

Kurzschrift zur Vorlesung
Mathematik I für MB, WI/MB und andere
Prof. Dr. Ulrich Reif

Inhalt:

1. Vektorrechnung
2. Lineare Gleichungssysteme
3. Matrizenrechnung
4. Lineare Abbildungen
5. Eigenwerte und -vektoren
6. Folgen
7. Reihen
8. Reelle Funktionen
9. Differenziation
10. Komplexe Zahlen

Vorbemerkung:

Im vorliegenden Kurzschrift werden wesentliche Begriffe, Resultate und Methoden der Vorlesung Mathematik I zusammen gestellt. Aufgrund des skizzenhaften Charakters kann es weder den Besuch der Vorlesung noch der Übungen ersetzen.

Korrekturen senden Sie bitte per Email an reif@mathematik.tu-darmstadt.de.

Mengen von Zahlen:

$\mathbb{N} = \{1, 2, 3, \dots\}$	natürliche Zahlen
$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$	natürliche Zahlen mit 0
$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$	ganze Zahlen
$\mathbb{Q} = \{p/q : p \in \mathbb{Z} \text{ und } q \in \mathbb{N}\}$	rationale Zahlen
$\mathbb{R} = \text{Menge aller Dezimalbrüche}$	reelle Zahlen
$\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$	positive reelle Zahlen
$\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$	nichtnegative reelle Zahlen
$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$	abgeschlossenes Intervall
$(a, b) = \{x \in \mathbb{R} : a < x < b\}$	offenes Intervall
$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$	halboffenes Intervall
$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$	halboffenes Intervall
$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$	Menge der komplexen Zahlen

1 Vektorrechnung

1.1 Punkte in \mathbb{R}^n : Die Menge aller *Punkte*

$$P = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \dots, x_n]^T$$

mit *Koordinaten* $x_i \in \mathbb{R}$ wird mit \mathbb{R}^n bezeichnet. Speziell erhält man für $n = 1$ die Zahlengerade \mathbb{R} , für $n = 2$ die Ebene \mathbb{R}^2 und für $n = 3$ den Raum \mathbb{R}^3 .

1.2 Rechnen mit Punkten: Sei $P = [x_1, \dots, x_n]^T$ und $Q = [y_1, \dots, y_n]^T$, dann gilt

- *Addition, Subtraktion:*

$$P + Q = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad P - Q = \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_n - y_n \end{bmatrix}$$

- *Skalarmultiplikation:*

$$\alpha P = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}, \quad \alpha \in \mathbb{R}$$

Insbesondere ist $1P = P$, $(-1)P = -P$ und $0P = \vec{0} := [0, \dots, 0]^T$.

- *Distributivgesetze:*

$$(\alpha + \beta)P = \alpha P + \beta P, \quad \alpha(P + Q) = \alpha P + \alpha Q, \quad \alpha, \beta \in \mathbb{R}$$

1.3 Vektoren in \mathbb{R}^n : Seien $P, Q \in \mathbb{R}^n$ zwei Punkte, dann ist die Differenz

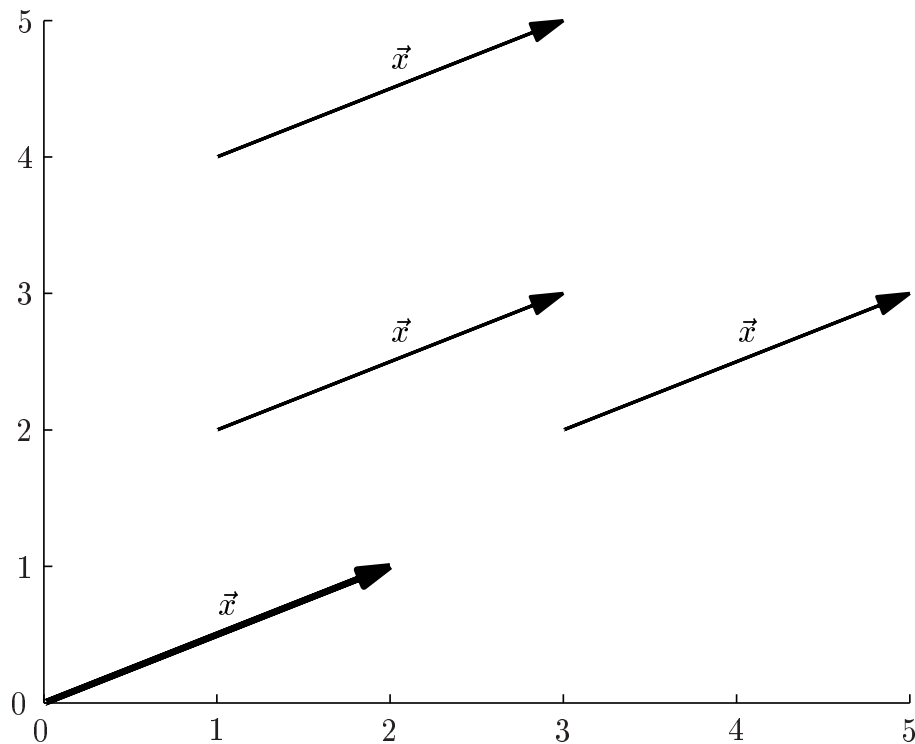
$$\vec{x} = P - Q$$

der *Vektor* vom Anfangspunkt Q zum Endpunkt P . Verschiedene Paare von Anfangs- und Endpunkten können denselben Vektor definieren. Die Lage eines Vektors im Raum ist also nicht eindeutig bestimmt, sondern kann durch eine Parallelverschiebung beliebig verändert werden. Im Folgenden ordnen wir einem Punkt $P \in \mathbb{R}^n$ einen Vektor \vec{x} dadurch eindeutig zu, indem wir als Anfangspunkt den Ursprung $Q = 0$ wählen,

$$\vec{x} = P - 0 = P.$$

Umgekehrt ist dadurch auch jedem Vektor eindeutig ein Punkt zugeordnet. Man nennt \vec{x} dann auch den *Ortsvektor* von P . Aufgrund dieser Identifizierung übertragen sich alle Rechenregeln und Gesetze für Punkte direkt auf Vektoren.

1.4 Beispiel:



Die verschiedenen Punktepaare erzeugen alle denselben Vektor, z.B.

$$\vec{x} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Fett gezeichnet ist der Ortsvektor zum Punkt $P = [2, 1]^T$.

1.5 Norm eines Vektors: Die *euklidische Norm* des Vektors $\vec{x} = [x_1, \dots, x_n]^T$ ist durch

$$\|\vec{x}\| := \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

definiert. Sie gibt die Länge des Vektors im geometrischen Sinne an und hat folgende Eigenschaften:

- *Positive Definitheit:*

$$\begin{aligned} \|\vec{x}\| &> 0 && \text{für } \vec{x} \neq \vec{0} \\ \|\vec{x}\| &= 0 && \text{für } \vec{x} = \vec{0} \end{aligned}$$

- *Homogenität:*

$$\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|, \quad \alpha \in \mathbb{R}$$

- *Dreiecksungleichung:*

$$\begin{aligned} \|\vec{x} + \vec{y}\| &\leq \|\vec{x}\| + \|\vec{y}\| \\ \|\vec{x} + \vec{y}\| &\geq \|\vec{x}\| - \|\vec{y}\| \end{aligned}$$

1.6 Normierung: Sei $\vec{x} \neq \vec{0}$ ein Vektor, dann erhält man durch die *Normierung*

$$\vec{x}_0 := \frac{\vec{x}}{\|\vec{x}\|}$$

einen Vektor \vec{x}_0 der Länge 1, der dieselbe Richtung wie \vec{x} besitzt.

1.7 Skalarprodukt: Das *Skalarprodukt* zweier Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ ist definiert durch

$$\langle \vec{x}, \vec{y} \rangle := x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^n x_i y_i.$$

Die Vektoren \vec{x}, \vec{y} heißen *orthogonal*, wenn

$$\langle \vec{x}, \vec{y} \rangle = 0.$$

Die Vektoren \vec{x} und \vec{y} stehen dann im geometrischen Sinne senkrecht aufeinander. Der Nullvektor ist orthogonal zu allen anderen Vektoren. Skalarprodukt und Norm sind durch die Formeln

$$\begin{aligned} \|\vec{x}\|^2 &= \langle \vec{x}, \vec{x} \rangle \\ \langle \vec{x}, \vec{y} \rangle &= \frac{1}{4} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2) \end{aligned}$$

verknüpft.

1.8 Eigenschaften:

- *Symmetrie:*

$$\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$$

- *Linearität*

$$\begin{aligned} \alpha \langle \vec{x}, \vec{y} \rangle &= \langle \alpha \vec{x}, \vec{y} \rangle = \langle \vec{x}, \alpha \vec{y} \rangle, \quad \alpha \in \mathbb{R} \\ \langle \vec{x}_1 + \vec{x}_2, \vec{y} \rangle &= \langle \vec{x}_1, \vec{y} \rangle + \langle \vec{x}_2, \vec{y} \rangle \\ \langle \vec{x}, \vec{y}_1 + \vec{y}_2 \rangle &= \langle \vec{x}, \vec{y}_1 \rangle + \langle \vec{x}, \vec{y}_2 \rangle \end{aligned}$$

- *Binomische Formel:*

$$\langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle$$

- Sei φ der Winkel zwischen \vec{x} und \vec{y} , dann gilt

$$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \cdot \|\vec{y}\| \cos \varphi.$$

- *Cauchy-Schwarz-Ungleichung:*

$$\langle \vec{x}, \vec{y} \rangle \leq \|\vec{x}\| \cdot \|\vec{y}\|$$

1.9 Vektorprodukt: Das *Vektorprodukt* (auch Kreuzprodukt genannt) zweier Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^3$ ist definiert durch

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}.$$

Das Ergebnis $\vec{z} = \vec{x} \times \vec{y}$ ist also wieder ein Vektor in \mathbb{R}^3 .

1.10 Eigenschaften:

- *Antisymmetrie:*

$$\vec{x} \times \vec{y} = -(\vec{y} \times \vec{x})$$

Daraus folgt insbesondere

$$\vec{x} \times \vec{x} = \vec{0}.$$

- *Linearität:*

$$\begin{aligned} \alpha(\vec{x} \times \vec{y}) &= (\alpha\vec{x}) \times \vec{y} = \vec{x} \times (\alpha\vec{y}), \quad \alpha \in \mathbb{R} \\ (\vec{x}_1 + \vec{x}_2) \times \vec{y} &= \vec{x}_1 \times \vec{y} + \vec{x}_2 \times \vec{y} \\ \vec{x} \times (\vec{y}_1 + \vec{y}_2) &= \vec{x} \times \vec{y}_1 + \vec{x} \times \vec{y}_2 \end{aligned}$$

- *Orthogonalität:*

$$\langle \vec{x} \times \vec{y}, \vec{x} \rangle = \langle \vec{x} \times \vec{y}, \vec{y} \rangle = 0$$

- Sei φ der Winkel zwischen \vec{x} und \vec{y} , dann gilt

$$\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \cdot \|\vec{y}\| \cdot |\sin \varphi|.$$

$\vec{x} \times \vec{y}$ ist also ein Vektor, der senkrecht auf \vec{x} und \vec{y} steht und als Länge den Flächeninhalt des von \vec{x} und \vec{y} aufgespannten Parallelogramms besitzt.

- Die drei Vektoren $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$ bilden in dieser Reihenfolge ein *Rechtssystem*. Das heißt anschaulich gesprochen folgendes: Wenn der Daumen und der Zeigefinger der rechten Hand in Richtung \vec{x} und \vec{y} zeigen, dann zeigt der Mittelfinger in Richtung $\vec{x} \times \vec{y}$.

