^2} - \frac{(y_0 - y_1)y_0}{b^2} = 0,$$

was durch Ausmultiplizieren leicht bestätigt werden kann. Das bedeutet

$$\frac{(y_0 - y_1)^2}{(x_0 - x_1)^2} = \frac{b^2}{a^2} \quad \text{und} \quad \frac{(y_0 - y_1)^2 y_0^2}{(x_0 - x_1)^2 x_0^2} = \frac{b^4}{a^4},$$

woraus $y_0^2 = \frac{b^2}{a^2} x_0^2$ folgt. Mit der Hyperbelgleichung $1 = \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2} = \frac{x_0^2}{a^2} - \frac{x_0^2 b^2}{a^2 b^2} = 0$ finden wir einen Widerspruch. Also ist \mathbf{p} der einzige gemeinsame Punkt von t und H .

Die Gerade t schneidet die x -Achse offensichtlich bei $x_t = \frac{a^2}{x_0}$. Sie teilt die Strecke $\|\mathbf{f}_1 - \mathbf{f}_2\|$ im Verhältnis

$$\frac{e - \frac{a^2}{x_0}}{\frac{a^2}{x_0} + e} = \frac{\frac{e}{a} x_0 - a}{\frac{e}{a} x_0 + a} = \frac{d_2}{d_1},$$

d. h., im Verhältnis der beim Winkel $\angle \mathbf{f}_1 \mathbf{p} \mathbf{f}_2$ anliegenden Seiten d_2 und d_1 . Daher halbiert t diesen Winkel (vergl. Übung 8.1.6). Daß es sich bei n um die Normale im Punkt \mathbf{p} handelt, ist eine einfache Übungsaufgabe.

Die Asymptoten der Hyperbel sind die Ursprungsgeraden mit den Gleichungen $a_{1/2}(x) = \pm \frac{b}{a} x$. Um das einzusehen, bestimmen wir die Funktionsgleichungen h_{\pm} der Hyperbel, indem wir (8.2) nach y auflösen:

$$h_{\pm}(x) = \pm \frac{b}{a} \sqrt{x^2 - a^2}.$$

h_{\pm} sind die Hyperbeläste oberhalb bzw. unterhalb der x -Achse. Wir zeigen nur den Fall $a_1(x) - h_+(x) \xrightarrow{x \rightarrow \infty} 0$. Für $x > 0$ erhalten wir, wenn wir uns an das dritte Binom erinnern,

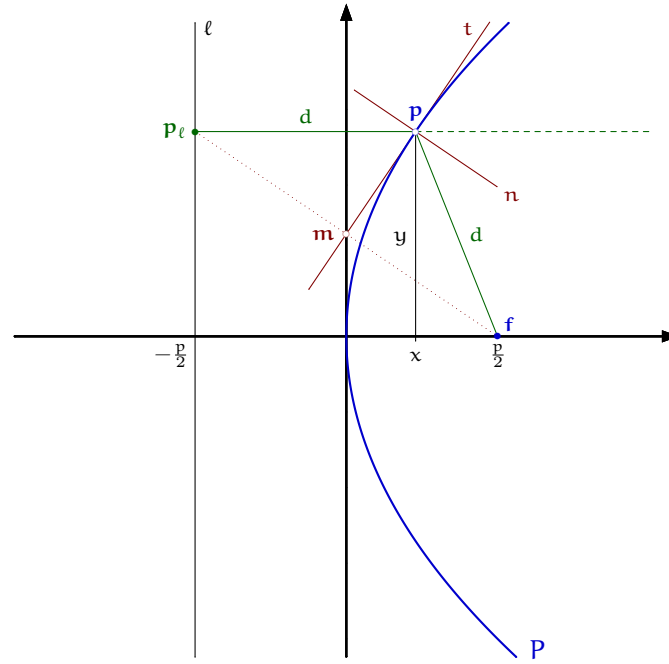
$$\begin{aligned} 0 \leq a_1(x) - h_+(x) &= \frac{b}{a} x - \frac{b}{a} |x| \sqrt{1 - \frac{a^2}{x^2}} = \frac{b}{a} x \left(1 - \sqrt{1 - \frac{a^2}{x^2}}\right) \\ &= \frac{b}{a} x \frac{\left(1 - \sqrt{1 - \frac{a^2}{x^2}}\right) \left(1 + \sqrt{1 - \frac{a^2}{x^2}}\right)}{\left(1 + \sqrt{1 - \frac{a^2}{x^2}}\right)} = \frac{b}{a} x \frac{\frac{a^2}{x^2}}{\left(1 + \sqrt{1 - \frac{a^2}{x^2}}\right)} \\ &\leq \frac{b}{a} x \frac{\frac{a^2}{x^2}}{1} = \frac{ab}{x} \xrightarrow{x \rightarrow \infty} 0. \end{aligned}$$

Für $x \rightarrow -\infty$ müssen wir

$$0 \leq h_-(x) - a_1(x) = -\frac{b}{a} |x| \sqrt{1 - \frac{a^2}{x^2}} - \frac{b}{a} x = \frac{b}{a} |x| \left(1 - \sqrt{1 - \frac{a^2}{x^2}}\right)$$

abschätzen, denn wegen $x < 0$ ist $-x = |x|$. Das verläuft ab jetzt genau so, wie wir es oben vorgeführt haben. \square

8.1.5 Parabel Eine Parabel P ist der geometrische Ort aller Punkte \mathbf{p} , für die der Abstand d zu einem festen Punkt \mathbf{f} , dem *Brennpunkt*, und zu einer Geraden ℓ , der sogenannten *Leitgeraden*, gleich groß ist.



Mit p bezeichnen wir den Abstand von f zu ℓ . Wir legen den Koordinatenursprung in die Mitte zwischen f und ℓ und die x -Achse senkrecht zu ℓ . Der Skizze entnehmen wir:

$$y^2 + \left(\frac{p}{2} - x\right)^2 = \left(\frac{p}{2} + x\right)^2, \text{ also } y^2 + x^2 - px + \frac{p^2}{4} = x^2 + px + \frac{p^2}{4}, \text{ d. h.:}$$

$$P = \{[x, y]^t \in \mathbb{R}^2 \mid y^2 = 2px\}, \quad (8.3)$$

$$t = \{[x, y]^t \in \mathbb{R}^2 \mid yy_0 = p(x + x_0)\} = \{[x, y]^t \in \mathbb{R}^2 \mid y_0(y - y_0) - p(x - x_0) = 0\},$$

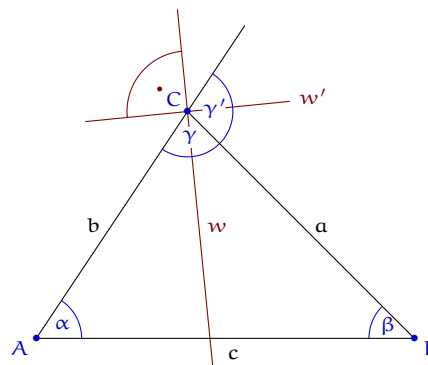
$$n = \{[x, y]^t \in \mathbb{R}^2 \mid p(y - y_0) + y_0(x - x_0) = 0\}.$$

t ist die Tangente an P in einem Parabelpunkt $\mathbf{p} = [x_0, y_0]^t$. Sie ist die Winkelhalbierende im Dreieck $\mathbf{f} \mathbf{p} \mathbf{p}_\ell$. Wird $p < 0$ gewählt, so entsteht eine nach links geöffnete Parabel. Wir haben nur noch die Tangenteneigenschaft von t zu zeigen. Dazu sei $\mathbf{p}_1 := [x_1, y_1]^t$ ein Punkt auf $t \cap H$, der von $\mathbf{p} := [x_0, y_0]^t$ verschieden ist (in der Skizze ist wieder x durch x_0 und y durch y_0 zu ersetzen). Wir haben also $y_0^2 = 2px_0$, $y_1^2 = 2px_1$ und $y_1y_0 = p(x_1 + x_0)$. Daraus folgt $(y_0 - y_1)^2 = y_0^2 - 2px_0 + y_1^2 - 2px_1 + 2px_0 + 2px_1 - 2y_1y_0 = 0$, also $y_0 = y_1$ und dann auch $x_0 = x_1$. Damit kann es nur einen gemeinsamen Punkt von P und t geben.

Der Mittelpunkt $\mathbf{m} = \frac{1}{2} \left(\left[-\frac{p}{2}, y_0\right]^t + \left[\frac{p}{2}, 0\right]^t \right) = \left[0, \frac{y_0}{2}\right]^t$ zwischen \mathbf{p}_ℓ und \mathbf{f} liegt auf der Tangente, denn er erfüllt die Tangentengleichung: $\frac{y_0}{2} y_0 = px_0$. Da das Dreieck $\mathbf{f} \mathbf{p} \mathbf{p}_\ell$ laut Parabeldefinition gleichschenkelig ist, handelt es sich bei t um die Winkelhalbierende in diesem Dreieck. Dessen Nebenwinkel hat die Normale n in \mathbf{p} als Winkelhalbierende. Diese Eigenschaft ist der Grund dafür, \mathbf{f} als *Brennpunkt* von P zu bezeichnen. Denken wir uns nämlich die Parabelinnenseite aus spiegelndem Material gefertigt, dann wird jeder parallel zur x -Achse einfallende Lichtstrahl in den Punkt \mathbf{f} reflektiert.

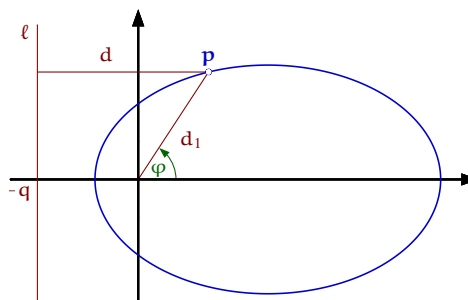
8.1.6 A Zwei elementare Eigenschaften der Winkelhalbierenden in einem Dreieck. Zeigen Sie:

- Eine Gerade w durch C ist genau dann die Winkelhalbierende, wenn sie die gegenüberliegende Seite c im Verhältnis $a : b$ der anliegenden Seiten a und b teilt.
- Die zu w orthogonale Gerade w' durch C ist die sogenannte *äußere* Winkelhalbierende. Sie halbiert den *Nebenwinkel* $\gamma' := \pi - \gamma$.



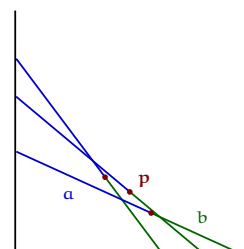
8.1.7 A Es gibt für Kegelschnitte einheitliche Definitionen, von denen wir zwei untersuchen wollen, nämlich die Konstruktion über eine Leitgerade und die Polardarstellung. Zeigen Sie:

- Ein Kegelschnitt K ist der geometrische Ort aller Punkte p , für die das Verhältnis $\frac{d_1}{d}$ aus dem Abstand d_1 von einem festen Punkt, dem Brennpunkt, und dem Abstand d von einer Geraden ℓ , der Leitgeraden, eine feste Zahl $\varepsilon > 0$ ist. Für $0 < \varepsilon < 1$ handelt es sich um eine Ellipse, für $\varepsilon = 1$ um eine Parabel und für $\varepsilon > 1$ um eine Hyperbel.
- $K = \left\{ \frac{p}{1 - \varepsilon \cos(\varphi)} e^{i\varphi} \mid 0 \leq \varphi < 2\pi \right\}$ ist die Polardarstellung eines Kegelschnitts. $\varepsilon = 0$ ergibt einen Kreis, $0 < \varepsilon < 1$ eine Ellipse, $\varepsilon = 1$ eine Parabel und $\varepsilon > 1$ eine Hyperbel.



8.1.8 A

Eine Leiter, die an eine Hauswand gelehnt ist, gleitet zu Boden. Welche Bahn beschreibt dabei ein beliebiger, fest gewählter Punkt p auf dieser Leiter?



8.2 Quadratische Formen

Der allgemeinste quadratische Ausdruck im Bereich gewöhnlicher reeller Funktionen ist durch $f(x) = ax^2 + bx + c$ gegeben, mit Konstanten $a \neq 0$, b und c . Die richtige Verallgemeinerung auf den \mathbb{R}^2 , \mathbb{R}^3 oder \mathbb{R}^n ist eine *quadratische Form*:

8.2.1 Definition Eine quadratische Form ist eine Funktion $Q: \mathbb{R}^n \rightarrow \mathbb{R}$ der folgenden Art

$$Q(x) = \langle x | Ax \rangle + \langle b | x \rangle + c \quad (8.4)$$

mit einer $n \times n$ -Matrix A , einem Vektor $b \in \mathbb{R}^n$ und einer Zahl $c \in \mathbb{R}$. Die zu Q gehörende Quadrik \mathcal{Q} ist die Menge

$$\mathcal{Q} := \{ x \in \mathbb{R}^n \mid Q(x) = 0 \}. \quad (8.5)$$

Die Matrix A der quadratischen Form Q kann o. B. d. A. selbstadjungiert gewählt werden, wie die folgende Rechnung zeigt:

$$\begin{aligned} Q(x) &= \langle x | Ax \rangle + \langle b | x \rangle + c = \langle A^* x | x \rangle + \langle b | x \rangle + c \\ &= \langle x | A^* x \rangle + \langle b | x \rangle + c \\ &= \frac{1}{2} (\langle x | Ax \rangle + \langle x | A^* x \rangle) + \langle b | x \rangle + c \\ &= \langle x | \frac{1}{2} (A + A^*) x \rangle + \langle b | x \rangle + c. \end{aligned}$$

Die quadratische Form ändert sich also nicht, wenn wir A durch die selbstadjungierte Matrix $\frac{1}{2} (A + A^*)$ ersetzen. Wir gehen daher für die folgenden Überlegungen von einer selbstadjungierten Matrix A aus und haben damit zur Untersuchung der Form die Eigenwerttheorie zur Verfügung. Bevor wir aber allgemeine quadratische Formen untersuchen, stellen wir den Zusammenhang mit den Kegelschnitten her. Die Gleichungen (8.1) und (8.2) für die Ellipse bzw. Hyperbel lassen sich leicht als Quadrik einer quadratischen Form identifizieren:

$$1 = \frac{x^2}{a^2} \pm \frac{y^2}{b^2} = \left\langle \begin{bmatrix} x \\ y \end{bmatrix} \mid \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \pm \frac{1}{b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle.$$

Die zugehörige Matrix

$$\tilde{A} := \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \pm \frac{1}{b^2} \end{bmatrix}$$

liegt offensichtlich bereits in Diagonalf orm vor. Aus den Eigenwerten $\frac{1}{a^2}$ und $\frac{1}{b^2}$ bzw. $-\frac{1}{b^2}$ lassen sich die Längen der Hauptachsen leicht ablesen. b und c sind Null. Eine um m_1 in x -Richtung und um m_2 in y -Richtung verschobene Ellipse bzw. Hyperbel hat die Gleichung

$$1 = \frac{(x - m_1)^2}{a^2} \pm \frac{(y - m_2)^2}{b^2} = \langle x - m | \tilde{A} (x - m) \rangle$$

$$= \langle \mathbf{x} | \tilde{A} \mathbf{x} \rangle - 2 \langle \tilde{A} \mathbf{m} | \mathbf{x} \rangle + \langle \mathbf{m} | \tilde{A} \mathbf{m} \rangle, \quad \mathbf{m} := \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}.$$

Offensichtlich hat also der Vektor \mathbf{b} in einer allgemeinen quadratischen Form etwas mit der Verschiebung der Quadrik zu tun, auch wenn aus ihm der Verschiebungsvektor meistens nicht direkt ablesbar ist. Hier ist $\mathbf{b} = -2\tilde{A}\mathbf{m}$. Wenn wir vor der Verschiebung $T(\mathbf{x}) := \mathbf{x} + \mathbf{m}$ noch eine Koordinatentransformation $B = [\mathbf{b}_1, \mathbf{b}_2]$ durchführen (mit der Orthonormalbasis $\mathcal{B} := \{\mathbf{b}_1, \mathbf{b}_2\}$), die eine Drehung der Quadrik bewirkt, dann haben wir die allgemeinste Form einer Ellipsen- bzw. Hyperbelgleichung gefunden:

$$1 = \langle B^*(\mathbf{x} - \mathbf{m}) | \tilde{A} B^*(\mathbf{x} - \mathbf{m}) \rangle = \langle \mathbf{x} - \mathbf{m} | B \tilde{A} B^*(\mathbf{x} - \mathbf{m}) \rangle.$$

Die Matrix $A := B \tilde{A} B^*$ wird nun i. Allg. nicht mehr diagonal sein (wohl aber selbstadjungiert – nachrechnen).

Auch wenn sich auf diese Weise nicht alle Quadriken gewinnen lassen, denn von Sonderfällen wie Doppelkegel, oder Doppelpunkten abgesehen, bilden die Parabeln eine Ausnahme, läßt sich aus der vorgestellten Konstruktion doch eine Methode erkennen, die auch zur Untersuchung allgemeiner Quadriken verwendet werden kann. Wir haben die Konstruktion nur schrittweise rückgängig zu machen.

Im ersten Schritt werden die Eigenwerte $\lambda_1, \dots, \lambda_n$ und die zugehörige ONB $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ aus Eigenvektoren der Matrix A in (8.4) bestimmt. Die Koordinatentransformation $\mathbf{x} = B\mathbf{x}_{\mathcal{B}}$ von der Koordinatendarstellung bzgl. \mathcal{B} in die kanonische Basis \mathcal{E} (gemäß unserer Vereinbarung auf Seite 155, daß sich eine Koordinatendarstellung \mathbf{x} ohne Index immer auf die kanonische Basis $\mathcal{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bezieht, führt die Gleichung $Q(\mathbf{x}) = 0$ aus (8.4) in

$$\begin{aligned} 0 &= \langle B\mathbf{x}_{\mathcal{B}} | AB\mathbf{x}_{\mathcal{B}} \rangle + \langle \mathbf{b} | B\mathbf{x}_{\mathcal{B}} \rangle + c \\ &= \langle \mathbf{x}_{\mathcal{B}} | B^*AB\mathbf{x}_{\mathcal{B}} \rangle + \langle B^*\mathbf{b} | \mathbf{x}_{\mathcal{B}} \rangle + c \\ &= \langle \mathbf{x}_{\mathcal{B}} | A_{\mathcal{B}}\mathbf{x}_{\mathcal{B}} \rangle + \langle \mathbf{b}_{\mathcal{B}} | \mathbf{x}_{\mathcal{B}} \rangle + c \end{aligned}$$

über. Dabei ist $A_{\mathcal{B}} := B^*AB = \text{diag}[\lambda_1, \dots, \lambda_n]$ eine Diagonalmatrix, deren Einträge die Eigenwerte $\lambda_1, \dots, \lambda_n$ von A sind. $\mathbf{b}_{\mathcal{B}} := B^*\mathbf{b}$ ist die Koordinatendarstellung von \mathbf{b} bzgl. \mathcal{B} .

Im zweiten Schritt wird obige Gleichung in Koordinatenform ausgeschrieben.

$$0 = \lambda_1 \tilde{x}_1^2 + \dots + \lambda_n \tilde{x}_n^2 + \tilde{b}_1 \tilde{x}_1 + \dots + \tilde{b}_n \tilde{x}_n + c.$$

Für die Eigenwerte $\lambda_i \neq 0$ kann $\lambda_i \tilde{x}_i^2 + \tilde{b}_i \tilde{x}_i$ durch quadratische Ergänzung in ein Binom verwandelt werden.

$$\lambda_i \tilde{x}_i^2 + \tilde{b}_i \tilde{x}_i = \lambda_i \left(\tilde{x}_i^2 + 2 \frac{\tilde{b}_i}{2\lambda_i} \tilde{x}_i + \frac{\tilde{b}_i^2}{4\lambda_i^2} \right) - \frac{\tilde{b}_i^2}{4\lambda_i} = \lambda_i (\tilde{x}_i - \tilde{m}_i)^2 - \lambda_i \tilde{m}_i^2,$$

mit $\tilde{m}_i := \frac{\tilde{b}_i}{2\lambda_i}$. Sind etwa im \mathbb{R}^2 alle Eigenwerte von Null verschieden, so erhält man die transformierte Gleichung der Quadrik in der Form einer verschobenen Ellipse oder Hyperbel:

$$\tilde{c} = \frac{(\tilde{x}_1 - \tilde{m}_1)^2}{1/\lambda_1} + \frac{(\tilde{x}_2 - \tilde{m}_2)^2}{1/\lambda_2}. \quad (8.6)$$

Dabei ist $\tilde{c} := \frac{\tilde{b}_1^2}{4\lambda_1} + \frac{\tilde{b}_2^2}{4\lambda_2} - c$. Jetzt erkennt man eine Ellipse daran, daß entweder alle Eigenwerte positiv oder alle negativ sind (sind sie z. B. alle positiv, dann muß $\tilde{c} > 0$ gelten, damit die Menge \mathcal{Q} nicht leer ist), während bei einer Hyperbel die Eigenwerte unterschiedliche Vorzeichen haben. In jedem Fall ergeben sich die Längen der Hauptachsen als $\sqrt{\frac{|\tilde{c}|}{|\lambda_i|}}$. Der Verschiebungsvektor der Quadrik hat in der Basis \mathcal{B} offensichtlich die Koordinaten \tilde{m}_1 und \tilde{m}_2 . Mit B wird er in die kanonische Basis zurück transformiert:

$$\mathbf{m} = B \begin{bmatrix} \tilde{m}_1 \\ \tilde{m}_2 \end{bmatrix}.$$

Ist einer der Eigenwerte Null, sagen wir λ_1 , dann läßt sich für \tilde{x}_1 keine quadratische Ergänzung durchführen. Statt (8.6) erhalten wir

$$0 = (\tilde{x}_2 - \tilde{m}_2)^2 + \frac{\tilde{b}_1}{\lambda_2} \tilde{x}_1 + \frac{c}{\lambda_2} - \tilde{m}_2^2.$$

Ist $\tilde{b}_1 \neq 0$, dann läßt sich das mit einem geeigneten \tilde{m}_1 in die Gestalt

$$(\tilde{x}_2 - \tilde{m}_2)^2 = -\frac{\tilde{b}_1}{\lambda_2} (\tilde{x}_1 - \tilde{m}_1)$$

bringen. Das ist eine verschobene Parabel. Man überlege sich, welche Möglichkeiten sich noch für $\tilde{b}_1 = 0$ ergeben.