1.11 Linearkombination: Seien $\vec{p}_1, \dots, \vec{p}_m$ Vektoren in \mathbb{R}^n und $\lambda_1, \dots, \lambda_m$ reelle Zahlen. Dann nennt man den Vektor

$$\vec{x} := \lambda_1 \vec{p}_1 + \dots + \lambda_m \vec{p}_m = \sum_{i=1}^m \lambda_i \vec{p}_i$$

eine *Linearkombination* der Vektoren \vec{p}_i mit Koeffizienten λ_i . Die Vektoren $\vec{p}_1, \dots, \vec{p}_m$ heißen *linear unabhängig*, wenn der Nullvektor dann und nur dann Ergebnis der Linearkombination ist, wenn alle Koeffizienten verschwinden,

$$\sum_{i=1}^m \lambda_i \vec{p}_i = \vec{0} \quad \Leftrightarrow \quad \lambda_1 = \dots = \lambda_m = 0.$$

Wählt man speziell die *Einheitsvektoren*

$$\vec{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n := \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

der Standardbasis, dann gilt für jeden Punkt $\vec{x} \in \mathbb{R}^n$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^n x_i \vec{e}_i.$$

1.12 Geraden in \mathbb{R}^n : Seien \vec{p} und \vec{r} Vektoren in \mathbb{R}^n und $\vec{r} \neq 0$. Die Gleichung

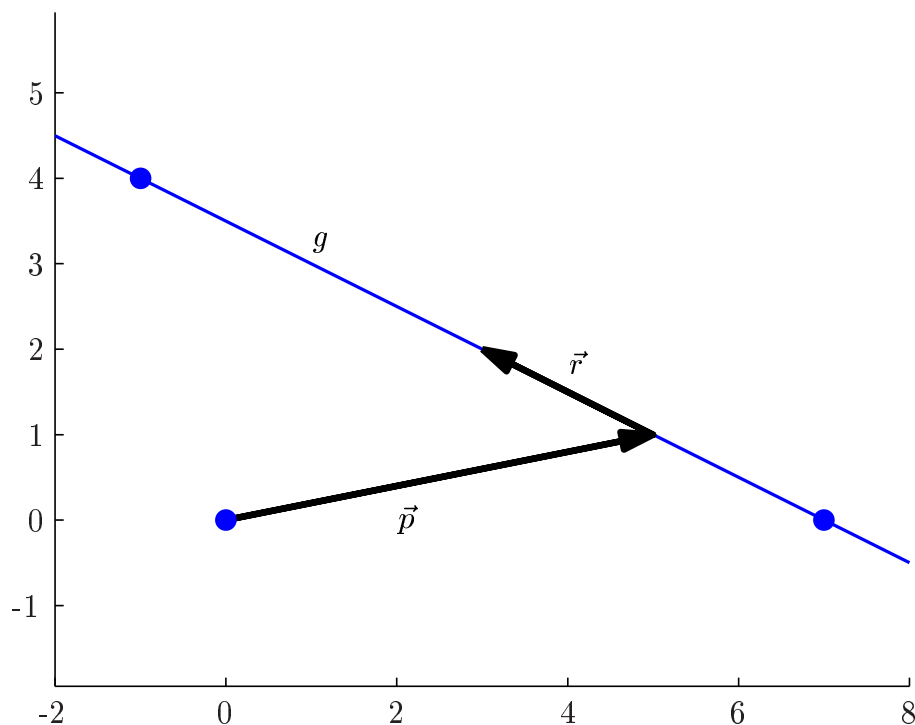
$$g : \vec{x} = \vec{p} + \lambda \vec{r}, \quad \lambda \in \mathbb{R},$$

beschreibt eine *Gerade in \mathbb{R}^n in parametrisierter Form*. Man bezeichnet \vec{p} als *Aufpunkt*, \vec{r} als *Richtungsvektor* und λ als *Parameter* der Geraden g .

1.13 Beispiel: Im Bild ist $\vec{p} = [5, 1]^T$ und $\vec{r} = [-2, 1]^T$, also

$$g : \vec{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

Für $\lambda = 3$ erhält man den Punkt $\vec{x} = [-1, 4]^T$ und für $\lambda = -1$ den Punkt $\vec{x} = [7, 0]^T$.



1.14 Abstand Punkt-Gerade: Der Abstand $d(\vec{q}, g)$ eines Punktes \vec{q} von der Geraden $g : \vec{x} = \vec{p} + \lambda\vec{r}$ ist definiert als

$$d(\vec{q}, g) := \min_{\lambda \in \mathbb{R}} \|\vec{x} - \vec{q}\|.$$

Dies ist also der kleinste Abstand, den ein Punkt auf der Geraden von \vec{q} haben kann. Der Punkt $\vec{x}^* = \vec{p} + \lambda^*\vec{r}$, für den dieses Minimum angenommen wird, ist dadurch gekennzeichnet, dass der Verbindungsvektor zum Punkt \vec{q} senkrecht zum Richtungsvektor \vec{r} der Geraden ist,

$$\langle \vec{x}^* - \vec{q}, \vec{r} \rangle = \langle \vec{p} - \vec{q} + \lambda^*\vec{r}, \vec{r} \rangle = 0.$$

Löst man diese Gleichung nach λ^* auf, so erhält man

$$\lambda^* = \frac{\langle \vec{q} - \vec{p}, \vec{r} \rangle}{\langle \vec{r}, \vec{r} \rangle}$$

und damit \vec{x}^* . Schließlich ist

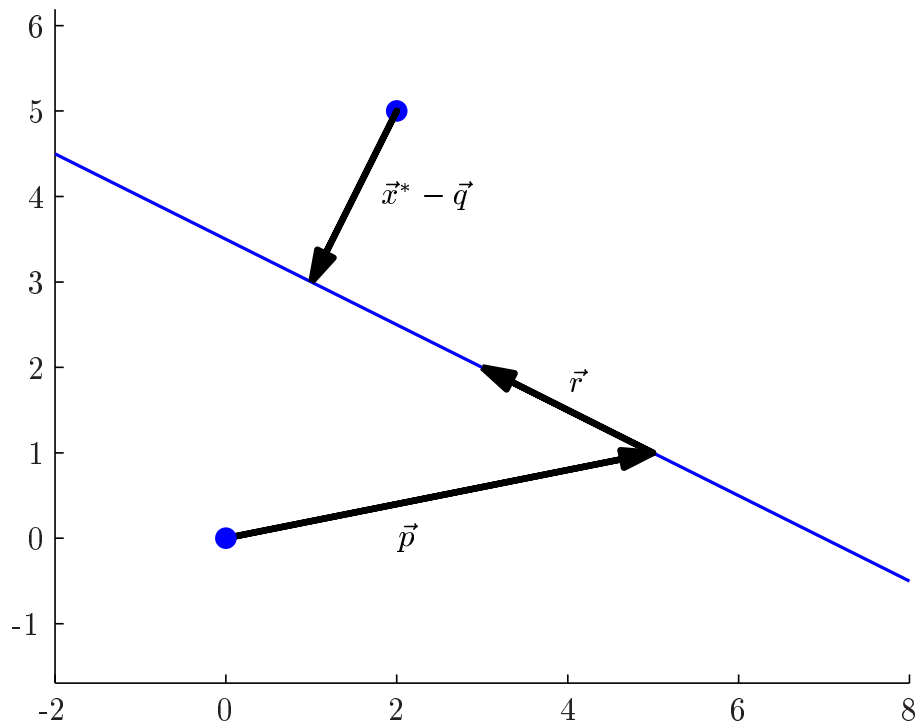
$$d(\vec{q}, g) = \|\vec{x}^* - \vec{q}\|.$$

1.15 Beispiel [\rightarrow 1.13]: Für $g : \vec{x} = [5, 1]^T + \lambda[-2, 1]^T$ und $\vec{q} = [2, 5]^T$ ist

$$\lambda^* = 2 \quad \text{und} \quad \vec{x}^* = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Damit erhält man

$$d(\vec{q}, g) = \|\vec{x}^* - \vec{q}\| = \sqrt{5}.$$



1.16 Abstand Gerade-Gerade: Gegeben seien zwei Geraden in \mathbb{R}^n ,

$$\begin{aligned} g_1 &: \vec{x}_1 = \vec{p}_1 + \lambda_1 \vec{r}_1, & \lambda_1 &\in \mathbb{R} \\ g_2 &: \vec{x}_2 = \vec{p}_2 + \lambda_2 \vec{r}_2, & \lambda_2 &\in \mathbb{R}. \end{aligned}$$

Der *Abstand* $d(g_1, g_2)$ der beiden Geraden voneinander ist definiert als

$$d(g_1, g_2) := \min_{\lambda_1, \lambda_2 \in \mathbb{R}} \|\vec{x}_1 - \vec{x}_2\|.$$

Dies ist also der kleinste Abstand, den zwei Punkte auf den Geraden voneinander haben können. Das Punktepaar $\vec{x}_1^* = \vec{p}_1 + \lambda_1^* \vec{r}_1$, $\vec{x}_2^* = \vec{p}_2 + \lambda_2^* \vec{r}_2$, für den dieses Minimum angenommen wird, ist dadurch gekennzeichnet, dass der Verbindungsvektor senkrecht zu beiden Richtungsvektoren \vec{r}_1 und \vec{r}_2 der Geraden ist. Dies führt auf zwei lineare Gleichungen für die beiden Parameter λ_1^* , λ_2^* , das man mit den Methoden, die im nächsten Kapitel beschrieben werden, lösen kann. Handelt es sich insbesondere um Geraden im Raum \mathbb{R}^3 , so kann das Problem einfacher wie folgt gelöst werden: Der Vektor $\vec{n} := \vec{r}_1 \times \vec{r}_2$ steht senkrecht auf \vec{r}_1 und \vec{r}_2 . Falls $\vec{n} \neq 0$, dann gilt also für einen ebenfalls unbekanntem Parameter $\mu \in \mathbb{R}$

$$\vec{x}_1^* - \vec{x}_2^* = \mu \vec{n}.$$

Multipliziert man diese Gleichung skalar mit \vec{n} , dann erhält man

$$\langle \vec{p}_1 - \vec{p}_2, \vec{n} \rangle = \mu \langle \vec{n}, \vec{n} \rangle \quad \Rightarrow \quad \mu = \frac{\langle \vec{p}_1 - \vec{p}_2, \vec{n} \rangle}{\|\vec{n}\|^2},$$

da $\langle \vec{r}_1, \vec{n} \rangle = \langle \vec{r}_2, \vec{n} \rangle = 0$. Der Abstand ist also

$$d(g_1, g_2) = \|\mu \vec{n}\| = |\mu| \cdot \|\vec{n}\| = \frac{|\langle \vec{p}_1 - \vec{p}_2, \vec{n} \rangle|}{\|\vec{n}\|}.$$

1.17 Geraden in impliziter Form: Sei

$$g : \vec{x} = \vec{p} + \lambda \vec{r}, \quad \lambda \in \mathbb{R},$$

eine Gerade in \mathbb{R}^2 und $\vec{n} \neq \vec{0}$ ein *Normalenvektor*. Dies ist ein Vektor, der senkrecht auf \vec{r} steht, also $\langle \vec{r}, \vec{n} \rangle = 0$. Multipliziert man die Gleichung der Geraden skalar mit \vec{n} , dann erhält man die *implizite Form*

$$g : \langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle.$$

Die Gerade g ist also die Menge aller Punkte $\vec{x} \in \mathbb{R}^2$, die diese Gleichung erfüllen. Auf der linken Seite steht eine Linearkombination der Komponenten von \vec{x} , und auf der rechten Seite steht die Konstante $d := \langle \vec{p}, \vec{n} \rangle \in \mathbb{R}$. Einen Normalenvektor \vec{n} erhält man beispielsweise gemäß

$$\vec{r} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \Rightarrow \quad \vec{n} := \begin{bmatrix} b \\ -a \end{bmatrix}.$$

Damit lautet die implizite Form

$$g : bx - ay = d.$$

1.18 Hessesche Normalform: Normiert man speziell den Normalenvektor \vec{n} auf Länge 1 [\rightarrow 1.6], das heißt

$$\vec{n}_0 := \frac{\vec{n}}{\|\vec{n}\|},$$

dann lautet die implizite Form

$$g : \langle \vec{x}, \vec{n}_0 \rangle = d_0, \quad \text{wobei} \quad d_0 := \langle \vec{p}, \vec{n}_0 \rangle = \frac{d}{\|\vec{n}\|}.$$

Diese bezeichnet man als die *Hessesche Normalform* der Geraden g . Sie ist dadurch ausgezeichnet, dass der Betrag der Konstanten d_0 den Abstand der Geraden vom Ursprung angibt, also

$$d(\vec{0}, g) = |d_0|.$$

Der Abstand eines beliebigen Punktes $\vec{q} \in \mathbb{R}^2$ von der Geraden ist durch

$$d(\vec{q}, g) = |d_0 - \langle \vec{q}, \vec{n}_0 \rangle|$$

gegeben.