Die gerade vorgestellten Ideen lassen sich auch auf höhere Dimensionen, aber vornehmlich auf die Dimension 3 erweitern. Da es dann drei verschiedene Eigenwerte geben kann, ergibt sich auch eine größere Vielfalt an möglichen Quadriken. Die einfachste Erweiterung der zweidimensionalen Situation entsteht, wenn alle Eigenwerte $\neq 0$ sind. Statt (8.6) erhalten wir dann

$$\tilde{c} = \frac{(\tilde{x}_1 - \tilde{m}_1)^2}{1/\lambda_1} + \frac{(\tilde{x}_2 - \tilde{m}_2)^2}{1/\lambda_2} + \frac{(\tilde{x}_3 - \tilde{m}_3)^2}{1/\lambda_3}. \quad (8.7)$$

Ist $\tilde{c} > 0$ und sind alle Eigenwerte positiv, dann handelt es sich um ein sog. *Ellipsoid*, d. h. um eine geschlossene Fläche um den Ursprung im \mathbb{R}^3 , deren Schnitt mit jeder Ursprungsebene eine Ellipse bildet. Sind dagegen ein oder zwei Eigenwerte negativ, so handelt es sich um ein sog. *Hyperboloid*, das im ersten Fall einschalig (wie etwa die Kühltürme von Kraftwerken) und im zweiten zweischalig ist. Weitere Formen und Sonderfälle entstehen, wenn Eigenwerte 0 sein dürfen, oder wenn $\tilde{c} = 0$ sein sollte. Darunter sind verschiedene Formen von Paraboloiden (einer Fläche, deren Schnitt mit einer oder mehreren Ebenen eine Parabel liefert), Doppelkegeln, elliptische, hyperbolische und parabolische Zylinder (also Zylinder, deren Querschnitte senkrecht zur Zylinderachse Ellipsen, Hyperbeln oder Parabeln darstellen), aber auch ausgeartete Figuren wie parallele Ebenen oder gar nur ein Paar Punkte.

8.3 Beispiele für Quadriken

8.3.1 Eine Ellipse

$$\begin{aligned} Q_1(\mathbf{x}) &:= \frac{1}{36}(8x^2 - 4xy + 5y^2) + \frac{1}{\sqrt{5}}(2x - 6y) + 12 \\ &= \frac{1}{36} \left\langle \mathbf{x} \left| \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix} \mathbf{x} \right\rangle + \frac{1}{\sqrt{5}} \left\langle \begin{bmatrix} 2 \\ -6 \end{bmatrix} \middle| \mathbf{x} \right\rangle + 12 \end{aligned}$$

Wir bestimmen zunächst die Eigenwerte der Matrix $A := \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$

$$\det \begin{bmatrix} 8 - \lambda & -2 \\ -2 & 5 - \lambda \end{bmatrix} = (8 - \lambda)(5 - \lambda) - 4 = \lambda^2 - 13\lambda + 36 = (\lambda - 4)(\lambda - 9)$$

zu $\lambda_1 := 4$ und $\lambda_2 := 9$. Der Eigenvektor \mathbf{b}_1 zu λ_1 ist eine Lösung des Gleichungssystems

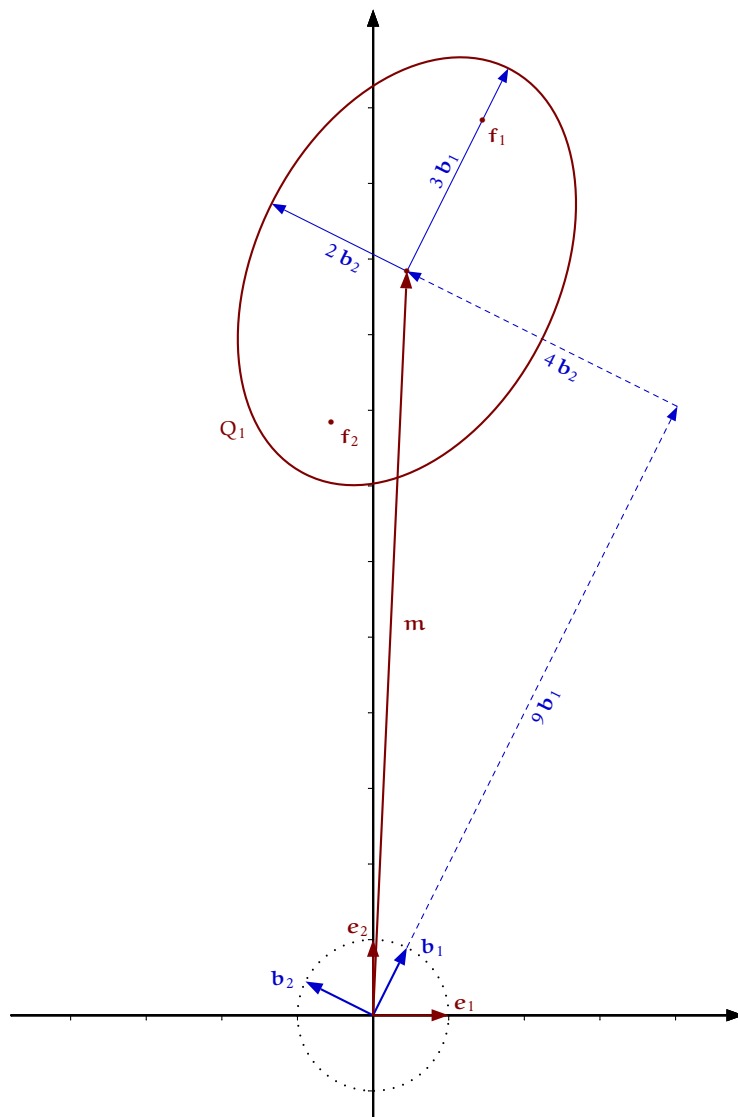
$$\begin{bmatrix} 8 - \lambda_1 & -2 \\ -2 & 5 - \lambda_1 \end{bmatrix} \mathbf{b}_1 = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \mathbf{b}_1 = \mathbf{0},$$

mit der offensichtlichen Lösung $\mathbf{b}_1 \sim \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Den auf 1 normierten Vektor $\frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ wählen wir als Eigenvektor zu λ_1 . Da A selbstadjungiert ist, gibt es eine ONB aus Eigenvektoren. Für den zweiten normierten Eigenvektor \mathbf{b}_2 gibt es daher nur noch die beiden Möglichkeiten $\mathbf{b}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ oder $\mathbf{b}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Wir wählen die erste, denn dann wird die ONB

$$\mathcal{B} := \{ \mathbf{b}_1, \mathbf{b}_2 \} = \left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$$

ein Rechtssystem. Mit der Matrix $B := \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$ führen wir die Koordinatentransformation $\mathbf{x} = B\mathbf{x}_{\mathcal{B}}$ durch:

$$\begin{aligned} \tilde{Q}_1(\mathbf{x}_{\mathcal{B}}) &= \frac{1}{36} \left\langle B\mathbf{x}_{\mathcal{B}} \left| \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix} B\mathbf{x}_{\mathcal{B}} \right\rangle + \frac{1}{\sqrt{5}} \left\langle \begin{bmatrix} 2 \\ -6 \end{bmatrix} \middle| B\mathbf{x}_{\mathcal{B}} \right\rangle + 12 \\ &= \frac{1}{36} \left\langle \mathbf{x}_{\mathcal{B}} \left| B^* \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix} B\mathbf{x}_{\mathcal{B}} \right\rangle + \frac{1}{5} \left\langle \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -6 \end{bmatrix} \middle| \mathbf{x}_{\mathcal{B}} \right\rangle + 12 \\ &= \frac{1}{36} \left\langle \mathbf{x}_{\mathcal{B}} \left| \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \mathbf{x}_{\mathcal{B}} \right\rangle - \left\langle \begin{bmatrix} 2 \\ 2 \end{bmatrix} \middle| \mathbf{x}_{\mathcal{B}} \right\rangle + 12 \\ &= \frac{1}{9} \tilde{x}^2 + \frac{1}{4} \tilde{y}^2 - 2\tilde{x} - 2\tilde{y} + 12 \\ &= \frac{1}{9} (\tilde{x}^2 - 18\tilde{x} + 9^2) + \frac{1}{4} (\tilde{y}^2 - 8\tilde{y} + 4^2) + 12 - 9 - 4 \\ &= \frac{(\tilde{x} - 9)^2}{9} + \frac{(\tilde{y} - 4)^2}{4} - 1. \end{aligned}$$



In der Basis \mathcal{B} handelt es sich bei unserer Quadrik $Q_1 := \{ \mathbf{x} \in \mathbb{R}^2 \mid Q_1(\mathbf{x}) = 0 \}$ also um eine um 9 Einheiten in Richtung \mathbf{b}_1 und um 4 Einheiten in Richtung \mathbf{b}_2 verschobene Ellipse mit der großen Halbachse $a = 3$ und der kleinen $b = 2$, sowie mit $e = \sqrt{3^2 - 2^2} = \sqrt{5}$. Der Mittelpunkt \mathbf{m} liegt bei

$$\mathbf{m} = 9\mathbf{b}_1 + 4\mathbf{b}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 9 - 8 \\ 18 + 4 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 22 \end{bmatrix}$$

und die beiden Brennpunkte bei $\mathbf{f}_{1/2} = \left[\frac{1}{\sqrt{5}} \pm 1, \frac{22}{\sqrt{5}} \pm 2 \right]^t$ (erhalten aus $\mathbf{m} \pm \sqrt{5} \mathbf{b}_1$).

8.3.2 A Untersuchen Sie die quadratische Form

$$Q_2(\mathbf{x}) := 180x^2 - 84xy + 145y^2 - \sqrt{13}(84x + 334y) + 1261.$$

8.3.3 Eine Hyperbel

$$\begin{aligned}
 Q_3(\mathbf{x}) &:= -221x^2 + 174xy + 11y^2 + \sqrt{10} (324x - 28y) - 1640 \\
 &= \left\langle \mathbf{x} \left| \begin{bmatrix} -221 & 87 \\ 87 & 11 \end{bmatrix} \mathbf{x} \right\rangle + \sqrt{10} \left\langle \begin{bmatrix} 324 \\ -28 \end{bmatrix} \middle| \mathbf{x} \right\rangle - 1640.
 \end{aligned}$$

Die Eigenwerte:

$$\begin{aligned}
 \det \begin{bmatrix} -221 - \lambda & 87 \\ 87 & 11 - \lambda \end{bmatrix} &= (11 - \lambda)(-221 - \lambda) - 87^2 = \lambda^2 + 210\lambda - 10000 \\
 &= (\lambda + 250)(\lambda - 40).
 \end{aligned}$$

Die Eigenwerte sind $\lambda_1 = 40$ und $\lambda_2 = -250$. Der Eigenvektor \mathbf{b}_1 zu λ_1 :

$$\begin{bmatrix} -261 & 87 \\ 87 & -29 \end{bmatrix} \mathbf{b}_1 = \begin{bmatrix} -3 \cdot 87 & 87 \\ 3 \cdot 29 & -29 \end{bmatrix} \mathbf{b}_1 = \mathbf{0}.$$

Eine Lösung ist offensichtlich $\mathbf{b}_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Damit ist der normierte Eigenvektor zu λ_2 z. B.

durch $\mathbf{b}_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ gegeben. Die Transformationsmatrix B ist daher

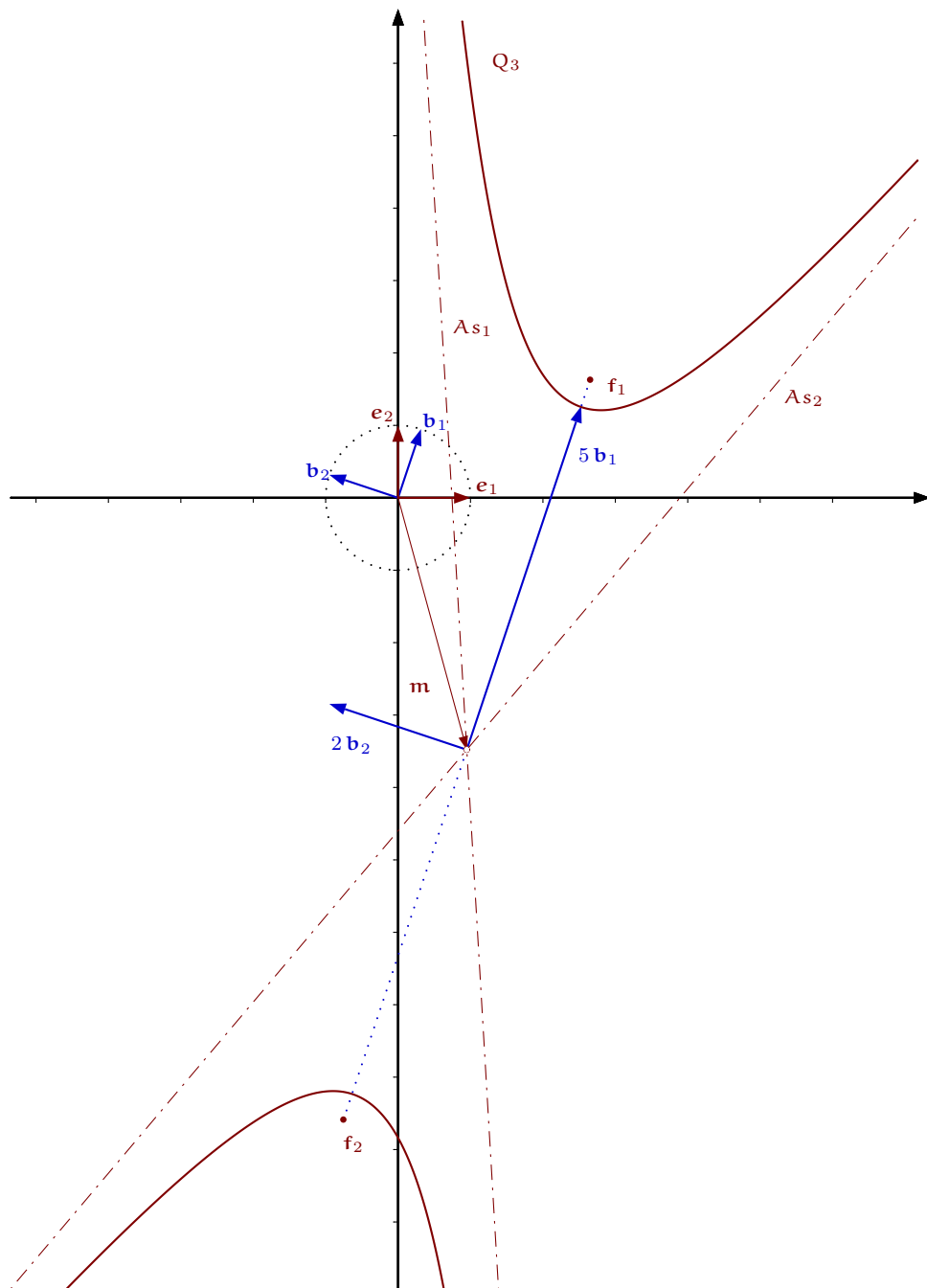
$$B := \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}.$$

Sie vermittelt die Koordinatentransformation $\mathbf{x} = B\mathbf{x}_B$:

$$\begin{aligned}
 \tilde{Q}_3(\mathbf{x}_B) &= \left\langle \mathbf{x}_B \left| B^* \begin{bmatrix} -221 & 87 \\ 87 & 11 \end{bmatrix} B \mathbf{x}_B \right\rangle + \left\langle \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 324 \\ -28 \end{bmatrix} \middle| \mathbf{x}_B \right\rangle - 1640 \\
 &= \left\langle \mathbf{x}_B \left| \begin{bmatrix} 40 & 0 \\ 0 & -250 \end{bmatrix}_B \mathbf{x}_B \right\rangle + \left\langle \begin{bmatrix} 240 \\ -1000 \end{bmatrix}_B \middle| \mathbf{x}_B \right\rangle - 1640 \\
 &= 40(\tilde{x}^2 + 6\tilde{x} + 3^2) - 250(\tilde{y}^2 + 4\tilde{y} + 2^2) - 360 + 1000 - 1640 \\
 &= 40(\tilde{x} + 3)^2 - 250(\tilde{y} + 2)^2 - 1000 \\
 &= 1000 \left[\frac{(\tilde{x} + 3)^2}{25} - \frac{(\tilde{y} + 2)^2}{4} - 1 \right].
 \end{aligned}$$

In der Basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ handelt es sich um eine Hyperbel, in \mathbf{b}_1 -Richtung um -3 Einheiten und in \mathbf{b}_2 -Richtung um -2 Einheiten verschoben. Die Halbachsen sind $a = 5$ und $b = 2$. Damit ist $e = \sqrt{5^2 + 2^2} = \sqrt{29}$. Der Mittelpunkt \mathbf{m} liegt bei $\mathbf{m} = -3\mathbf{b}_1 - 2\mathbf{b}_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -11 \end{bmatrix}$ und

die Brennpunkte bei $\mathbf{f}_{1/2} = \mathbf{m} \pm \sqrt{29} \mathbf{b}_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \pm \sqrt{29} \\ -11 \pm 3\sqrt{29} \end{bmatrix}$.

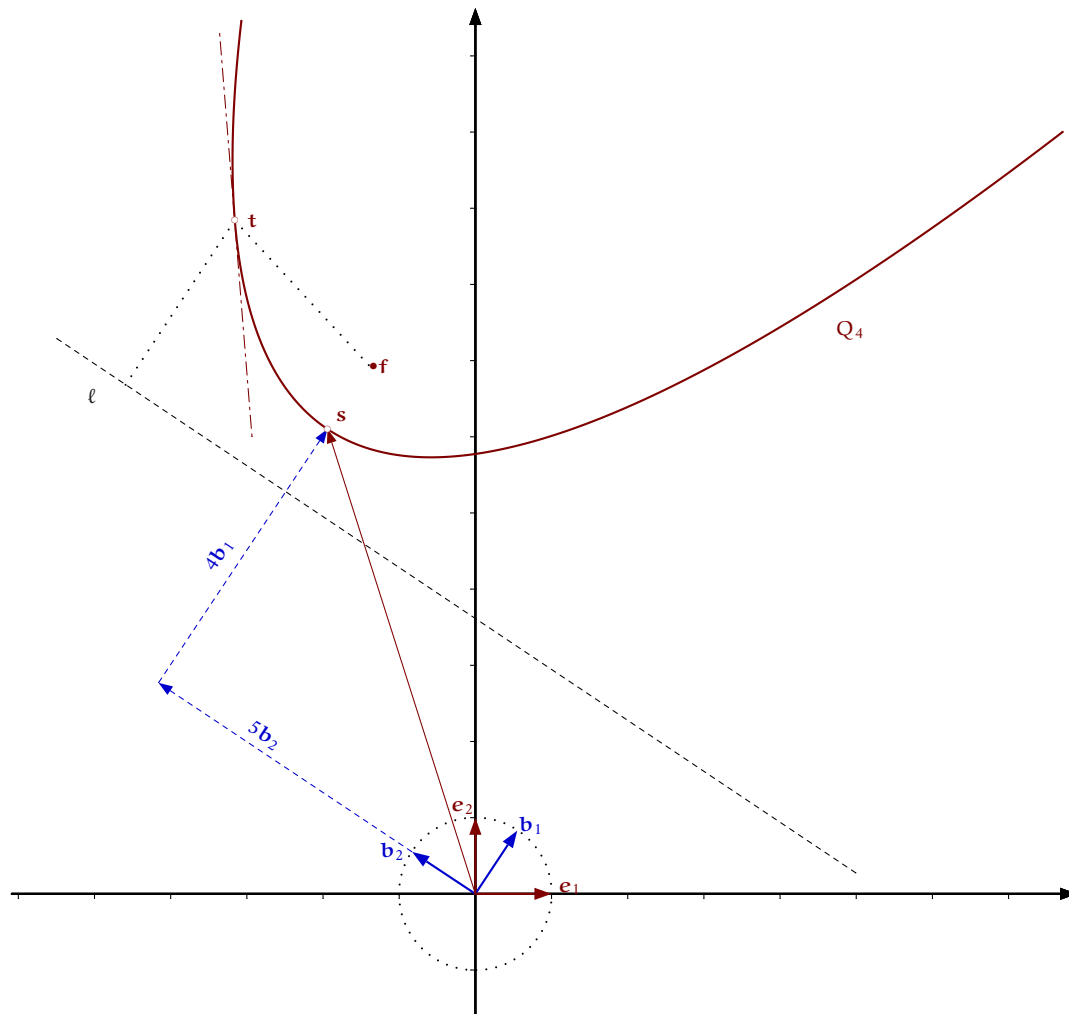


Die Asymptoten haben die Gleichungen $As_1 : 17x + y = 4\sqrt{10}$, $As_2 : 13x - 11y = 16\sqrt{10}$. Ein Richtungsvektor z. B. der Geraden As_1 ist durch $5\mathbf{b}_1 + 2\mathbf{b}_2 \sim \begin{bmatrix} -1 \\ 17 \end{bmatrix}$ gegeben. Daher können wir als Normalenvektor $\begin{bmatrix} 17 \\ 1 \end{bmatrix}$ nehmen. Die Geradengleichung lautet dann $17x + y = c$ mit

einem noch zu bestimmenden c . Der Punkt \mathbf{m} liegt auf beiden Asymptoten. Einsetzen ergibt $c = \frac{1}{\sqrt{10}} (3 \cdot 17 - 11) = \frac{40}{\sqrt{10}} = 4\sqrt{10}$. Dieselbe Rechnung mit dem Richtungsvektor $5\mathbf{b}_1 - 2\mathbf{b}_2$ liefert die Gleichung für $\mathcal{A}s_2$.