1.19 Beispiel [\rightarrow 1.15]: Sei $g : \vec{x} = [5, 1]^T + \lambda[-2, 1]^T$ und $\vec{q} = [2, 5]^T$. Man erhält

$$\vec{r} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \Rightarrow \vec{n} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow d = \langle \vec{p}, \vec{n} \rangle = 7$$

und damit die implizite Form

$$g : x + 2y = 7.$$

Beispielsweise erfüllen die Punkte $x = 7, y = 0$ und $x = -1, y = 4$ diese Gleichung [\rightarrow 1.13]. Die Normierung

$$\|\vec{n}\| = \sqrt{5} \Rightarrow \vec{n}_0 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow d_0 = \langle \vec{p}, \vec{n}_0 \rangle = \frac{7}{\sqrt{5}}$$

liefert die Hessesche Normalform

$$g : \frac{1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y = \frac{7}{\sqrt{5}}.$$

Der Abstand der Geraden vom Ursprung ist also $d_0 = 7/\sqrt{5}$. Der Abstand des Punktes $\vec{q} = [2, 5]^T$ ist wie zuvor

$$d(\vec{q}, g) = |7/\sqrt{5} - \langle [2, 5]^T, [1, 2]^T \rangle / \sqrt{5}| = \sqrt{5}.$$

1.20 Ebenen in \mathbb{R}^3 : Seien $\vec{p}, \vec{r}_1, \vec{r}_2$ Vektoren in \mathbb{R}^3 und $\vec{n} := \vec{r}_1 \times \vec{r}_2 \neq \vec{0}$. Die Gleichung

$$E : \vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2, \quad \lambda_1, \lambda_2 \in \mathbb{R},$$

beschreibt eine Ebene in \mathbb{R}^3 in parametrisierter Form. Man bezeichnet \vec{p} als *Aufpunkt*, \vec{r}_1, \vec{r}_2 als *Richtungsvektoren*, \vec{n} als *Normalenvektor* und λ_1, λ_2 als *Parameter* der Ebene.

1.21 Abstand Punkt-Ebene: Der Abstand $d(\vec{q}, E)$ eines Punktes \vec{q} von der Ebene $\vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2$ ist definiert als

$$d(\vec{q}, E) := \min_{\lambda_1, \lambda_2 \in \mathbb{R}} \|\vec{x} - \vec{q}\|.$$

Dies ist also der kleinste Abstand, den ein Punkt auf der Ebene von \vec{q} haben kann. Der Punkt $\vec{x}^* = \vec{p} + \lambda_1^* \vec{r}_1 + \lambda_2^* \vec{r}_2$, für den dieses Minimum angenommen wird, ist dadurch gekennzeichnet, dass der Verbindungsvektor zum Punkt \vec{q} senkrecht zu beiden Richtungsvektoren \vec{r}_1, \vec{r}_2 der Ebene ist, d.h.,

$$\vec{x}^* - \vec{q} = \mu \vec{n}.$$

Multipliziert man diese Gleichung skalar mit \vec{n} , dann erhält man

$$\langle \vec{p} - \vec{q}, \vec{n} \rangle = \mu \langle \vec{n}, \vec{n} \rangle \quad \Rightarrow \quad \mu = \frac{\langle \vec{p} - \vec{q}, \vec{n} \rangle}{\|\vec{n}\|^2},$$

da $\langle \vec{r}_1, \vec{n} \rangle = \langle \vec{r}_2, \vec{n} \rangle = 0$. Der Abstand ist also

$$d(\vec{q}, E) = \|\mu \vec{n}\| = |\mu| \cdot \|\vec{n}\| = \frac{|\langle \vec{p} - \vec{q}, \vec{n} \rangle|}{\|\vec{n}\|}.$$

Man beachte die Analogie zur Berechnung des Abstands Gerade-Gerade.

1.22 Beispiel: Sei $\vec{p} = [1, 1, 5]^T$, $\vec{r}_1 = [3, 0, 1]^T$ und $\vec{r}_2 = [1, 2, -1]^T$, also

$$E : \vec{x} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} + \lambda_1 \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Dann ist der Normalenvektor gegeben durch

$$\vec{n} = \vec{r}_1 \times \vec{r}_2 = \begin{bmatrix} -2 \\ 4 \\ 6 \end{bmatrix} \quad \Rightarrow \quad \|\vec{n}\| = \sqrt{56}.$$

Der Abstand des Punktes $\vec{q} = [1, 0, 7]^T$ von der Ebene ist

$$d(\vec{q}, E) = \frac{|-8|}{\sqrt{56}} = \frac{4}{\sqrt{14}}.$$

Dabei ist $\mu = -1/7$ und $\vec{x}^* = [9/7, -4/7, 43/7]^T$.

1.23 Ebenen in impliziter Form: Sei

$$E : \vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2, \quad \lambda_1, \lambda_2 \in \mathbb{R},$$

eine Ebene in \mathbb{R}^3 mit Normalenvektor \vec{n} . Multipliziert man die Gleichung der Ebene skalar mit \vec{n} , dann erhält man die *implizite Form*

$$E : \langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle.$$

Die Ebene E ist also die Menge aller Punkte $\vec{x} \in \mathbb{R}^3$, die diese Gleichung erfüllen. Auf der linken Seite steht eine Linearkombination der Komponenten von \vec{x} , und auf der rechten Seite steht die Konstante $d := \langle \vec{p}, \vec{n} \rangle \in \mathbb{R}$.

1.24 Hessesche Normalform: Normiert man speziell den Normalenvektor \vec{n} auf Länge 1, d.h.,

$$\vec{n}_0 := \frac{\vec{n}}{\|\vec{n}\|},$$

dann lautet die implizite Form

$$E : \langle \vec{x}, \vec{n}_0 \rangle = d_0, \quad \text{wobei} \quad d_0 := \langle \vec{p}, \vec{n}_0 \rangle = \frac{d}{\|\vec{n}\|}.$$

Diese bezeichnet man als die *Hessesche Normalform* der Ebene E . Sie ist dadurch ausgezeichnet, dass der Betrag der Konstanten d_0 den Abstand der Ebene vom Ursprung angibt, also

$$d(\vec{0}, E) = |d_0|.$$

Der Abstand eines beliebigen Punktes $\vec{q} \in \mathbb{R}^3$ von der Ebene ist durch

$$d(\vec{q}, E) = |d_0 - \langle \vec{q}, \vec{n}_0 \rangle|$$

gegeben.

1.25 Beispiel [\rightarrow 1.22]: Mit $\vec{p} = [1, 1, 5]^T$, $\vec{n} = [-2, 4, 6]^T$ und $\vec{x} = [x, y, z]^T$ erhält man $d = \langle \vec{p}, \vec{n} \rangle = 32$ und damit die implizite Form

$$E : -2x + 4y + 6z = 32.$$

Die Normierung

$$\|\vec{n}\| = \sqrt{56} = 2\sqrt{14} \quad \Rightarrow \quad \vec{n}_0 = \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \quad \Rightarrow \quad d_0 = \langle \vec{p}, \vec{n}_0 \rangle = \frac{16}{\sqrt{14}}$$

liefert die Hessesche Normalform

$$E : \frac{-1}{\sqrt{14}} x + \frac{2}{\sqrt{14}} y + \frac{3}{\sqrt{14}} z = \frac{16}{\sqrt{14}}.$$

Der Abstand der Ebene vom Ursprung ist also $d_0 = 16/\sqrt{14}$, und er Abstand des Punktes $\vec{q} = [1, 0, 7]^T$ von der Ebene ist wie zuvor

$$d(\vec{q}, E) = |16/\sqrt{14} - 20/\sqrt{14}| = 4/\sqrt{14}.$$

1.26 Schnitt Ebene-Gerade: Zur Berechnung des Schnittpunkts \vec{x}^* einer Ebene E mit einer Geraden g in \mathbb{R}^3 verwendet man zweckmäßigerweise für die Ebene die implizite und für die Gerade die parametrische Form,

$$\begin{aligned} E &: \langle \vec{x}, \vec{n} \rangle = d \\ g &: \vec{x} = \vec{p} + \lambda \vec{r}, \quad \lambda \in \mathbb{R}. \end{aligned}$$

Setzt man die Geradengleichung in die Ebenengleichung ein, so erhält man die Bedingung

$$\langle \vec{p}, \vec{n} \rangle + \lambda^* \langle \vec{r}, \vec{n} \rangle = d \quad (1.1)$$

für den Parameter λ^* des Schnittpunkts. Nun sind folgende Fälle zu unterscheiden:

- Falls $\langle \vec{r}, \vec{n} \rangle \neq 0$, dann ist der Schnittpunkt eindeutig bestimmt und durch

$$\vec{x}^* = \vec{p} + \frac{d - \langle \vec{p}, \vec{n} \rangle}{\langle \vec{r}, \vec{n} \rangle} \vec{r}$$

gegeben.

- Falls $\langle \vec{r}, \vec{n} \rangle = 0$ und $\langle \vec{p}, \vec{n} \rangle = d$, dann ist Gleichung (1.1) für alle $\lambda^* \in \mathbb{R}$ erfüllt; es gibt also unendlich viele Lösungen. Dies bedeutet, dass die Gerade parallel zur Ebene ist und vollständig in dieser liegt.
- Falls $\langle \vec{r}, \vec{n} \rangle = 0$ und $\langle \vec{p}, \vec{n} \rangle \neq d$, dann ist Gleichung (1.1) für kein $\lambda^* \in \mathbb{R}$ erfüllt; es gibt also keinen Schnittpunkt. Dies bedeutet, dass die Gerade parallel zur Ebene ist und nicht in dieser liegt.

2 Lineare Gleichungssysteme

2.1 Beispiel: Man berechne den Schnittpunkt der drei Ebenen

$$E_1 : 2x + y + z = 1$$

$$E_2 : 3x + y + z = 2$$

$$E_3 : 4x + 2y + 3z = 0$$

Subtrahiert man das dreifache der ersten Zeile vom doppelten der zweiten Zeile, so erhält man die Bedingung

$$-y - z = 1.$$

Subtrahiert man das doppelte der ersten Zeile von der dritten Zeile, so erhält man die Bedingung

$$z = -2.$$

Setzt man dies in die vorherige Gleichung ein, so erhält man die Bedingung

$$-y + 2 = 1$$

und damit den Wert $y = 1$. Setzt man dies in die erste Gleichung ein, so erhält man die Bedingung

$$2x + 1 - 2 = 1$$

und damit den Wert $x = 1$. Der Schnittpunkt ist also $\vec{x} = [1, 1, -2]^T$.

2.2 Beispiel [\rightarrow 2.1]: Es soll die Schnittmenge der Ebenen E_1 und E_2 berechnet werden. Wie zuvor erhält man aus E_1 und E_2 die Bedingung

$$-y - z = 1.$$

Da keine weiteren Bedingungen vorhanden sind, kann man beispielweise der Variablen z einen beliebigen Wert zuordnen, sagen wir

$$z = t, \quad t \in \mathbb{R}.$$

Damit ergibt sich

$$-y - t = 1 \quad \Rightarrow \quad y = -1 - t.$$

Setzt man dies in die Gleichung von E_1 ein, so erhält man die Bedingung

$$2x + (-t - 1) + t = 1$$

und damit $x = 1$. Die Menge aller Schnittpunkte ist also gegeben durch

$$\vec{x} = \begin{bmatrix} 1 \\ -t - 1 \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Die Lösungsmenge ist also eine Gerade.

2.5 Beispiel:

- In Bsp. 2.1 ergibt sich ein quadratisches Gleichungssystem mit

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \\ 4 & 2 & 3 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

- In Bsp. 2.2 ergibt sich ein unterbestimmtes Gleichungssystem mit

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

2.6 Elementare Umformungen: Es ist zweckmäßig, das LGS $A\vec{x} = \vec{b}$ in folgendem Schema zu notieren:

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \\ \vdots \\ \boxed{m} : \end{array} \begin{array}{cccc|c} x_1 & x_2 & \cdots & x_n & \vec{b} \\ a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}$$

Die Zahlen in den Kästchen enthalten fortlaufende Zeilennummern, die nur der Kennzeichnung dienen. In dem Schema sind die folgenden *elementaren Umformungen* erlaubt:

- *Zeilenvertauschung:* Zwei Zeilen

$$\boxed{i} \leftrightarrow \boxed{j}$$

dürfen vertauscht werden.

- *Spaltenvertauschung:* Zwei Spalten

$$x_i \leftrightarrow x_j$$

dürfen vertauscht werden. Dabei ist zu beachten, dass auch die Einträge in der Kopfzeile vertauscht werden.

- *Linearkombination:* Die i -te Zeile darf durch die Linearkombination

$$\boxed{i} \leftarrow p \times \boxed{i} - q \times \boxed{j}$$

ersetzt werden, sofern $p \neq 0$. Insbesondere kann man $q = 0$ wählen und so eine Skalierung der i -ten Zeile erreichen.

Durch geeignete elementare Umformungen kann man ein gegebenes LGS in eine einfachere Form überführen, deren Lösung sich unmittelbar ablesen lässt.