8.3.4 Eine Parabel

$$\begin{aligned} Q_4(\mathbf{x}) &:= 9x^2 - 12xy + 4y^2 + \sqrt{13}(22x - 32y) + 533 \\ &= \left\langle \mathbf{x} \left| \begin{bmatrix} 9 & -6 \\ -6 & 4 \end{bmatrix} \mathbf{x} \right\rangle + \sqrt{13} \left\langle \begin{bmatrix} 22 \\ -32 \end{bmatrix} \middle| \mathbf{x} \right\rangle + 533. \end{aligned}$$



Die Eigenwerte der Matrix sind $\lambda_1 = 0$ und $\lambda_2 = 13$, mit den zugehörigen Eigenvektoren $\mathbf{b}_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ und $\mathbf{b}_2 = \frac{1}{\sqrt{13}} \begin{bmatrix} -3 \\ 2 \end{bmatrix}$. Die Transformationsmatrix: $B = \frac{1}{\sqrt{13}} \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$.

$$\tilde{Q}_4(\mathbf{x}_B) = \left\langle \mathbf{x}_B \left| \begin{bmatrix} 0 & 0 \\ 0 & 13 \end{bmatrix}_B \mathbf{x}_B \right\rangle + \left\langle \begin{bmatrix} -52 \\ -130 \end{bmatrix}_B \middle| \mathbf{x}_B \right\rangle + 533$$

$$\begin{aligned}
&= 13\tilde{y}^2 - 52\tilde{x} - 130\tilde{y} + 533 \\
&= 13(\tilde{y}^2 - 10\tilde{y} + 5^2) - 52\tilde{x} + 208 \\
&= 13\left[(\tilde{y} - 5)^2 - 4(\tilde{x} - 4)\right].
\end{aligned}$$

Da der \tilde{x}^2 -Term fehlt, lässt sich die quadratische Ergänzung nur für \tilde{y} durchführen. Wir erhalten die um 5 Einheiten in \mathbf{b}_2 - und um 4 Einheiten in \mathbf{b}_1 -Richtung verschobene Parabel $\tilde{y}^2 = 4\tilde{x}$.

Damit ist $p = 2$. Der Scheitel liegt bei $\mathbf{s} = 5\mathbf{b}_2 + 4\mathbf{b}_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} -7 \\ 22 \end{bmatrix}$ und der Brennpunkt bei

$\mathbf{s} + \frac{p}{2} \mathbf{b}_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} -5 \\ 25 \end{bmatrix}$. Die Leitgerade ℓ geht durch den Punkt $\mathbf{s} - \frac{p}{2} \mathbf{b}_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} -9 \\ 19 \end{bmatrix}$ und hat den Normalenvektor \mathbf{b}_1 . Damit ist $\ell = \{ [x, y]^t \mid 2x + 3y = 3\sqrt{13} \}$.

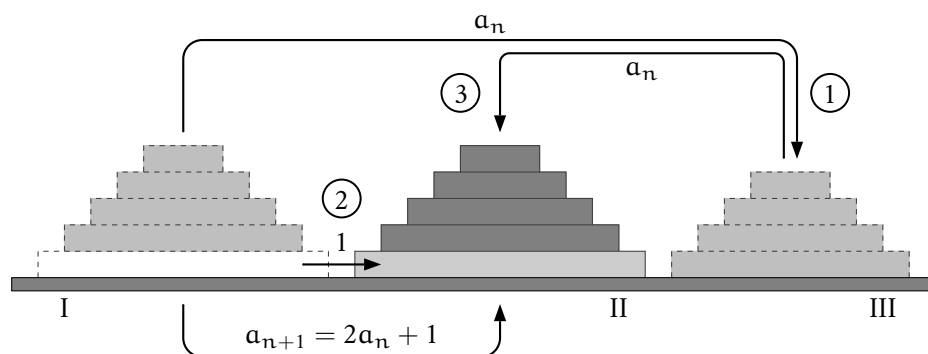
8.3.5 A Untersuchen Sie die quadratischen Formen

$$\begin{aligned}
Q_5(\mathbf{x}) &:= 17x^2 - 16xy + 17y^2 - 104x - 4y - 13, \\
Q_6(\mathbf{x}) &:= 25x^2 + 120xy + 144y^2 + 286x + 247y + 676.
\end{aligned}$$

8.3.6 A Untersuchen Sie die quadratische Form

$$Q_7(\mathbf{x}) := 43x^2 + 36y^2 + 43z^2 + 36xy - 36yz - 22zx - 120x + 288y - 264z + 936.$$

9 Rekurrenzgleichungen



Bei dem Spiel *Turm von Hanoi* geht es darum, einen Turm aus Steinen abnehmender Größe Stein für Stein von Position I in die Position II zu versetzen, ohne einen größeren auf einen kleineren Stein zu legen. Dabei steht der Platz III als Zwischenablage zur Verfügung. a_n sei die Anzahl der Züge, die nötig sind, um einen Turm aus n Steinen zu versetzen. Für einen Turm aus $n+1$ Steinen werden a_n Züge gebraucht, um die ersten n Steine auf Platz III zu transportieren, einen weiteren, um den letzten Stein nach II zu bringen und anschließend wieder a_n Züge, um die n Steine von Platz III in ihre endgültige Position II zu verschieben. Also ist

$$a_{n+1} = 2a_n + 1. \quad (9.1)$$

Mit der offensichtlichen Anfangsbedingung $a_1 = 1$ haben wir für das Problem eine sog. *Rekurrenzgleichung* (oder *Rekursionsgleichung*) aufgestellt. Das Merkmal einer solchen Gleichung ist, daß die Kenntnis von a_{n+1} auf die Kenntnis von a_n zurückgreift und die wiederum auf die von a_{n-1} usw., bis man bei der Anfangsbedingung a_1 (manchmal auch bei a_0) angekommen ist. Das Ziel ist eine explizite Formel für a_n . Eine mögliche Methode dafür, die bei einfachen Gleichungen erfolgreich sein kann, ist die Rückwärtsentwicklung:

$$\begin{aligned} a_n &= 2a_{n-1} + 1 = 2(2a_{n-2} + 1) + 1 = 4a_{n-2} + 2 + 1 \\ &= 4(2a_{n-3} + 1) + 2 + 1 = 2^3 a_{n-3} + 2^2 + 2^1 + 2^0 \\ &= 2^3(2a_{n-4} + 1) + 2^2 + 2^1 + 2^0 \\ &= 2^4 a_{n-4} + 2^3 + 2^2 + 2^1 + 2^0 \\ &= \dots \\ &= 2^{n-1} a_1 + 2^{n-2} + \dots + 2^2 + 2^1 + 2^0 = \sum_{k=0}^{n-1} 2^k. \end{aligned}$$

In diesem Fall hat sie zwar eine Formel in Form einer Summe geliefert, da die Summenbildung jedoch selbst ein rekursiver Vorgang ist, wird nur dann ein echter Fortschritt erzielt, wenn für

sie eine explizite Formel gefunden werden kann. Das ist meistens nicht der Fall. Glücklicherweise handelt es sich bei diesem Beispiel um die sog. *geometrische Reihe*, für die sehr wohl eine geschlossene Formel bekannt ist. Um sie zu finden setzt man $S_n := \sum_{k=0}^n q^k$ (im Beispiel ist natürlich $q = 2$). Dann ist

$$q \cdot S_n = q + q^2 + q^3 + \cdots + q^n + q^{n+1}.$$

Zieht man davon $S_n = 1 + q + q^2 + \cdots + q^n$ ab, so ergibt sich

$$(q - 1)S_n = q^{n+1} - 1$$

und daraus, für $q \neq 1$,

$$\sum_{k=0}^n q^k = \frac{q^{n+1} - 1}{q - 1}. \quad (9.2)$$

Im vorliegenden Beispiel erhalten wir

$$a_n = \sum_{k=0}^{n-1} 2^k = \frac{2^n - 1}{2 - 1} = 2^n - 1.$$

Demnach sind für einen Turm aus 5 Steinen, bei der vorgestellten Lösungsstrategie, 31 Züge nötig.

9.1 Lineare Rekurrenzgleichungen 1. Ordnung

Wir bringen die Gleichung (9.1) in die systematische Form einer linearen Rekurrenzgleichung 1. Ordnung

$$a_{n+1} - 2a_n = 1. \quad (9.3)$$

Im Hinblick auf unsere Lösungstheorie linearer Gleichungen lösen wir zunächst die homogene lineare Rekurrenzgleichung

$$a_{n+1} - 2a_n = 0 \quad (9.4)$$

durch den Ansatz

$$a_n^{(0)} := \lambda^n, \quad (9.5)$$

mit einem $\lambda \neq 0$, das noch zu bestimmen ist. Durch Einsetzen von (9.5) in (9.4) erhalten wir

$$\lambda^{n+1} - 2\lambda^n = \lambda^n(\lambda - 2) = 0.$$

Das bedeutet $\lambda = 2$, so daß die allgemeine homogene Lösung die Form

$$a_n^{(0)} = t 2^n$$

hat, mit einem noch beliebigen Faktor t . Dieser wird es uns erlauben unseren Ansatz für die allgemeine Lösung an die Anfangsbedingung $a_1 = 1$ anzupassen. Doch zunächst brauchen wir eine spezielle inhomogene Lösung $a_n^{(1)}$ von (9.3). Da die rechte Seite dieser Gleichung einfach

eine Konstante ist, versuchen wir den Ansatz einer konstanten Lösung $a_n^{(1)} = d = \text{konst.}$ Wenn sich d bestimmen läßt, ist dieser Ansatz gerechtfertigt. Einsetzen in (9.3) ergibt

$$d - 2d = 1,$$

oder $d = -1$. Damit lautet die allgemeine Lösung von (9.3):

$$a_n = t \cdot 2^n - 1.$$

Die freie Konstante t bestimmen wir aus der Anfangsbedingung $a_1 = 1$,

$$1 = 2t - 1,$$

zu $t = 1$. Daher ist die Lösung von (9.3), wie wir bereits gesehen haben, durch $a_n = 2^n - 1$ gegeben.

9.1.1 Konstante Inhomogenität

$$a_n - c a_{n-1} = d, \quad d : \text{konstant.} \quad (9.6)$$

Der übliche Ansatz $a_n^{(0)} = \lambda^n$ für die homogene Lösung führt auf $\lambda = c$ (deshalb wurde das Minuszeichen vor dem c in die Definition von Rekurrenzgleichungen 1. Ordnung mit aufgenommen, da die homogene Lösung sonst $(-c)^n$ wäre).

Für die inhomogene Lösung versuchen wir den Ansatz $a_n^{(1)} = k$, mit einer noch zu bestimmenden Konstanten k . Einsetzen in (9.6) ergibt die Gleichung $k - ck = d$, die nur für $c \neq 1$ nach k auflösbar ist. Wir untersuchen daher vorerst den Fall $c \neq 1$. Dann ist $k = \frac{d}{1-c}$. Damit lautet die allgemeine Lösung von (9.6) bisher

$$a_n = t c^n + \frac{d}{1-c}.$$

t bestimmen wir aus der Anfangsbedingung a_0 :

$$a_0 = t + \frac{d}{1-c}, \quad \text{also} \quad t = a_0 - \frac{d}{1-c}.$$

Das ergibt für den Fall $c \neq 1$ die endgültige Lösung:

$$a_n = a_0 c^n + \frac{d}{1-c} (1 - c^n). \quad (9.7)$$

Für $c = 1$: Der Versuch, die inhomogene Lösung durch eine Konstante zu bestimmen scheitert, wie wir oben gesehen haben. Wir versuchen $a_n^{(1)} = k \cdot n$. In (9.6) eingesetzt:

$$kn - k(n-1) = d,$$

woraus $k = d$ folgt. Damit haben wir durch $a_n^{(1)} = dn$ eine spezielle inhomogene Lösung gefunden. Zusammen mit der homogenen Lösung $a_n^{(0)} = 1$ (s.o.) lautet die allgemeine Lösung $a_n = t + dn$. Die Anfangsbedingung a_0 bestimmt t zu a_0 , so daß die Lösung für $c = 1$ einfach durch

$$a_n = a_0 + dn \quad (9.8)$$

gegeben ist.

9.1.2 Variable Inhomogenität

Wir untersuchen nur Gleichungen der Form

$$a_n - c a_{n-1} = d \cdot n. \quad (9.9)$$

Sei zunächst $c \neq 1$. Die homogene Lösung ist wieder $a_n^{(0)} = c^n$. Für die inhomogene machen wir den Ansatz $a_n^{(1)} = x \cdot n + y$ und versuchen die Unbekannten x und y zu bestimmen. Einsetzen in (9.9) ergibt

$$\begin{aligned} xn + y - cx(n-1) - cy &= dn, \\ x(1-c)n + cx + y(1-c) &= dn. \end{aligned}$$

Da das für alle $n \geq 0$ gelten muß, ergibt sich für $n = 0$

$$y = -\frac{cx}{1-c}.$$

Vergleich der Koeffizienten von n zeigt

$$x = \frac{d}{1-c} \quad \text{und damit} \quad y = -\frac{cd}{(1-c)^2}.$$

Daher ist

$$a_n = tc^n + \frac{d}{1-c} n - \frac{cd}{(1-c)^2}.$$

Berücksichtigen wir noch die Anfangsbedingung a_0 , so erhalten wir endgültig für $c \neq 1$:

$$a_n = a_0 c^n + \frac{cd}{(1-c)^2} (c^n - 1) + \frac{d}{1-c} n. \quad (9.10)$$

Für den Fall $c = 1$ führt der Ansatz $a_n^{(1)} = x \cdot n + y$ auf die unmögliche Bestimmungsgleichung $x = d$. Wir versuchen $a_n^{(1)} = x \cdot n^2 + y \cdot n + z$. Einsetzen in (9.9) ergibt

$$2xn + y - x = dn,$$

d. h. $x = \frac{d}{2}$ und $y = \frac{d}{2}$. Da z aus der Bestimmungsgleichung herausfällt, können wir es ohne Bedenken weglassen. Genau wie oben vorgeführt erhält man jetzt unschwer

$$a_n = a_0 + \frac{d}{2} n(n+1). \quad (9.11)$$

9.2 Lineare Rekurrenzgleichungen 2. Ordnung

Wir beginnen mit einem Beispiel, der *FIBONACCI-Folge*. Dabei handelt es sich um ein stark vereinfachtes Populationsmodell einer Kaninchenzucht (etwa aus dem Jahre 1200): Ausgehend von einem jungen Kaninchenpaar bringt jedes erwachsene Paar jährlich ein weiteres zur Welt. Erwachsen ist ein Kaninchenpaar bereits im zweiten Lebensjahr. Außerdem wird davon abgesehen, daß Paare auch sterben könnten. a_n soll die Anzahl der Paare nach n Jahren wiedergeben. Dann ist $a_0 = 1$ und $a_1 = 1$. Im zweiten Jahr hat das Kaninchenpaar Nachkommen, so daß

$a_2 = 2$ ist. Im nächsten Jahr hat wieder nur das erste Paar Nachkommen, denn der Nachwuchs ist noch nicht erwachsen: $a_3 = 3$. Im vierten Jahr sind die Nachkommen des zweiten Jahrs erwachsen und haben zusammen mit dem Elternpaar jeweils ein Kaninchenpaar d. h. a_2 Paare als Nachkommen. Zusammen mit den a_3 Paaren des dritten Jahrs sind inzwischen $a_4 = a_3 + a_2$ Paare vorhanden. Im $n + 1$ -ten Jahr sind die a_n Paare des vorausgehenden Jahrs und die a_{n-1} Nachkommen der in diesem Jahr erwachsen gewordenen Paare vorhanden. Damit haben wir die Rekursionsgleichung

$$a_{n+1} = a_n + a_{n-1}, \quad a_0 = 1, a_1 = 1, \quad (9.12)$$

der FIBONACCI-Folge gefunden. Wenn man alles auf eine Seite bringt und die Indizierung um eins verschiebt, erkennt man, daß es sich um eine homogene Rekurrenzgleichung handelt:

$$a_n - a_{n-1} - a_{n-2} = 0, \quad a_0 = 1, a_1 = 1. \quad (9.13)$$

Sie ist von zweiter Ordnung, da sie auf die zwei direkten Vorläufer a_{n-1} und a_{n-2} der Folge zurückgreift, um das aktuelle Folgenglied a_n zu berechnen. Die ersten 11 Glieder der Folge sind schnell bestimmt:

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89 \dots$$

Der systematische Lösungsansatz dieser homogenen Rekurrenzgleichung 2. Ordnung ist $a_n = \lambda^n$:

$$\lambda^n - \lambda^{n-1} - \lambda^{n-2} = \lambda^{n-2}(\lambda^2 - \lambda - 1) = 0. \quad (9.14)$$

$\lambda^2 - \lambda - 1 = 0$ hat die beiden Lösungen $\lambda_1 = \Phi$ und $\lambda_2 = -\frac{1}{\Phi}$, mit

$$\Phi := \frac{1}{2}(\sqrt{5} + 1) \approx 1.61803398875. \quad (9.15)$$

Es ist eine gute Übung nachzurechnen, daß die zweite Lösung $\lambda_2 = \frac{1}{2}(1 - \sqrt{5})$ tatsächlich mit $-\frac{1}{\Phi} = 1 - \Phi$ übereinstimmt. Die Zahl $\varphi := \frac{1}{\Phi} \approx 0.61803398875$ wird als *goldener Schnitt* bezeichnet. Die allgemeine Lösung der homogenen Gleichung (9.13) ist wegen deren Linearität durch eine Linearkombination der Lösungen Φ^n und $(-\frac{1}{\Phi})^n = (1 - \Phi)^n$ gegeben:

$$a_n = t_1 \Phi^n + t_2 (1 - \Phi)^n. \quad (9.16)$$

t_1 und t_2 werden aus den Anfangsbedingungen $a_0 = 1$ und $a_1 = 1$ bestimmt:

$$\begin{aligned} t_1 + t_2 &= 1, \\ t_1 \Phi + t_2 (1 - \Phi) &= 1. \end{aligned}$$

Dieses lineare Gleichungssystem führt auf $t_2 = 1 - t_1$ und das auf $1 = t_1 \Phi + (1 - t_1)(1 - \Phi) = t_1(2\Phi - 1) + 1 - \Phi = t_1 \sqrt{5} + 1 - \Phi$, was leicht nach t_1 aufgelöst werden kann:

$$t_1 = \frac{\Phi}{\sqrt{5}}, \quad t_2 = \frac{\sqrt{5}}{\sqrt{5}} - \frac{\Phi}{\sqrt{5}} = \frac{\frac{1}{2}(\sqrt{5} - 1)}{\sqrt{5}} = -\frac{1 - \Phi}{\sqrt{5}}.$$

In Gleichung (9.16) eingesetzt ergibt das die endgültige Lösung von (9.13),

$$a_n = \frac{1}{\sqrt{5}} (\Phi^{n+1} - (1 - \Phi)^{n+1}). \quad (9.17)$$

Da der zweite Summand $(1 - \Phi)^{n+1} = (-1)^{n+1} \Phi^{-n-1}$ rasch klein wird, kann (9.17) zur schnellen Berechnung großer FIBONACCI-Zahlen verwendet werden. Dazu bestimmt man nämlich nur

$$a_n \approx \frac{1}{\sqrt{5}} \Phi^{n+1}$$

und rundet zur nächsten ganzen Zahl auf bzw. ab. Schon für $n = 10$ und $n = 11$ ergibt sich

$$a_{10} \approx \frac{1}{\sqrt{5}} \Phi^{11} \approx 88.997, \quad a_{11} \approx 144.001,$$

d. h. $a_{10} = 89$ und $a_{11} = 144$. Für $n = 26$ kann man den Unterschied von $a_{26} \approx \frac{1}{\sqrt{5}} \Phi^{27}$ zu $a_{26} = 196418$ kaum noch darstellen. Für $a_{38} \approx \frac{1}{\sqrt{5}} \Phi^{39}$ zeigen die meisten Rechner nur das exakte Ergebnis 63245986 an.

Wir untersuchen nun die allgemeine Rekurrenzgleichung 2. Ordnung mit konstanter Inhomogenität:

$$a a_n + b a_{n-1} + c a_{n-2} = d, \quad a \neq 0. \quad (9.18)$$

Der Ansatz $a_n^{(0)} = \lambda^n$ für die homogene Lösung führt jetzt auf

$$a\lambda^n + b\lambda^{n-1} + c\lambda^{n-2} = \lambda^{n-2}(a\lambda^2 + b\lambda + c) = 0,$$

d. h. auf die gewöhnliche quadratische Gleichung

$$a\lambda^2 + b\lambda + c = 0, \quad (9.19)$$

die bekanntlich die Lösungen

$$\lambda_{1/2} = \frac{1}{2a} \left(-b \pm \sqrt{b^2 - 4ac} \right) \quad (9.20)$$

hat. Wir kümmern uns nicht darum, ob der Ausdruck unter der Wurzel positiv oder negativ ist (im letzteren Fall merken wir uns vor, die Lösung komplex aufzuschreiben). Allerdings müssen wir uns sehr wohl um den Sonderfall $b^2 = 4ac$ Gedanken machen, in dem (9.20) nur die eine Lösung $\lambda = -\frac{b}{2a}$ hat.