2.7 Beispiel [\rightarrow **2.1**]: Das Schema zu dem angegebenen LGS hat die Form

$$\begin{array}{l} \boxed{1}: \\ \boxed{2}: \\ \boxed{3}: \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 2 \\ 4 & 2 & 3 & 0 \end{array}$$

Durch Linearkombination können die jeweils ersten Koeffizienten der zweiten und der dritten Zeile zu Null gemacht werden:

$$\begin{array}{l} \boxed{1}: \\ 2 \times \boxed{2} - 3 \times \boxed{1} = \boxed{4}: \\ 1 \times \boxed{3} - 2 \times \boxed{1} = \boxed{5}: \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & -2 \end{array}$$

Das Schema hat nun gestaffelte Form und kann schrittweise aufgelöst werden:

$$\begin{array}{l} \boxed{5}: \\ \boxed{4}: \\ \boxed{1}: \end{array} \begin{array}{l} z = -2 \\ -y - z = 1 \Rightarrow -y + 2 = 1 \Rightarrow y = 1 \\ 2x + y + z = 1 \Rightarrow 2x + 1 - 2 = 1 \Rightarrow x = 1 \end{array}$$

Die Lösung ist also $\vec{x} = [1, 1, -2]^T$.

2.8 Beispiel [\rightarrow **2.2**]: Das Schema zu dem angegebenen LGS hat die Form

$$\begin{array}{l} \boxed{1}: \\ \boxed{2}: \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 2 \end{array}$$

Durch Linearkombination kann der erste Koeffizient der zweiten Zeile zu Null gemacht werden:

$$\begin{array}{l} \boxed{1}: \\ 2 \times \boxed{2} - 3 \times \boxed{1} = \boxed{4}: \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \end{array}$$

In der letzten Zeile kann entweder der Wert von y oder der Wert von z frei gewählt werden. Wir setzen z.B. $z = t$ für eine beliebige Zahl $t \in \mathbb{R}$ und erhalten damit

$$\begin{array}{l} \boxed{4}: \\ \boxed{1}: \end{array} \begin{array}{l} -y - z = 1 \Rightarrow -y - t = 1 \Rightarrow y = -1 - t \\ 2x + y + z = 1 \Rightarrow 2x + (-1 - t) + t = 1 \Rightarrow x = 1 \end{array}$$

Die Lösungsmenge ist also die Gerade

$$\vec{x} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

2.9 Beispiel [\rightarrow **2.3**]: Das LGS hat hier die Form

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \\ \boxed{3} : \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{array}$$

Elimination der Einträge in der ersten Spalte mittels Linearkombination ergibt

$$\begin{array}{l} \boxed{1} : \\ 2 \times \boxed{2} - 3 \times \boxed{1} = \boxed{4} : \\ 2 \times \boxed{3} - 1 \times \boxed{1} = \boxed{5} : \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & 1 & 5 \end{array}$$

Elimination in der zweiten Spalte ergibt die gestaffelte Form

$$\begin{array}{l} \boxed{1} : \\ \boxed{4} : \\ \boxed{5} + \boxed{4} = \boxed{6} : \end{array} \begin{array}{ccc|c} x & y & z & \vec{b} \\ 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 6 \end{array}$$

Aus der letzten Zeile ergibt sich der Widerspruch

$$0x + 0y + 0z = 6.$$

Es existiert also keine Lösung.

2.10 Gestaffelte Form: Wie in den Beispielen zuvor gesehen, lässt sich die Lösung eines LGS einfach bestimmen, indem man es durch elementare Umformungen auf *gestaffelte Form* bringt:

$$\begin{array}{cccccc|c} \tilde{x}_1 & \tilde{x}_2 & \cdots & \tilde{x}_r & \tilde{x}_{r+1} & \cdots & \tilde{x}_n & \vec{b} \\ \bullet & * & \cdots & * & * & \cdots & * & * \\ 0 & \bullet & \cdots & * & * & \cdots & * & * \\ 0 & 0 & \cdots & * & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \bullet & * & \cdots & * & * \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \times \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \times \end{array}$$

Dabei sind

- $\tilde{x}_1, \dots, \tilde{x}_n$ eine Umordnung der gesuchten Komponenten x_1, \dots, x_n , die durch Spaltenvertauschungen entsteht,
- alle mit \bullet markierten Einträge von Null verschieden,
- alle mit $*$ markierten Einträge beliebig,
- alle mit \times markierten Einträge beliebig.

Die Existenz von Lösungen hängt von den mit \times markierten Einträgen ab:

- Wenn ein einziger dieser Einträge von Null verschieden ist, dann besitzt das LGS keine Lösung.
- Wenn es keine Nullzeilen und damit keine derartigen Einträge gibt oder wenn alle diese Einträge gleich Null sind, dann existieren Lösungen. Diese sind wie folgt gegeben: Die Werte von $\tilde{x}_{r+1}, \dots, \tilde{x}_n$ können beliebig vorgegeben werden,

$$\tilde{x}_{r+1} = t_1, \quad \dots, \quad \tilde{x}_n = t_{n-r}, \quad t_1, \dots, t_{n-r} \in \mathbb{R}.$$

Davon ausgehend können der Reihe nach die Werte von $\tilde{x}_r, \tilde{x}_{r-1}, \dots, \tilde{x}_1$ bestimmt werden.

2.11 Gauss-Algorithmus: Der *Gauss-Algorithmus* gibt Regeln an, mit denen ein LGS auf gestaffelte Form gebracht werden kann:

1. Suche ein Element $a_{i,j} \neq 0$. Vertausche die erste mit der j -ten Spalte und vertausche die erste mit der i -ten Zeile.
2. Ersetze alle Zeilen mit Index $i \geq 2$ durch die Linearkombination

$$a_{1,1} \times \boxed{i} - a_{i,1} \times \boxed{1}.$$

Damit haben die erste Zeile und die erste Spalte die gewünschte Form. Sie werden im weiteren Verlauf des Algorithmus nicht mehr verändert. Nun wendet man das Verfahren analog auf die zweite Zeile und die zweite Spalte an, wobei zu beachten ist, dass die erste Zeile nicht mehr für Zeilenvertauschungen verwendet werden darf. So verfährt man weiter, bis die gestaffelte Form erreicht ist.

2.12 Beispiel: Für einen reellen Parameter $\alpha \in \mathbb{R}$ sei das folgende LGS gegeben:

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \\ \boxed{3} : \\ \boxed{4} : \end{array} \begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \vec{b} \\ \hline 1 & 1 & 3 & -2 & 4 & -1 \\ -1 & -1 & -3 & 2 & -4 & \alpha \\ 0 & 0 & -2 & 1 & 0 & 3 \\ 0 & 0 & 4 & -2 & 0 & -6 \end{array}$$

Elimination der Einträge in der ersten Spalte mittels Linearkombination ergibt

$$\begin{array}{l} \boxed{1} : \\ \boxed{5} : \\ \boxed{3} : \\ \boxed{4} : \end{array} \begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \vec{b} \\ \hline 1 & 1 & 3 & -2 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & \alpha - 1 \\ 0 & 0 & -2 & 1 & 0 & 3 \\ 0 & 0 & 4 & -2 & 0 & -6 \end{array}$$

Um die zweite Zeile in die gewünschte Form zu bringen, wird in den Zeilen $\boxed{5}$, $\boxed{3}$, $\boxed{4}$ ein von Null verschiedener Eintrag gesucht. Wir wählen z.B. den Eintrag $a_{3,4} = 1$. Vertauschung der zweiten und der vierten Spalte sowie der Zeilen $\boxed{5}$ und $\boxed{3}$ ergibt

$$\begin{array}{l} \boxed{1} : \\ \boxed{3} : \\ \boxed{5} : \\ \boxed{4} : \end{array} \begin{array}{ccccc|c} x_1 & x_4 & x_3 & x_2 & x_5 & \vec{b} \\ \hline 1 & -2 & 3 & 1 & 4 & -1 \\ 0 & 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & \alpha - 1 \\ 0 & -2 & 4 & 0 & 0 & -6 \end{array}$$

Nun werden die Einträge der Zeilen $\boxed{5}$, $\boxed{4}$ in der zweiten Spalte zu Null gemacht:

$$\begin{array}{l} \boxed{1} : \\ \boxed{3} : \\ \boxed{5} : \\ \boxed{6} : \end{array} \begin{array}{ccccc|c} x_1 & x_4 & x_3 & x_2 & x_5 & \vec{b} \\ \hline 1 & -2 & 3 & 1 & 4 & -1 \\ 0 & 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & \alpha - 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Damit ist die gestaffelte Form erreicht. Es ist $r = 2$, und die Umordnung der Lösungskomponenten ist hier

$$\tilde{x}_1 = x_1, \quad \tilde{x}_2 = x_4, \quad \tilde{x}_3 = x_3, \quad \tilde{x}_4 = x_2, \quad \tilde{x}_5 = x_5.$$

Nun sind zwei Fälle zu unterscheiden:

- Wenn $\alpha \neq 1$, dann gibt es keine Lösung.
- Wenn $\alpha = 1$, dann gibt es einen Lösungsraum mit $n - r = 3$ freien Parametern,

$$x_3 = t_1, \quad x_2 = t_2, \quad x_5 = t_3.$$

Durch Einsetzen in die Zeilen $\boxed{3}$ und $\boxed{1}$ erhält man schließlich die Lösung

$$\vec{x} = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad t_1, t_2, t_3 \in \mathbb{R}.$$

2.13 Homogene LGS: Ein LGS $A\vec{x} = \vec{b}$ heißt *homogen*, wenn die rechte Seite der Nullvektor ist und anderenfalls *inhomogen*. Ein homogenes LGS besitzt stets mindestens eine Lösung, nämlich den Nullvektor. Betrachtet man die gestaffelte Form, dann sind alle mit \times markierten Einträge Null. Man kann also die Werte

$$\tilde{x}_{r+1} = t_1, \quad \dots, \quad \tilde{x}_n = t_{n-r}, \quad t_1, \dots, t_{n-r} \in \mathbb{R}$$

beliebig vorgeben und erhält somit eine Lösungsmenge mit $(n - r)$ freien Parametern. Diese bezeichnet man auch als *Kern von A* und schreibt dafür

$$\ker A := \{\vec{x} : A\vec{x} = \vec{0}\}.$$

Die Zahl $\dim \ker A := n - r$ wird *Dimension* des Kerns genannt. Die Zahl r , also die Anzahl der von Null verschiedenen Zeilen in der gestaffelten Form wird als *Rang* von A bezeichnet und man schreibt $\text{rang } A := r$. Es gilt also

$$\dim \ker A + \text{rang } A = n,$$

d.h., die Dimension des Kerns und der Rang der Matrix ergeben zusammen die Spaltenzahl.

2.14 Beispiel [\rightarrow 2.12]: Sei

$$A = \begin{bmatrix} 1 & 1 & 3 & -2 & 4 \\ -1 & -1 & -3 & 2 & -4 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 4 & -2 & 0 \end{bmatrix},$$

dann erhält man für das homogene LGS $A\vec{x} = \vec{0}$ die gestaffelte Form

$$\begin{array}{l} \boxed{1} : \\ \boxed{3} : \\ \boxed{5} : \\ \boxed{6} : \end{array} \begin{array}{c} x_1 \quad x_4 \quad x_3 \quad x_2 \quad x_5 \quad \vec{b} \\ \hline 1 \quad -2 \quad 3 \quad 1 \quad 4 \quad 0 \\ 0 \quad 1 \quad -2 \quad 0 \quad 0 \quad 0 \\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \end{array} .$$

Hier ist wie zuvor $r = 2$ und damit

$$\text{rang } A = 2 \quad \text{und} \quad \dim \ker A = 3.$$

Mit

$$\vec{x}_1 := \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{x}_2 := \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{x}_3 := \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

ist der Kern von A gegeben durch

$$\ker A = \{t_1\vec{x}_1 + t_2\vec{x}_2 + t_3\vec{x}_3, t_1, t_2, t_3 \in \mathbb{R}\}.$$

2.15 Inhomogene LGS: Sei \vec{x}_s eine Lösung des LGS $A\vec{x} = \vec{b}$ und $\vec{x}_h \in \ker A$ eine Lösung des zugehörigen homogenen Systems, dann ist auch $\vec{x} := \vec{x}_s + \vec{x}_h$ eine Lösung. Sind umgekehrt \vec{x} und \vec{x}_s Lösungen von $A\vec{x} = \vec{b}$, dann ist $\vec{x}_h := \vec{x} - \vec{x}_s \in \ker A$ eine Lösung des homogenen Systems. Man kann also jede Lösung von $A\vec{x} = \vec{b}$ in der Form

$$\vec{x} = \vec{x}_s + \vec{x}_h, \quad \vec{x}_h \in \ker A,$$

darstellen. Mit anderen Worten gilt: Die allgemeine Lösung eines inhomogenen Systems erhält man als Summe einer speziellen Lösung dieses Systems und der allgemeinen Lösung des zugehörigen homogenen Systems. Dieser grundlegende Sachverhalt wird als *Superpositionsprinzip* bezeichnet.