$\lambda_1 \neq \lambda_2$ (d. h. $b^2 \neq 4ac$): Die allgemeine homogene Lösung hat die Form

$$a_n^{(0)} = t_1 \lambda_1^n + t_2 \lambda_2^n. \quad (9.21)$$

Für die inhomogene Lösung machen wir, wie bisher immer, zunächst den Ansatz $a_n^{(1)} = x$, mit einer zu bestimmenden Konstanten x . Einsetzen in (9.18) ergibt $x = \frac{1}{a+b+c}$, falls $a+b+c \neq 0$ ist. Für $a+b+c = 0$ führt der Ansatz $a_n^{(1)} = x \cdot n$ zu $x = -\frac{d}{b+2c}$ (b kann wegen $b^2 \neq 4ac$ nicht mit $-2c$ übereinstimmen, so daß hier keine weitere Fallunterscheidung zu treffen

ist). Wie im Beispiel der FIBONACCI-Folge vorgeführt, läßt sich nun die allgemeine Lösung aus der homogenen und der inhomogenen zusammensetzen. Die Anfangsbedingungen a_0 und a_1 führen auf ein lineares Gleichungssystem für die beiden Unbekannten t_1 und t_2 .

$\lambda_1 = \lambda_2$ (d. h. $b^2 = 4ac$ und $\lambda = \lambda_1 = -\frac{b}{2a}$): Jetzt haben wir das Problem, daß uns ein Ansatz der Form (9.21) für die homogene Lösung zunächst nicht möglich ist, da uns nur eine Lösung λ^n der homogenen Rekurrenzgleichung zur Verfügung steht. Um aber einen Lösungsansatz a_n an die beiden Anfangsbedingungen a_0 und a_1 anzupassen, benötigen wir zwei Lösungen der homogenen Rekurrenzgleichung, damit wir sie, wie in (9.21), mit zwei Unbekannten t_1 und t_2 kombinieren können. Erst dann erhalten wir ein Gleichungssystem aus zwei linearen Gleichungen für t_1 und t_2 .

Wir zeigen nun, daß durch $n \cdot \lambda^n$ eine weitere Lösung der homogenen Rekurrenzgleichung gegeben ist. Einsetzen in $aa_n + ba_{n-1} + ca_{n-2}$ müßte Null ergeben. Dabei können wir natürlich $a\lambda^n + b\lambda^{n-1} + c\lambda^{n-2} = 0$ verwenden, denn λ wurde ja gerade so bestimmt, daß diese Gleichung erfüllt ist.

$$\begin{aligned} a n \lambda^n + b(n-1)\lambda^{n-1} + c(n-2)\lambda^{n-2} &= n(a\lambda^n + b\lambda^{n-1} + c\lambda^{n-2}) - b\lambda^{n-1} - 2c\lambda^{n-2} \\ &= -\lambda^{n-2}(b\lambda + c) = -\lambda^{n-2}\left(-\frac{b^2}{2a} + 2c\right) \\ &= -\lambda^{n-2}\left(-\frac{4ac}{2a} + 2c\right) = 0. \end{aligned}$$

Der Ansatz für die homogene Lösung lautet jetzt

$$a_n^{(0)} = (t_1 + t_2 n) \lambda^n, \quad \lambda = -\frac{b}{2a}. \quad (9.22)$$

Nun kann man mit der Bestimmung einer inhomogenen Lösung fortfahren, so wie es oben vorgestellt wurde.

9.2.1 A i) Zeigen Sie (mit Induktion), daß die maximale Anzahl a_n der Teile, in die man eine Ebene durch n Geraden zerlegen kann, der Rekursion $a_{n+1} = a_n + n + 1$, oder $a_n - a_{n-1} = n$, mit $a_0 = 1$ genügt. Geben Sie eine Lösung dieser Rekursionsgleichung an.

ii) Zeigen Sie die sogenannte *Mitternachtsformel*

$$x_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (9.23)$$

für die Lösungen x_1 und x_2 der quadratischen Gleichung $ax^2 + bx + c = 0$, falls die sogenannte *Diskriminante* $b^2 - 4ac$ nicht negativ ist. Leiten Sie daraus die *pq-Formel*

$$x_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} \quad (9.24)$$

für die Lösungen der quadratischen Gleichung in Normalform $x^2 + px + q = 0$ ab.

Nomenclature

$f', \frac{d}{dx}$	Ableitung	287
A^*	adjungierte Matrix	147
\Leftrightarrow	äquivalent	3
\sim	Äquivalenzrelation	22
B_k	BERNOULLI-Zahlen	332
$ z $	Betrag	78
$\text{im } f$	Bild, Wertebereich	25
$\binom{\alpha}{k}$	Binomialkoeffizient, verallgemeinerter	330
$\binom{n}{k}$	Binomialkoeffizient	15
D_f	Definitionsbereich	25
\det	Determinante	73
$A \setminus B$	Differenzmenge	10
\dim	Dimension	125
\oplus	direkte Summe	138
\mathbb{C}_1	Komplexer Einheitskreis	80
$\mathbb{1}$	Einheitsmatrix	138
U_ε	ε -Umgebung	230
e	EULER-Zahl	262
\exists	es existiert	5
$(x_n)_{n \in \mathbb{N}}, (x_n)$	Folge	213
\Rightarrow	folgt	2
$x \mapsto f(x)$	Funktion	25
\forall	für alle	5
$[y]$	GAUSS-Klammer	247
ggT	größter gemeinsamer Teiler	36
φ, Φ	goldener Schnitt	209
grad	Grad eines Polynoms	340
$\lim_{x \rightarrow a^\pm}$	Grenzwert, links-/rechtsseitiger	268
$\lim_{n \rightarrow \infty}, \xrightarrow{n \rightarrow \infty}$	Grenzwert	216
id	identische Funktion	26

i	imaginäre Einheit	76
$(a, b), [a, b]$	Intervalle	8
R^{-1}	inverse Relation	24
\mathcal{E}	kanonische Basis $\{e_1, \dots, e_n\}$	127
\ker	Kern	141
kgV	kleinstes gemeinsames Vielfaches	48
\leq	kleiner oder gleich	21
A^c	Komplement einer Menge	10
\mathbb{C}	komplexe Zahlen	77
\bar{z}	konjugiert komplexe Zahl	78
$\mathbf{x} \times \mathbf{y}$	Kreuzprodukt	68
δ_{ik}	KRONECKER-Symbol	132
\mathbb{K}	Körper	90
\emptyset	leere Menge	10
$\liminf, \varliminf_{n \rightarrow \infty}$	Limes Inferior	250
$\limsup, \varlimsup_{n \rightarrow \infty}$	Limes Superior	250
lh	lineare Hülle	124
$\ \cdot\ _p$	L^p -Norm	97
L^p	L^p -Raum	98
$M_n(\mathbb{K}), M_n$	$n \times n$ -Matrizen	143
\max, \min	Maximum, Minimum	222
$x = y \pmod p, =_p$	Gleichheit modulo p	40
$\binom{n}{k_1, k_2, \dots, k_r}$	Multinomialkoeffizient	53
\mathbb{N}_n	natürliche Zahlen $\leq n$	30
\mathbb{N}, \mathbb{N}_0	natürliche Zahlen	12
$\neg A, \bar{A}$	Negation	2
$\ \mathbf{x}\ $	Norm, Länge	65
o. B. d. A.	ohne Beschränkung der Allgemeinheit	8
$V^\perp, \mathbf{x}^\perp$	orthogonaler Teilraum	139
\mathcal{S}_n	Permutationen, symmetrische Gruppe	30
\mathcal{P}_n	Polynome vom grad $\leq n$	350
\mathbb{P}	Primzahlen	9
$A \times B$	Produktmenge	10
\prod	Produktzeichen	15
$[P]$	Projektionsraum	181
ran	Rang einer Matrix	175

\mathbb{Q}	rationale Zahlen	9
$\operatorname{Re}(z)$, $\operatorname{Im}(z)$	Realteil, Imaginärteil	77
aRb , $R(a_1, \dots, a_n)$	Relation	19
$R_{f,n}$	Restglied	322
$\mathbb{Z}/p\mathbb{Z}$	Restklassen modulo p	24
\mathbb{R}^n	\mathbb{R}^n	10
\cap , \cup	Schnitt, Vereinigung von Mengen	10
sgn	Vorzeichen einer Permutation	32
sgn	Vorzeichen einer reellen Zahl	88
$\langle \mathbf{x} \mathbf{y} \rangle$	Skalarprodukt	65
$\sigma(A)$	Spektrum	177
\sum	Summenzeichen	15
\sup , \inf	Supremum, Infimum	222
$T_{f,n}$, T_f	TAYLOR-Polynom, TAYLOR-Reihe	322
$a \mid b$	a teilt b	35
\subseteq , \supseteq	Teilmenge, Obermenge	7
$T(\mathbf{a}, \mathbf{p}, \mathbf{b})$	Teilverhältnis	121
A^t	Transponierte	105
t_{ij}	Transposition	30
f^{-1}	Umkehrfunktion	28
\wedge , \vee	und, oder	2
$f^{-1}(B)$	Urbild	26
$\mathbf{x} = [x_1, \dots, x_n]^t$	Vektor	61
$g \circ f$	Verkettung von Funktionen	25
$S \circ R$	Verkettung von Relationen	24
$\sqrt[n]{}$	n -te Wurzel	224
\mathbb{Z}_p^*	endlicher Körper zu $p \in \mathbb{N}$	41
$[a]_{\sim}$, $[a]$	Äquivalenzklasse	22

Index

- Abbildung
 - lineare, 135
 - längentreue, 149
 - normale, 185
 - positive, 183
 - schiefsymmetrische, 188
 - selbstadjungierte, 179
 - winkeltreue, 149
- Abstand
 - Punkt zu Ebene, 70
 - Punkt zu Gerade, 75
- Additionssätze
 - der Trigonometrie, 79, 138
- Adjungierte
 - Matrix, 147, 148
- Adjunkte, 172
- Algorithmus
 - euklidischer, 36
 - erweiterter, 39, 46
- antilinear, 89, 93
- Antisymmetrie, 93
- Aufspann, x
- Aussageform, 5, 8
- Aussagen, 1
- Aussagenlogik, 1
- Banachraum, 95
- Basis, 124
 - duale, 150, 152
 - Entwicklung nach einer, 127, 150
 - kanonische, 127
- Basiswechsel, 150
- Betrag
 - einer komplexen Zahl, 78
- Beweis
 - direkter, 3
 - durch Widerspruch, 4
- bijektiv, 27
- Binom
 - drittes, 29, 193
- Binomialkoeffizient, 15, 44, 53
- binomischer Lehrsatz, 15
- Blockmatrix, 164
- Brennpunkt, 189, 191, 193
- Cardanische Formeln, 76, 84
- CAUCHY-SCHWARZ-Ungleichung, 67, 93, 94, 97
- charakteristisches Polynom, 178
- Chinesischer Restsatz, 47
- Definitheit, 93
- δ_{ik} , 132, 138
- Determinante, 73, 160
 - Summendarstellung, 162
- Determinanten-Produktsatz, 164
- Diagramm
 - kommutierendes, 157
- Dimension
 - eines Vektorraums, 125
- direkte Summe, 138
- disjunkt, 10
- Disjunktion, 2
- Diskriminante, 76, 84, 211
- Drehung, 185–187
 - ebene, 137
- Dreibein, 72
- Dreiecks-Blockmatrix
 - obere/untere, 165
- Dreiecksform
 - obere, 104, 116
- Dreiecksmatrix

- obere/untere, 163
- Dreiecksungleichung, 78, 93
 - umgekehrte, 78, 96
- Dreieckszahl, 9
- 3-er Probe, 41
- duale Basis, 150, 152
- Eigenvektor, 177
- Eigenwert, 177
- Eigenwertgleichung, 177
- Einheitskreis, 80
- Einheitsmatrix, 138
- Einheitswurzel
 - n-te, 81
- 1, 138
- 11-er Probe, 42
- Ellipse, 189
 - Tangente, 189
- Ellipsoid, 198
- Entwicklungskoeffizienten, 127
- ε -Umgebung, 96
- erzeugend, 124
- EULER-Formel, 79
- Fakultät, 15
- Faßkreisbogen
 - Satz vom, 152
- FERMAT
 - kleiner Satz von, 44
- FIBONACCI-Folge, 208, 210
- Funktion, 21, 25
 - Einschränkung der, 26
 - identische, 26
- Funktionswert, 25
- GAUSS-Verfahren, 100, 107
 - erweitertes, 144
 - komplexes, 128
- GAUSSsche Zahlenebene, 77, 78
- Gleichheit
 - modulo, 40
- Gleichung
 - kubische, 76, 84
- Gleichungssystem
 - in/homogenes, 100, 106
 - lineares, 99
- GOLDBACHsche Vermutung, 1
- Goldener Schnitt, 121
- goldener Schnitt, 209
- GRAM-SCHMIDT-Verfahren, 133
- Graph einer Funktion, 25
- GRASSMANN-Identität, 188
- GRASSMANN-PLÜCKER-Identität, 188
- Gruppe, 43
 - symmetrische, 30
- Gruppenhomomorphismus, 32
- HERONS Formel, 80
- HESSE-Form, 71
- HESSE-Form, 70
- Hilbertraum, 95, 132
- homogene Lösung, 106
- Hyperbel, 191
 - Tangente, 191
- Hyperboloid, 198
- Höhe
 - im Dreieck, 153
- HÖLDER-Ungleichung, 97
- identische Relation, 24
- Implikation, 2
- Induktion
 - vollständige, 13
- Induktionsprinzip, 13
- inhomogene Lösung, 106
- inhomogenes LGS
 - Lösungsstruktur, 106
- injektiv, 27
- Inkreis, 80
- Inkreismittelpunkt, 123
- Inkreisradius, 80
- Intervall, 8
 - abgeschlossenes, offenes, 8
 - halboffenes, 8
- Inverse, 142
 - Matrix, 143, 144, 171, 173
 - modulo p , 42
 - Relation, 24
- Inversion, 30

- kanonische Basis, 127
- Kegelschnitt, 189
 - Polardarstellung, 195
- Kern
 - einer linearen Abbildung, 141
- Koeffizient, 99
- Kombinatorik, 51
- kommutierendes Diagramm, 157
- konjugiert komplex, 78
- Konjunktion, 2
- Koordinatenform
 - der Kugelgleichung, 74
 - einer Ebene, 71
- Koordinatentransformation, 157
- Koordinatenvektor
 - bzgl. einer Basis, 156
- Kosinussatz, 66
- Kreuzprodukt, 68, 175, 188
- KRONECKER-Symbol, 132, 138
- kubische Gleichung, 76, 84
- Kugel
 - offene, abgeschlossene, 96
- Kugelgleichung, 74
- Körper, 43

- l. a., 114
- l. u., 114
- LAGRANGE-Identität, 188
- leere Menge, 10
- Leitgerade, 193, 195
- Lemma
 - von EUKLID, 37, 38, 46
- lh, 124
- linear, 106
- linear abhängig, 114
- linear unabhängig, 74, 114
- lineare Abbildung
 - orthogonale, 149
 - unitäre, 149
- lineare Abbildung, 135
 - unitäre, 185
- lineare Hülle, 124
- lineares Gleichungssystem, 99
- Linearform, 97

- Linearkombination, 63, 91
- Lotto, 59
- L^p -Norm, 96, 97
- L^p -Raum, 98
- Länge
 - eines Vektors, 65

- magisches Quadrat, 109
- Matrix, 105
 - adjungierte, 147, 148
 - inverse, 143, 144, 171, 173, 175
 - längentreue, 149
 - orthogonale, 149
 - quadratische, 143
 - unitäre, 149
 - winkeltreue, 149
- Matrixprodukt, 137
- MENELAOS
 - Satz von, 122
- MINKOWSKI-Ungleichung, 98
- Minoren, 170
- Mittelpunkt
 - einer Kugel, 74
- Mitternachtsformel, 76, 211
- modulo, 40
- Monom, 116
- DE MORGANSche Regeln
 - für Mengen, 11
- DE MORGANSche Regeln, 4
- multilinear, 159
- Multilinearform
 - alternierende, 159
- Multinomialkoeffizient, 53
- Multiplikation
 - Matrix·Matrix, 137
 - Matrix·Vektor, 106

- \mathbb{N} , 12
- \mathbb{N}_0 , 13
- natürliche Zahlen, 12
- Nebenwinkel, 195
- 9-er Probe, 41
- \mathbb{N}_n , 30
- Norm, 65, 93

- auf \mathbb{C}^n , 89
- auf \mathbb{R}^n , 75
- euklidische, 65, 75, 89
- Normalenform
 - einer Ebene, 69
- Normalenvektor, 69
- normiert, 67
- normierter Raum, 95
- Nullabbildung, 138
- Nullkombination, 114
 - triviale, 114
- Nullmatrix, 138
- Nullpolynom, 117
- nullteilerfrei, 78
- Nullvektor, 61

- o. B. d. A., 8
- ONB, 132
- Ordnungsrelation, 21
- Orientierung
 - von Vektoren, 73
- orthogonal, 67
 - Matrix, 149
- Orthonormalbasis, 132

- Parabel, 193
- Parameterdarstellung
 - einer Ebene, 65
- PASCALSches Dreieck, 16
- PAULI-Matrizen, 180
- Permutation, 30, 52, 161
 - (un)gerade, 32
- Permutationsmatrix, 187
- Pfadbild
 - der Determinante, 162
- Pfadregel, 163, 164
- Polardarstellung, 79
- Polarisationsgleichung, 149, 184
- Polynomdivision, 29
- positiv orientiert, 73
- Potenzfunktion, 28
- pq-Formel, 211
- Primzahl, 9
- Produktmenge, 9

- Produktzeichen, 15
- Projektion, 180
 - eindimensionale, 182
 - nicht triviale, 181
 - Spektrum der, 181
- Projektionsraum, 180
- Projektor, 180
- PYTHAGORAS
 - Satz des, 65

- Quader, 10
- quadratische Form, 196
- Quadratwurzel, 88
- Quadrik, 196, 199
- Quantor, 5
- Quersumme, 41
 - alternierende, 42

- Radius, 74
- Rang
 - einer Matrix, 175
 - einer Matrix, 173
 - einer Menge, 173
- rationale Zahlen, 9
- Raum
 - mit Skalarprodukt, 95
 - normierter, 95
- Rechte-Hand-Regel, 73
- Reihe
 - geometrische, 206
- Rekurrenzgleichung, 205, 208
 - 1. Ordnung, 206
 - 2. Ordnung, 208, 209
- Rekursionsgleichung, 205, 208
- Relation, 19
 - bijektive, 21
 - funktionale, 21
 - homogene, 21
 - identische, 24
 - injektive, 21
 - inverse, 24
 - linkstotale, 21
 - rechtstotale, 21
 - reflexive, 21

- surjektive, 21
- symmetrische, 21
- transitive, 21
- Repräsentant
 - einer Äquivalenzrelation, 22
- Richtungsvektor, 64
- RSA-Verschlüsselung, 45
- Rückwärtsentwicklung, 205
- Satz
 - vom Faßkreisbogen, 152
 - von RIESZ, 97
 - VON MENELAOS, 122
 - VON PYTHAGORAS, 65
- Schlüssel
 - öffentlicher/privater, 45
- Schnitt, 23
- selbstadjungiert, 179
- sesquilinear, 93
- Sesquilinearform, 93
- Signum, 88
- Skalarmultiplikation, 61
- Skalarprodukt, 65, 93
 - auf \mathbb{C}^n , 89
 - auf \mathbb{R}^n , 75
 - euklidisches, 75, 89
- Skalarprodukt-Raum, 95
- Spaltenvektor, 105
- Spatprodukt, 71
- Spektrum, 177
- Spiegelung, 185
- Strahlensatz, 119
- Stützvektor, 64
- Summationsindex, 15
- Summenzeichen, 15
- surjektiv, 27
- symmetrische Gruppe, 30
- Tangentialebene
 - an die Kugel, 75
- Tautologie, 3
- Teilen mit Rest, 35
- Teiler, 35
 - größter gemeinsamer, 36
- teilerfremd, 36
- Teilmenge, 7
 - einer n -elementigen Menge, 53
- Teilraum, 91
 - orthogonaler, 139
- Teilverhältnis, 121
- Torus, 11
- transponieren, 105
- Transposition, 30
- Tupel, 17
- Turm von Hanoi, 205
- Umkehrfunktion, 28
- Umkreis, 152
 - Radius, 152
- unendlichdimensional, 125
- Ungleichung
 - CAUCHY-SCHWARZsche, 93, 94, 97
 - MINKOWSKI-, 98
 - HÖLDERSche, 97
 - YOUNGSche, 98
- unitär
 - Matrix, 149
- Unterdeterminante, 170
- Urbild, 25, 26
- VANDERMONDSche Identität, 58, 59
- Vektor, viii, 61, 90
- Vektorraum, viii, 90
- Verkettung
 - von Funktionen, 25
 - von Relationen, 24
- Verschlüsselung
 - RSA-, 45
- Vielfaches
 - kleinstes gemeinsames, 48
- vollständige Induktion, 13
- Vorzeichenfunktion, 88
- Vorzeichenmatrix, 171
- Wahrheitswert, 1
- Widerspruchsbeweis, 4
- Winkelhalbierende, 195
 - äußere, 195
- wohldefiniert, 23

- Wurzel
 - n-te, [29](#)
 - einer linearen Abbildung, [184](#)
- YOUNG-Ungleichung, [98](#)
- $\mathbb{Z}/p\mathbb{Z}$, [24](#)
- Zahl
 - komplexe, [77](#)
 - konjugiert komplexe, [78](#)
- Zahlen
 - natürliche, [12](#)
 - rationale, [9](#)
- Zeilenumformung, [107](#)
 - äquivalente, [109](#)
- Zerlegung
 - einer Zahl, [54](#), [55](#)
- \mathbb{Z}_p , [41](#)
- Zwei-Quadrate-Satz, [78](#)
- Zylinder, [10](#)
- Äquivalenzklasse, [40](#)
- Äquivalenzrelation, [21](#), [22](#), [40](#)
- Überdeckung
 - disjunkte, [22](#)