2.16 Beispiel [\rightarrow **2.12**]: Sei speziell $\alpha = 1$. Man rechnet leicht nach, dass z.B.

$$\vec{x}_s := \begin{bmatrix} 0 \\ 0 \\ -5 \\ -7 \\ 0 \end{bmatrix}$$

das gegebene inhomogene LGS

$$\begin{bmatrix} 1 & 1 & 3 & -2 & 4 \\ -1 & -1 & -3 & 2 & -4 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 4 & -2 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} -1 \\ 1 \\ 3 \\ -6 \end{bmatrix}$$

löst. Zusammen mit dem in Beispiel 2.14 bestimmten Kern von A erhält man somit die Lösungsmenge

$$\vec{x} = \vec{x}_s + t_1 \vec{x}_1 + t_2 \vec{x}_2 + t_3 \vec{x}_3, \quad t_1, t_2, t_3 \in \mathbb{R}.$$

Diese Darstellung unterscheidet sich von der in Beispiel 2.12 angegebenen Form. Die Gesamtheit der Lösungen ist aber in beiden Fällen genau dieselbe. Dies sieht man, indem man in der hier angegebenen Lösung den freien Parameter t_1 durch $t_1 + 5$ ersetzt.

2.17 Determinante: Sei A eine $(n \times n)$ -Matrix. Dann kann man die eindeutige Lösbarkeit des LGS $A\vec{x} = \vec{b}$ mit Hilfe der *Determinante* von A entscheiden. Die Determinante ist eine reelle Zahl, die wie folgt definiert ist: Wenn A eine (1×1) -Matrix ist, dann ist $\det A := a_{1,1}$. Anderenfalls gilt

$$\det A := \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det A_{i,j}.$$

Dabei ist i ein beliebiger Zeilenindex und $A_{i,j}$ eine $(n-1) \times (n-1)$ -Matrix, die durch Streichen der i -ten Zeile und der j -ten Spalte entsteht. Damit ist die Berechnung der Determinante auf ein Problem niedrigerer Dimension zurückgeführt und wiederholte Anwendung führt schließlich auf Determinanten von Matrizen der Dimension (1×1) . Anstelle der oben angegebenen Formel, die man auch *Entwicklung nach der i -ten Zeile* nennt, kann man auch nach der j -ten Spalte entwickeln,

$$\det A := \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det A_{i,j}.$$

Es gilt: Das quadratische LGS $A\vec{x} = \vec{b}$ ist genau dann eindeutig lösbar, wenn $\det A \neq 0$. Äquivalent hierzu sind die Aussagen

$$\det A \neq 0 \Leftrightarrow \dim \ker A = 0 \Leftrightarrow \text{rang } A = n.$$

2.18 Spezialfälle:

- $n = 2$:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \det A = ad - bc.$$

- $n = 3$: Entwicklung nach der ersten Zeile ergibt

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \quad \det A = a(ei - hf) - b(di - gf) + c(dh - ge).$$

Alternativ verwendet man die *Regel von Sarrus*.

- Wenn A eine obere oder untere *Dreiecksmatrix* ist, also

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n,n} \end{bmatrix} \quad \text{oder} \quad A = \begin{bmatrix} a_{1,1} & 0 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{bmatrix},$$

dann ist $\det A$ das Produkt der Diagonalelemente,

$$\det A = a_{1,1}a_{2,2} \cdots a_{n,n}.$$

2.19 Beispiel:

-

$$\det \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} = 2$$

-

$$\det \begin{bmatrix} 1 & 3 & 2 \\ 2 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix} = 9$$

-

$$\det \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix} = 1 \cdot 3 \cdot 4 \cdot 2 = 24$$

3 Matrizenrechnung

3.1 Transponierter Vektor: Die Notation $\vec{x} \in \mathbb{R}^n$ bezieht sich per Definition¹ immer auf einen stehenden Vektor,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Der *transponierte Vektor* \vec{x}^T ist das zugehörige liegende Zahlenschema

$$\vec{x}^T := [x_1, x_2, \dots, x_n],$$

das man auch als *liegenden Vektor* bezeichnet. Die Transposition eines liegenden Vektors ergibt wieder einen stehenden Vektor,

$$\vec{x} = (\vec{x}^T)^T = [x_1, x_2, \dots, x_n]^T.$$

3.2 Matrix: Ein Zahlenschema der Form

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

heißt $(m \times n)$ -Matrix. Im Fall $n = m$ heißt die Matrix *quadratisch*. Die Einträge $a_{i,j} \in \mathbb{R}$ heißen *Elemente* der Matrix. Analog zur Schreibweise $\vec{x} \in \mathbb{R}^n$ für Vektoren schreiben wir $A \in \mathbb{R}^{m \times n}$ für $(m \times n)$ -Matrizen. Die *Spalten* der Matrix A sind Vektoren in \mathbb{R}^m ,

$$A = [\vec{a}^1, \vec{a}^2, \dots, \vec{a}^n], \quad \vec{a}^j := \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}.$$

Die *Zeilen* der Matrix A sind liegende Vektoren in \mathbb{R}^n ,

$$A = \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_m^T \end{bmatrix}, \quad \vec{a}_i^T := [a_{i,1}, a_{i,2}, \dots, a_{i,n}].$$

3.3 Beispiel: Die (3×4) -Matrix $A \in \mathbb{R}^{3 \times 4}$ sei gegeben durch

$$A = \begin{bmatrix} 3 & 1 & 4 & 5 \\ 1 & 5 & 2 & 0 \\ 0 & 7 & 3 & 2 \end{bmatrix}.$$

Dann ist

$$a_{3,2} = 7, \quad \vec{a}^3 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{a}_2^T = [1, 5, 2, 0], \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 5 \\ 2 \\ 0 \end{bmatrix}.$$

¹Diese Definition bezieht sich auf das vorliegende Skript und ist keineswegs allgemeingültig.

3.4 Addition und Skalarmultiplikation: Seien $A, B \in \mathbb{R}^{m \times n}$ zwei Matrizen gleicher Dimension, dann gilt Folgendes:

- *Addition, Subtraktion:* $C := A \pm B$ ist eine $(m \times n)$ -Matrix mit Elementen

$$c_{i,j} = a_{i,j} \pm b_{i,j}.$$

- *Skalarmultiplikation:* Für $\alpha \in \mathbb{R}$ ist $C := \alpha A$ eine $(m \times n)$ -Matrix mit Elementen

$$c_{i,j} = \alpha a_{i,j}.$$

Insbesondere ist $1A = A$, $(-1)A = -A$ und

$$0A = 0_{m,n} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

die $(m \times n)$ -Nullmatrix. Wenn die Dimension der Nullmatrix aus dem Zusammenhang klar ist, schreiben wir anstelle von $0_{m,n}$ auch einfach 0.

- *Distributivgesetze:* Für $\alpha, \beta \in \mathbb{R}$ gilt

$$(\alpha + \beta)A = \alpha A + \beta A, \quad \alpha(A + B) = \alpha A + \alpha B.$$

3.5 Matrizenprodukt: Sei A eine $(m \times n)$ -Matrix und B eine $(n \times k)$ -Matrix. Dann ist das *Matrizenprodukt* $C := A \cdot B$ eine $(m \times k)$ -Matrix, die durch

$$c_{i,j} = \langle \vec{a}_i, \vec{b}^j \rangle = \sum_{j=1}^n a_{i,j} b_{j,k}$$

definiert ist. Das Element $c_{i,j}$ ist also das Skalarprodukt der i -ten Zeile von A mit der j -ten Spalte von B . Insbesondere macht das Matrizenprodukt nur dann Sinn, wenn die Spaltenzahl von n mit der Zeilenzahl von B übereinstimmt, da anderenfalls das Skalarprodukt nicht erklärt ist. Der Mal-Punkt wird meist weggelassen, wenn Verwechslungen ausgeschlossen sind, also $AB = A \cdot B$.

3.6 Beispiel: Für

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 3 \\ -4 & 1 & 2 \end{bmatrix}$$

ist

$$AB = \begin{bmatrix} -3 & 3 & 14 \\ 6 & -1 & 4 \end{bmatrix}, \quad B \cdot B = B^2 = \begin{bmatrix} 1 & 3 & 7 \\ -10 & 4 & 7 \\ -15 & -2 & 3 \end{bmatrix}.$$

Die Produkte $A \cdot A$ und $B \cdot A$ sind nicht definiert.

3.7 Vektoren als spezielle Matrizen: Ein Vektor $\vec{x} \in \mathbb{R}^n$ kann als $(n \times 1)$ -Matrix aufgefasst werden, also als Matrix mit nur einer Spalte. Genauso kann der liegende Vektor \vec{x}^T als $(1 \times n)$ -Matrix aufgefasst werden, also als Matrix mit nur einer Zeile. In diesem Sinne ist für $A \in \mathbb{R}^{m \times n}$ das Matrizenprodukt $\vec{b} = A \cdot \vec{x}$ ein Vektor in \mathbb{R}^m mit Komponenten

$$b_i = \langle \vec{a}_i, \vec{x} \rangle = \sum_{j=1}^n a_{i,j} x_j.$$

Dies erklärt insbesondere die in 2.4 eingeführte Schreibweise für lineare Gleichungssysteme. Man kann den Vektor \vec{b} auch als Linearkombination der Spalten von A deuten,

$$\vec{b} = A\vec{x} = \sum_{j=1}^n \vec{a}^j x_j.$$

Analog ist für $\vec{y} \in \mathbb{R}^m$ das Matrizenprodukt $\vec{c}^T = \vec{y}^T A$ ein liegender Vektor mit Komponenten

$$c_j = \langle \vec{y}, \vec{a}^j \rangle,$$

der auch als Linearkombination der Zeilen von A gedeutet werden kann,

$$\vec{c}^T = \vec{y}^T A = \sum_{i=1}^m \vec{a}_i^T y_i.$$

Sind $\vec{x}, \vec{z} \in \mathbb{R}^n$ zwei Vektoren gleicher Länge, dann ist das Matrizenprodukt

$$\vec{x}^T \cdot \vec{z} = \sum_{j=1}^n x_j z_j = \langle \vec{x}, \vec{z} \rangle$$

gerade das Skalarprodukt der beiden Vektoren. Für beliebige Vektoren $\vec{x} \in \mathbb{R}^n$ und $\vec{y} \in \mathbb{R}^m$ ist aber auch das Produkt

$$B = \vec{x} \cdot \vec{y}^T = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_m \end{bmatrix}$$

definiert. *Merke:*

- Liegender Vektor mal stehender Vektor ergibt eine reelle Zahl.
- Stehender Vektor mal liegender Vektor ergibt eine Matrix.

3.8 Beispiel: Für

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \vec{z} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

ist

$$A\vec{x} = \begin{bmatrix} 11 \\ 3 \end{bmatrix}, \quad \vec{y}^T A = [2 \quad 4 \quad 5], \quad \vec{x}^T \vec{z} = 5, \quad \vec{x} \vec{y}^T = \begin{bmatrix} 6 & -3 \\ 4 & -2 \\ 2 & -1 \end{bmatrix}.$$

Meist werden die Komponenten liegender Vektoren durch Kommata und Matrixelemente durch Zwischenraum getrennt. Da einzeilige Matrizen aber liegenden Vektoren entsprechen, werden hier beide Varianten verwendet. Man schreibt also auch $\vec{y}^T A = [2, 4, 5]$. Dies entspricht im Übrigen den Konventionen der Programmiersprache Matlab, bei der Matrixelemente einer Zeile entweder durch ein Leerzeichen oder durch ein Komma getrennt werden können.

3.9 Rechenregeln:

- Für die Matrizenmultiplikation gilt das *Assoziativgesetz*, d.h., es gilt

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

Nachdem die Reihenfolge der Berechnung beliebig ist, schreibt man für das Produkt auch kurz ABC .

- Das *Kommutativgesetz* gilt dagegen *nicht*, d.h., im Allgemeinen ist

$$AB \neq BA.$$

- Aus $AB = 0$ folgt *nicht* $A = 0$ oder $B = 0$.

3.10 Beispiel: Für

$$A = \begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

ist

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix}.$$

3.11 Transposition: Die *Transponierte* einer $(m \times n)$ -Matrix A ist eine $(n \times m)$ -Matrix, die mit A^T bezeichnet wird. Die Spalten von A^T sind die transponierten Zeilen von A ,

$$A = \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_m^T \end{bmatrix} \Rightarrow A^T = [\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m].$$

Es gilt

- $(A^T)^T = A$.
- $(AB)^T = B^T A^T$.
- $\langle A\vec{x}, \vec{y} \rangle = (A\vec{x})^T \cdot \vec{y} = \vec{x}^T \cdot (A^T \vec{y}) = \langle \vec{x}, A^T \vec{y} \rangle$.

Eine quadratische Matrix A heißt *symmetrisch*, wenn $A^T = A$. Es gilt

- Die Summe symmetrischer Matrizen ist symmetrisch.
- Das Produkt symmetrischer Matrizen ist im Allgemeinen *nicht* symmetrisch.
- Für beliebiges A ist sowohl $A^T + A$ als auch $A^T \cdot A$ symmetrisch.

3.12 Beispiel: Für

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

ist

$$A^T = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad A^T + A = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 10 & 6 & 1 \\ 6 & 8 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Insbesondere sind $A^T + A$ und $A^T A$ symmetrisch.

3.13 Determinante [\rightarrow 2.17]: Seien A und B zwei $(n \times n)$ -Matrizen, dann gilt

- *Vielfaches:*

$$\det(\alpha A) = \alpha^n \det A, \quad \alpha \in \mathbb{R}$$

Achtung, Exponent von α beachten!

- *Produkt:*

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

- *Transponierte:*

$$\det(A^T) = \det A$$

3.14 Matrix-Gleichungssysteme: Ein LGS der Form

$$AX = B$$

heißt auch *Matrix-Gleichungssystem*. Dabei sind $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{m \times k}$ gegeben und $X \in \mathbb{R}^{n \times k}$ gesucht. Die Bestimmung der Lösung X erfolgt vollkommen analog zum Lösen linearer Gleichungssystem gemäß Kapitel 2, indem A auf gestaffelte Form gebracht wird. Nun sind auf der rechten Seite aber alle Spalten von B umzuformen, und die Kopfzeile des Lösungsschemas enthält die Zeilenvektoren von X . Die Kriterien für die Lösbarkeit sind vollkommen analog. Insbesondere ist auch ein Matrix-Gleichungssystem mit einer quadratischen Matrix A genau dann eindeutig lösbar, wenn $\det A \neq 0$.

3.15 Beispiel: Gegeben sei das Matrix-Gleichungssystem $AX = B$ mit

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 7 \\ 3 & 1 \\ 3 & 5 \end{bmatrix}.$$

Die Lösung X ist also eine (3×2) -Matrix. Für das Schema

$$\begin{array}{c|ccc|cc} & \vec{x}_1^T & \vec{x}_2^T & \vec{x}_3^T & \vec{b}^1 & \vec{b}^2 \\ \hline \boxed{1} & 2 & 0 & 1 & 2 & 7 \\ \boxed{2} & 1 & 2 & 1 & 3 & 1 \\ \boxed{3} & 2 & 1 & 2 & 3 & 5 \end{array}$$

liefert der Gauss-Algorithmus

$$\begin{array}{c|ccc|cc} & \vec{x}_1^T & \vec{x}_2^T & \vec{x}_3^T & \vec{b}^1 & \vec{b}^2 \\ \hline \boxed{1} & 2 & 0 & 1 & 2 & 7 \\ \boxed{4} & 0 & 4 & 1 & 4 & -5 \\ \boxed{5} & 0 & 1 & 1 & 1 & -2 \\ \hline \boxed{1} & 2 & 0 & 1 & 2 & 7 \\ \boxed{4} & 0 & 4 & 1 & 4 & -5 \\ \boxed{6} & 0 & 0 & 3 & 0 & -3 \end{array}$$

Damit ergibt sich

$$\begin{aligned} \boxed{6} & : 3\vec{x}_3^T = [0, -3] \Rightarrow \vec{x}_3^T = [0, -1] \\ \boxed{4} & : 4\vec{x}_2^T + \vec{x}_3^T = [4, -5] \Rightarrow \vec{x}_2^T = [1, -1] \\ \boxed{1} & : 2\vec{x}_1^T + \vec{x}_3^T = [2, 7] \Rightarrow \vec{x}_1^T = [1, 4] \end{aligned}$$

und schließlich die Lösung

$$X = \begin{bmatrix} 1 & 4 \\ 1 & -1 \\ 0 & -1 \end{bmatrix}.$$

3.16 Einheitsmatrix: Die aus den Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_n$ [\rightarrow 1.11] gebildete $(n \times n)$ -Matrix

$$E_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n]$$

heißt *Einheitsmatrix*. Wenn die Dimension der Einheitsmatrix aus dem Zusammenhang klar ist, schreiben wir für E_n auch kurz E . Für eine beliebige $(m \times n)$ -Matrix A gilt

$$AE_n = E_m A = A.$$

Für die Determinante gilt [\rightarrow 2.18] $\det E = 1$.

3.17 Inverse Matrix: Sei A eine $(n \times n)$ -Matrix, mit $\det A \neq 0$. Dann ist die Lösung des Matrix-Gleichungssystems

$$AX = E$$

eindeutig bestimmt. Sie wird *inverse Matrix* oder auch kurz *Inverse* von A genannt und mit A^{-1} bezeichnet. Matrizen mit $\det A = 0$ oder nicht-quadratische Matrizen besitzen keine Inverse. Es gilt

- $AA^{-1} = A^{-1}A = E$
- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$
- $\det A^{-1} = \frac{1}{\det A}$, da $1 = \det E = \det(AA^{-1}) = \det A \cdot \det A^{-1}$ [\rightarrow 3.13].

Sei $AX = B$ ein beliebiges Gleichungssystem, dann erhält man nach Multiplikation von links mit A^{-1} die Lösung X ,

$$AX = B \quad \Rightarrow \quad A^{-1}AX = A^{-1}B \quad \Rightarrow \quad EX = A^{-1}B \quad \Rightarrow \quad X = A^{-1}B.$$

Die Berechnung der Inversen lohnt sich immer dann, wenn wiederholt Gleichungssysteme mit derselben Matrix A und verschiedenen rechten Seiten gelöst werden müssen.

3.18 Beispiel: Für $n = 2$ gilt

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Im Nenner steht gerade $\det A = ad - bc$, sodass die angegebene Inverse für $\det A \neq 0$ definiert ist.

3.19 Beispiel: Für

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}$$

liefert der Gauss-Algorithmus

	\vec{x}_1^T	\vec{x}_2^T	\vec{x}_3^T	\vec{e}_1	\vec{e}_2	\vec{e}_3
1	1	2	1	1	0	0
2	1	1	2	0	1	0
3	2	1	1	0	0	1
4	0	-1	1	-1	1	0
5	0	-3	-1	-2	0	1
6	0	0	-4	1	-3	1

und damit

$$A^{-1} = X = \frac{1}{4} \begin{bmatrix} -1 & -1 & 3 \\ 3 & -1 & -1 \\ -1 & 3 & -1 \end{bmatrix}.$$

3.20 Orthogonale Matrizen: Eine $(n \times n)$ -Matrix A heißt *orthogonal*, wenn

$$A^T A = E.$$

Für orthogonale Matrizen A und B gilt:

- $A^{-1} = A^T$.
- $AA^T = E$.
- A^T ist orthogonal.
- AB ist orthogonal.
- $\det A = \pm 1$, da $1 = \det E = \det(AA^T) = \det A \cdot \det A^T = (\det A)^2$.

Eine Matrix ist genau dann orthogonal, wenn ihre Spaltenvektoren ein *Orthonormalsystem* bilden, d.h., wenn die Vektoren Länge 1 haben und paarweise orthogonal sind,

$$\langle \vec{a}^i, \vec{a}^j \rangle = \delta_{i,j} := \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j. \end{cases}$$

$\delta_{i,j}$ bezeichnet man als *Kronecker-Symbol*. Es hat den Wert 1 für $i = j$ und anderenfalls den Wert 0. Die Spaltenvektoren bilden genau dann ein Orthonormalsystem, wenn auch die Zeilenvektoren ein Orthonormalsystem bilden, d.h., wenn $\langle \vec{a}_i, \vec{a}_j \rangle = \delta_{i,j}$.

3.21 Beispiel: Für beliebige Winkel φ ist die Matrix

$$A = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

orthogonal.

3.22 Beispiel: Die Matrix

$$A = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$$

ist orthogonal.

3.23 Cramer'sche Regel: Wir betrachten nochmals das lineare Gleichungssystem $A\vec{x} = \vec{b}$ der Dimension $n \times n$. Es gilt stets $A\vec{e}_2 = \vec{a}^2, \dots, A\vec{e}_n = \vec{a}^n$. Also können wir schreiben

$$A[\vec{x}, \vec{e}_2, \dots, \vec{e}_n] = A_1 \quad \text{wobei} \quad A_1 := [\vec{b}, \vec{a}^2, \dots, \vec{a}^n].$$

Die $(n \times n)$ -Matrix A_1 auf der rechten Seite entsteht also dadurch, dass man die erste Spalte von A durch \vec{b} ersetzt. Der zweite Faktor auf der rechten Seite ist ebenfalls eine $(n \times n)$ -Matrix, die untere Dreiecksform hat, siehe 2.18. Ihre Determinante ist

gleich dem Produkt der Diagonalelemente, also x_1 . Man erhält schließlich für die erste Lösungskomponente mit Hilfe der Produktregel die Formel

$$x_1 = \frac{\det A_1}{\det A},$$

sofern $\det A \neq 0$. Bezeichne allgemein A_k die $(n \times n)$ -Matrix, die entsteht, wenn man die k -te Spalte von A durch \vec{b} ersetzt, dann gilt analog

$$x_k = \frac{\det A_k}{\det A}.$$

Diese sogenannte Cramer'sche Regel ist in der Regel nur dann effizient, wenn man an einzelnen Lösungskomponenten, nicht aber am kompletten Vektor \vec{x} interessiert ist.

3.24 Orthogonalisierungsverfahren nach Gram-Schmidt: Eine $(n \times n)$ -Matrix A lässt sich stets als Produkt einer orthogonalen Matrix Q und einer oberen Dreiecksmatrix R schreiben,

$$A = QR, \quad Q = [\vec{q}^1, \dots, \vec{q}^n], \quad R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{bmatrix}.$$

Die Spaltenvektoren $\vec{q}^1, \dots, \vec{q}^n$ und die zugehörigen Spalten der Matrix R werden wie folgt berechnet: Für die k -te Spalte ergibt sich die Bedingung

$$\vec{a}^k = r_{1k}\vec{q}^1 + \cdots + r_{kk}\vec{q}^k.$$

Multipliziert man diese skalar mit \vec{q}^i für einen beliebigen Index $i < k$, dann erhält man aufgrund der Orthogonalität von Q die Koeffizienten

$$r_{ik} = \langle \vec{a}^k, \vec{q}^i \rangle. \quad (3.1)$$

Mit den so bestimmten Werten definieren wir den Hilfsvektor

$$\vec{p}^k := \vec{a}^k - \sum_{i=1}^{k-1} r_{ik}\vec{q}^i \quad (3.2)$$

uns lösen die resultierende Gleichung $\vec{p}^k = r_{kk}\vec{q}^k$ durch

$$\vec{q}^k := \frac{\vec{p}^k}{\|\vec{p}^k\|}, \quad r_{kk} := \|\vec{p}^k\|. \quad (3.3)$$

Die tatsächliche Berechnung der Matrizen Q und R erfolgt nun sukzessive:

- Man beginnt bei der ersten Spalte. Hier ist $k = 1$ und $\vec{p}^1 = \vec{a}^1$. Die Gleichung (3.3) liefert die Werte für \vec{q}^1 und r_{11} .

- Nun betrachtet man $k = 2$. Für die Bestimmung von r_{12} gemäß (3.1) ist nur der bereits bekannte Wert von \vec{q}^1 erforderlich. Also kann auch \vec{p}^2 gemäß (3.2) und damit \vec{q}^2 und r_{22} gemäß (3.3) berechnet werden.
- Für $k = 3$ lassen sich \vec{p}^3, \vec{q}^3 und alle r_{i3} berechnen, da die Werte der ersten und der zweiten Spalte bereits bekannt sind.
- Für beliebiges k verwendet man die zuvor berechneten Werte der Spalten 1 bis $k - 1$.

Mit Hilfe der QR-Zerlegung kann man das Gleichungssystem $A\vec{x} = \vec{b}$ umwandeln in

$$R\vec{x} = Q^T\vec{b}.$$

Dieses System ist einfach zu lösen, da R gestaffelte Form hat. Die QR-Zerlegung ist insbesondere auch bei der näherungsweise Lösung überbestimmter Gleichungssysteme nützlich (Stichwort „*Ausgleichsrechnung*“).

4 Lineare Abbildungen

4.1 Lineare Abbildung: Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt *lineare Abbildung* von \mathbb{R}^n nach \mathbb{R}^m , wenn für alle \vec{x}_1, \vec{x}_2 und alle $\alpha \in \mathbb{R}$ gilt

$$\begin{aligned} f(\alpha x_1) &= \alpha f(x_1) \\ f(x_1 + x_2) &= f(x_1) + f(x_2). \end{aligned}$$

Im Fall $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ nennen wir f auch eine lineare Abbildung *in* \mathbb{R}^n .

4.2 Beispiel: Für einen gegebenen Vektor $\vec{a} \in \mathbb{R}^n$ ist die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle$$

linear.

4.3 Beispiel: Für eine gegebene Matrix $A \in \mathbb{R}^{m \times n}$ ist die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit

$$f(\vec{x}) = A\vec{x}$$

linear.

4.4 Matrixform: Seien $\vec{e}_1, \dots, \vec{e}_n$ die Einheitsvektoren in \mathbb{R}^n dann kann jeder Vektor $\vec{x} \in \mathbb{R}^n$ in der Form

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^n x_j \vec{e}_j$$

dargestellt werden [\rightarrow 1.11]. Für den Funktionswert der linearen Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ an der Stelle \vec{x} gilt dann

$$f(\vec{x}) = f\left(\sum_{j=1}^n x_j \vec{e}_j\right) = \sum_{j=1}^n x_j f(\vec{e}_j).$$

Er ist also durch die Funktionwerte $f(\vec{e}_1), \dots, f(\vec{e}_n)$ der Einheitsvektoren vollständig bestimmt. Verwenden wir diese Funktionswerte als Spaltenvektoren einer $(m \times n)$ -Matrix A , d.h.,

$$A = [\vec{a}^1, \vec{a}^2, \dots, \vec{a}^n], \quad \vec{a}^j := f(\vec{e}_j),$$

dann gilt

$$A\vec{x} = A \sum_{j=1}^n x_j \vec{e}_j = \sum_{j=1}^n x_j A\vec{e}_j = \sum_{j=1}^n x_j \vec{a}^j = \sum_{j=1}^n x_j f(\vec{e}_j) = f(\vec{x}).$$

Eine lineare Abbildung f kann also stets in der Matrixform $f(\vec{x}) = A\vec{x}$ geschrieben werden. Die Matrix-Abbildungen gemäß Beispiel 4.3 umfassen also tatsächlich die Menge *aller* linearen Abbildungen.

4.5 Beispiel [\rightarrow 4.2]: Sei $\vec{a} = [a_1, \dots, a_n]^T$, dann ist $\langle \vec{a}, \vec{e}_j \rangle = a_j$. Die Matrix A ist also durch den Zeilenvektor

$$A = [f(\vec{e}_1), \dots, f(\vec{e}_n)] = [a_1, \dots, a_n] = \vec{a}^T$$

gegeben, $f(\vec{x}) = \vec{a}^T \vec{x}$.

4.6 Verkettung: Sei $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times k}$ sowie $f(\vec{x}) = A\vec{x}$ und $g(\vec{y}) = B\vec{y}$ die zugehörigen linearen Abbildungen. Dann ist die *verkettete Abbildung* $h := f \circ g : \mathbb{R}^k \rightarrow \mathbb{R}^m$ gegeben durch

$$h(\vec{y}) = f(g(\vec{y})) = f(B\vec{y}) = AB\vec{y}.$$

Die Matrizenmultiplikation entspricht also der Verkettung der zugehörigen linearen Abbildungen.

4.7 Fixpunkt: Sei A eine $(n \times n)$ -Matrix. Ein Punkt $\vec{v} \in \mathbb{R}^n$ heißt *Fixpunkt* von A , wenn $A\vec{v} = \vec{v}$. Die Menge aller Fixpunkte wird mit

$$\text{fix } A := \{\vec{v} \in \mathbb{R}^n : A\vec{v} = \vec{v}\}$$

bezeichnet. Wegen $A\vec{0} = \vec{0}$ ist der Nullvektor stets ein Fixpunkt, $\vec{0} \in \text{fix } A$. Wenn $\vec{v} \in \text{fix } A$ ein Fixpunkt von A ist, dann sind auch alle Punkte der Form $t\vec{v}$ Fixpunkte von A ,

$$\vec{v} \in \text{fix } A \quad \Rightarrow \quad t\vec{v} \in \text{fix } A, \quad t \in \mathbb{R}.$$

Wenn $\vec{v} \neq \vec{0}$, dann nennt man die Gerade $g : t\vec{v}$, $t \in \mathbb{R}$, eine *Fixpunktgerade* von A . Zur Bestimmung von Fixpunkten schreibt man

$$A\vec{v} = \vec{v} \quad \Rightarrow \quad A\vec{v} = E\vec{v} \quad \Rightarrow \quad (A - E)\vec{v} = \vec{0}.$$

Es gilt also

$$\text{fix } A = \ker(A - E).$$

4.8 Beispiel [\rightarrow 3.22]: Es gilt

$$A = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix} \quad \Rightarrow \quad A - E = \frac{1}{3} \begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 2 & -1 & -1 \end{bmatrix}$$

und die Lösung der Fixpunktgleichung $(A - E)\vec{v} = \vec{0}$ ist die Fixpunktgerade

$$\text{fix } A = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

4.9 Spezielle lineare Abbildungen in \mathbb{R}^n :

- Die *Streckung* $f(\vec{x}) = \lambda\vec{x}$ mit Streckfaktor $\lambda \in \mathbb{R}$ ist gegeben durch $A = \lambda E$. Insbesondere erhält man für $\lambda = 1$ die identische Abbildung $f(\vec{x}) = \vec{x}$, für $\lambda = 0$ die Nullabbildung $f(\vec{x}) = \vec{0}$ und für $\lambda = -1$ die Punktspiegelung $f(\vec{x}) = -\vec{x}$.

- Eine lineare Abbildung $f(\vec{x}) = A\vec{x}$ heißt *Drehung*, wenn A orthogonal ist und $\det A = 1$ gilt. Drehungen sind *normerhaltend*, d.h., $\|A\vec{x}\| = \|\vec{x}\|$, denn

$$\|A\vec{x}\|^2 = (A\vec{x})^T \cdot (A\vec{x}) = \vec{x}^T (A^T A) \vec{x} = \vec{x}^T E \vec{x} = \vec{x}^T \vec{x} = \|\vec{x}\|^2.$$

- In \mathbb{R}^2 ist eine Drehung um den Ursprung um den Winkel φ gegeben durch [→ 3.21]

$$A = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}.$$

- In \mathbb{R}^3 ist eine Drehung um die z -Achse um den Winkel φ gegeben durch

$$A = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Für allgemeine Drehmatrizen in \mathbb{R}^3 ist die Drehachse durch die Fixpunktgerade gegeben. Der Drehwinkel φ bestimmt sich gemäß der Formel

$$2 \cos \varphi + 1 = \text{spur } A,$$

wobei $\text{spur } A := a_{1,1} + a_{2,2} + a_{3,3}$ die Summe der Diagonalelemente von A ist.

- Eine lineare Abbildung $f(\vec{x}) = A\vec{x}$ heißt *Projektion*, wenn

$$A^2 = A$$

gilt. In diesem Fall ist jeder Bildpunkt $\vec{v} = A\vec{x}$ ein Fixpunkt von A , denn

$$A\vec{v} = A(A\vec{x}) = A^2\vec{x} = A\vec{x} = \vec{v}.$$

Das heißt, jeder Punkt \vec{x} wird durch einmalige Anwendung der Abbildung f auf die Menge $\text{fix } A$ abgebildet und bleibt bei weiteren Anwendungen der Abbildung dann unverändert. Die Projektion heißt *senkrecht*, wenn

$$\langle A\vec{x} - \vec{x}, A\vec{x} \rangle = 0$$

für alle $\vec{x} \in \mathbb{R}^n$. Eine Projektion ist genau dann senkrecht, wenn A symmetrisch ist.

- Die senkrechte Projektion auf die Ursprungsgerade $g : t\vec{v}$, $t \in \mathbb{R}$, ist durch

$$A_g := \frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2}$$

gegeben. Insbesondere ist $d(\vec{x}, g) = \|A\vec{x} - \vec{x}\|$ der Abstand des Punktes \vec{x} von der Geraden g [→ 1.14].

- Sei $M : \langle \vec{x}, \vec{n} \rangle = 0$ eine implizit gegebene Menge, also z.B. eine Gerade in \mathbb{R}^2 oder eine Ebene in \mathbb{R}^3 . Die senkrechte Projektion auf M ist durch

$$A_M := E - \frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2}$$

gegeben. Insbesondere ist $d(\vec{x}, M) = \|A\vec{x} - \vec{x}\|$ der Abstand des Punktes \vec{x} von der Menge M [→ 1.18], [→ 1.24].

- Eine lineare Abbildung $g(\vec{x}) = B\vec{x}$ heißt *Spiegelung*, wenn

$$B^2 = E$$

gilt. Zweimaliges Spiegeln führt also auf den Ausgangspunkt zurück. Wenn A eine Projektion ist, dann ist

$$B := 2A - E$$

eine Spiegelung an der Menge $\text{fix } A = \text{fix } B$, denn

$$B^2 = (2A - E) \cdot (2A - E) = 4A^2 - 2AE - 2EA + E^2 = E$$

und

$$B\vec{v} = \vec{v} \Leftrightarrow 2A\vec{v} - \vec{v} = \vec{v} \Leftrightarrow A\vec{v} = \vec{v}.$$

Umgekehrt ist für eine Spiegelung B die Matrix

$$A := \frac{1}{2}(B + E)$$

eine Projektion, denn

$$A^2 = \frac{1}{4}(B + E) \cdot (B + E) = \frac{1}{4}(B^2 + BE + EB + E^2) = \frac{1}{4}(2B + 2E) = A.$$

- Die Spiegelung an der Geraden $g : t\vec{v}, t \in \mathbb{R}$, ist gegeben durch

$$B_g := 2A_g - E = 2 \frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2} - E.$$

- Die Spiegelung an der Menge $M : \langle \vec{x}, \vec{n} \rangle = 0$ ist gegeben durch

$$B_M := 2A_M - E = E - 2 \frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2}.$$

Diese Abbildung wird auch *Householder-Transformation* genannt.

4.10 Beispiel [\rightarrow 3.22]: Die Matrix

$$A = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$$

ist orthogonal und es gilt $\det A = 1$. Also ist A eine Drehung. Gemäß Beispiel 4.8 ist die Drehgerade gegeben durch $g : t[1, 1, 1]^T, t \in \mathbb{R}$, und für den Drehwinkel gilt

$$2 \cos \varphi + 1 = 2 \Rightarrow \cos \varphi = 1/2 \Rightarrow \varphi = \pm\pi/3.$$

Das Vorzeichen des Drehwinkels hängt davon ab, aus welcher Richtung man auf die Drehachse schaut.

4.11 Beispiel: Sei $\vec{v} = [2, 1]^T$. Die senkrechte Projektion auf die Gerade $g : t\vec{v}, t \in \mathbb{R}$, ist

$$A_g = \frac{1}{5} \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

und die Spiegelung an der Geraden ist

$$B_g = 2A - E = \frac{1}{5} \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}.$$

4.12 Beispiel: Sei $\vec{n} = [1, 2, -1]^T$. Die senkrechte Projektion auf die Ebene $M : \langle \vec{x}, \vec{n} \rangle = 0$ ist

$$A_M = \frac{1}{6} \begin{bmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

und die Spiegelung an der Ebene ist

$$B_M = 2A_M - E = \begin{bmatrix} 2 & -2 & 1 \\ -2 & -1 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

4.13 Basiswechsel: Ein Vektor $\vec{x} \in \mathbb{R}^n$ kann als Linearkombination der Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_n$ dargestellt werden [\rightarrow 1.11],

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^n x_j \vec{e}_j.$$

Die Komponenten x_1, \dots, x_n bezeichnen wir als die *kartesischen Koordinaten* von \vec{x} . Wir versuchen nun, \vec{x} als Linearkombination eines anderen Systems $\vec{v}_1, \dots, \vec{v}_n$ von n Vektoren in \mathbb{R}^n darzustellen,

$$\vec{x} = \sum_{j=1}^n y_j \vec{v}_j = V\vec{y}.$$

Dabei ist

$$V = [\vec{v}_1, \dots, \vec{v}_n]$$

eine $(n \times n)$ -Matrix mit den Vektoren $\vec{v}_1, \dots, \vec{v}_n$ als Spalten. Bei gegebenem \vec{x} ist also der Vektor $\vec{y} = [y_1, \dots, y_n]^T$ der unbekanntem Koeffizienten gesucht. Eine Lösung existiert und ist eindeutig, wenn $\det V \neq 0$. In diesem Fall nennen wir die Vektoren $\vec{v}_1, \dots, \vec{v}_n$ eine *Basis* des \mathbb{R}^n und erhalten $\vec{y} = V^{-1}\vec{x}$. Die Komponenten y_1, \dots, y_n von \vec{y} bezeichnen wir als die *V-Koordinaten* von \vec{x} . Die Matrix V beschreibt den Übergang der *V-Koordinaten* zu den *kartesischen Koordinaten* und V^{-1} den umgekehrten Vorgang,

$$\vec{x} = V\vec{y}, \quad \vec{y} = V^{-1}\vec{x}.$$

Diesen Zusammenhang nennt man *Basiswechsel*. Sei nun $f(\vec{x}) = A\vec{x}$ eine lineare Abbildung in \mathbb{R}^n . Die Darstellung von f im *V-Koordinatensystem* hat die Form

$$\tilde{f}(\vec{y}) = \tilde{A}\vec{y}, \quad \tilde{A} := V^{-1}AV.$$

Die Matrizen A und \tilde{A} heißen *äquivalent*, da sie dieselbe lineare Abbildung (allerdings bezüglich verschiedener Koordinatensysteme) beschreiben. Durch geeignete Wahl eines Koordinatensystems kann man erreichen, dass die Matrix einer linearen Abbildung eine besonders einfache Gestalt erhält.

Bilden die Vektoren v_1, \dots, v_n speziell ein Orthonormalsystem [\rightarrow 3.20], dann ist die Matrix V orthogonal und es gilt

$$\tilde{f}(\vec{y}) = \tilde{A}\vec{y}, \quad \tilde{A} := V^T A V.$$

Für $\vec{x}_1 = V\vec{y}_1$ und $\vec{x}_2 = V\vec{y}_2$ erhalten wir dann für das Skalarprodukt

$$\langle \vec{x}_1, \vec{x}_2 \rangle = \langle V\vec{y}_1, V\vec{y}_2 \rangle = \langle \vec{y}_1, V^T V \vec{y}_2 \rangle = \langle \vec{y}_1, \vec{y}_2 \rangle.$$

In diesem Fall bleiben also unter dem Basiswechsel alle Skalarprodukte und damit auch alle Längen und Winkel erhalten.

4.14 Beispiel [\rightarrow 4.11]: Sei $\vec{v}_1 = [2, 1]^T$ und $\vec{v}_2 = [-1, 2]$, dann ist

$$V = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad V^{-1} = \frac{1}{5} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$$

Die Projektion A_g und die Spiegelung B_g haben in V -Koordinaten die Form

$$\tilde{A}_g = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{B}_g = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

4.15 Beispiel [\rightarrow 4.12]: Die Matrix

$$V = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{bmatrix}$$

ist orthogonal, $VV^T = E$. Die Drehmatrix A geht durch Basiswechsel über in

$$\tilde{A} = V^T A V = \begin{bmatrix} 1/2 & \sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Mit $\varphi = -\pi/3$ ist dies gerade die in 4.9 angegebene Form einer Drehung um die dritte Koordinatenachse. Diese ist hier durch den dritten Basisvektor $\vec{v}_3 = \sqrt{2}[1, 1, 1]^T$ gegeben und stimmt also mit der zuvor bestimmten Drehachse überein.

4.16 Beispiel [\rightarrow 4.15]: Sei V wie zuvor und

$$f(\vec{x}) := \vec{x} \times [1, 1, 1]^T.$$

Die Matrixform von f ist

$$f(\vec{x}) = [f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3)]\vec{x} = A\vec{x}, \quad A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

und man erhält

$$\tilde{A} = V^T A V = \begin{bmatrix} 0 & \sqrt{3} & 0 \\ -\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Diese Matrix lässt sich in der Form $\tilde{A} = \tilde{A}_3 \cdot \tilde{A}_2 \cdot \tilde{A}_1$ in Faktoren zerlegen, wobei

$$\tilde{A}_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{A}_2 := \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \tilde{A}_3 := \sqrt{3}E.$$

Die Abbildung \tilde{A}_1 ist eine Projektion in die $\vec{v}_1\vec{v}_2$ -Ebene, \tilde{A}_2 ist eine Drehung um die \vec{v}_3 -Achse um den Winkel $-\pi/2$, und \tilde{A}_3 ist eine Streckung um den Faktor $\sqrt{3}$. Die Abbildung f lässt sich also als Verkettung einer Projektion, einer Drehung und einer Streckung deuten.

5 Eigenwerte und -vektoren

5.1 Fixgeraden: Eine Ursprungsgerade $g : t\vec{v}, t \in \mathbb{R}$, heißt *Fixgerade* der linearen Abbildung $f(\vec{x}) = A\vec{x}$, wenn jeder Punkt auf g wieder auf einen Punkt auf g abgebildet wird. Anders als bei der *Fixpunktgeraden* müssen hier Urbild und Bild nicht übereinstimmen. Betrachten wir das Bild $f(\vec{v}) = A\vec{v}$ des Richtungsvektors. Da es wieder auf der Geraden g liegt, muss es eine Zahl $\lambda \in \mathbb{R}$ geben, sodass

$$A\vec{v} = \lambda\vec{v}.$$

Wenn dies der Fall ist, dann gilt natürlich für jeden anderen Punkt auf der Geraden die analoge Beziehung $A(t\vec{v}) = \lambda(t\vec{v})$. Der Faktor λ gibt also das Streckverhältnis zwischen Bild und Urbild für Punkte auf der Fixgeraden an. Insbesondere handelt es sich um eine Fixpunktgerade, wenn $\lambda = 1$.

5.2 Eigenwert und -vektor: Sei A eine $(n \times n)$ -Matrix. Ein Vektor $\vec{v} \neq \vec{0}$ heißt *Eigenvektor* von A zum *Eigenwert* $\lambda \in \mathbb{R}$, wenn

$$A\vec{v} = \lambda\vec{v}.$$

Wenn \vec{v} ein Eigenvektor von A ist, dann ist auch jedes Vielfache $t\vec{v}$ ein Eigenvektor, sofern $t \neq 0$. Zur Bestimmung von Eigenwerten und Vektoren betrachtet man analog zur Bestimmung von Fixpunkten [→ 4.7] die Gleichung

$$(A - \lambda E)\vec{v} = \vec{0}. \quad (5.1)$$

Die Lösbarkeit dieses LGS hängt von der Determinante

$$p(\lambda) := \det(A - \lambda E)$$

ab. Die Funktion p ist ein Polynom vom Grad n in der Variablen λ und wird als *charakteristische Polynom* von A bezeichnet. Betrachten wir nun einen festen Wert λ .

- Falls $p(\lambda) \neq 0$, dann ist das LGS (5.1) eindeutig lösbar und es folgt $\vec{v} = \vec{0}$. Folglich ist \vec{v} kein Eigenvektor und λ auch kein Eigenwert.
- Falls $p(\lambda) = 0$, dann besitzt (5.1) nichttriviale Lösungen. Es gibt also einen Vektor $\vec{v} \neq \vec{0}$ mit $(A - \lambda E)\vec{v} = \vec{0}$ und dies ist gerade ein Eigenvektor zum Eigenwert λ .

Die Eigenwerte der Matrix A sind also die Nullstellen des charakteristischen Polynoms $p(\lambda) = \det(A - \lambda E)$. Die zugehörigen Eigenvektoren bestimmen sich aus dem LGS $(A - \lambda E)\vec{v} = \vec{0}$.

5.3 Beispiel: Für

$$A = \begin{bmatrix} 5 & -9 & -3 \\ 4 & -9 & -4 \\ -6 & 15 & 8 \end{bmatrix}$$

ist das charakteristische Polynom

$$p(\lambda) = \det \begin{bmatrix} (5 - \lambda) & -9 & -3 \\ 4 & (-9 - \lambda) & -4 \\ -6 & 15 & (8 - \lambda) \end{bmatrix} = -\lambda^3 + 4\lambda^2 - \lambda - 6.$$

Wenn es ganzzahlige Nullstellen gibt, so müssen sie Teiler des konstanten Terms -6 sein. Deshalb überprüft man zunächst die Teiler von -6 : Es ist $\lambda = 1$ beispielsweise keine Nullstelle, aber $\lambda_1 := 2$ liefert $p(\lambda_1) = 0$. Nun teilt man p durch den Linearfaktor $(2 - \lambda)$ und erhält

$$q(\lambda) := p(\lambda) : (2 - \lambda) = \lambda^2 - 2\lambda - 3.$$

Die beiden restlichen Eigenwerte von A sind Nullstellen von q ,

$$q(\lambda) = 0 \quad \Rightarrow \quad \lambda_2 := 3, \quad \lambda_3 := -1.$$

Insgesamt erhält man also die Faktorisierung

$$p(\lambda) = (2 - \lambda)(3 - \lambda)(-1 - \lambda).$$

Der Eigenvektor zu λ_1 ergibt sich aus

$$(A - \lambda_1 E)\vec{v}_1 = \begin{bmatrix} 3 & -9 & -3 \\ 4 & -11 & -4 \\ -6 & 15 & 6 \end{bmatrix} \vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{v}_1 := \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Analog erhält man

$$\vec{v}_2 := \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} \quad \text{und} \quad \vec{v}_3 := \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Natürlich können auch beliebige Vielfache der angegebenen Eigenvektoren verwendet werden, z.B. $\vec{v}_2 = [0, -1, 3]^T$ oder $\vec{v}_3 = [7, 7, -7]^T$.

5.4 Regeln:

- Wenn \vec{v}_1 und \vec{v}_2 Eigenvektoren von A zum *selben* Eigenwert λ sind, dann ist auch $t_1\vec{v}_1 + t_2\vec{v}_2$ ein Eigenvektor von A zum Eigenwert λ , sofern $t_1\vec{v}_1 + t_2\vec{v}_2 \neq \vec{0}$.
- Wenn \vec{v} ein Eigenvektor von A zum Eigenwert λ ist, dann ist \vec{v} ein Eigenvektor von A^k zum Eigenwert λ^k . Falls A invertierbar ist, darf dabei auch $k = -1$ gewählt werden, d.h., die Eigenwerte der Inversen sind die Kehrwerte der Eigenwerte der gegebenen Matrix.
- Die Summe der Eigenwerte ist gleich der Summe der Diagonalelemente von A ,

$$\text{spur } A := a_{1,1} + \cdots + a_{n,n} = \lambda_1 + \cdots + \lambda_n.$$

Im Beispiel ist $\text{spur } A = 5 - 9 + 8 = 4$ und $\lambda_1 + \lambda_2 + \lambda_3 = 2 + 3 - 1 = 4$.

- Das Produkt der Eigenwerte ist gleich der Determinante von A ,

$$\det A = \lambda_1 \cdots \lambda_n.$$

Das heißt, eine Matrix ist genau dann invertierbar, wenn alle Eigenwerte von Null verschieden sind. Im Beispiel ist $\det A = -6$ und $\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = 2 \cdot 3 \cdot (-1) = -6$.

- Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.
- Die Eigenwerte (aber nicht die Eigenvektoren!) von A und A^T stimmen überein.

5.5 Spezialfälle:

- Wenn A eine obere oder untere Dreiecksmatrix ist [\rightarrow 2.18], dann stimmen die Eigenwerte mit den Diagonalelementen überein,

$$p(\lambda) = (a_{1,1} - \lambda)(a_{2,2} - \lambda) \cdots (a_{n,n} - \lambda), \quad \lambda_1 = a_{1,1}, \lambda_2 = a_{2,2}, \dots, \lambda_n = a_{n,n}.$$

Insbesondere sind die Eigenwerte der Einheitsmatrix $\lambda_1 = \dots = \lambda_n = 1$.

- Wenn A symmetrisch ist, dann sind alle Eigenwerte reell und die zugehörigen Eigenvektoren können durch geeignete Normierung so gewählt werden, dass sie ein Orthonormalsystem bilden.
- Wenn ein Eigenwert eine k -fache Nullstelle des charakteristischen Polynoms ist, dann gibt es hierzu mindestens einen und höchstens k linear unabhängige Eigenvektoren.

5.6 Beispiel:

$$A = \begin{bmatrix} 3 & 2 & -8 \\ 0 & 7 & 1 \\ 0 & 0 & -5 \end{bmatrix} \Rightarrow p(\lambda) = (3 - \lambda)(7 - \lambda)(-5 - \lambda) \Rightarrow \lambda_1 = 3, \lambda_2 = 7, \lambda_3 = -5.$$

5.7 Beispiel:

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix} \Rightarrow p(\lambda) = (1 - \lambda)(2 - \lambda)^2 \Rightarrow \lambda_1 = 1, \lambda_2 = \lambda_3 = 2.$$

Der Eigenvektor zu $\lambda_1 = 1$ ist $\vec{v}_1 = [0, 1, 1]^T$. Zu der doppelten Nullstelle $\lambda_2 = \lambda_3 = 2$ ergibt sich das Gleichungssystem

$$\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Dieses hat zwei linear unabhängige Lösungen, z.B. $\vec{v}_2 = [-1, 0, 1]^T$ und $\vec{v}_3 = [0, 1, 0]^T$.

5.8 Beispiel:

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow p(\lambda) = (1 - \lambda)^3 \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 1.$$

Zu der dreifachen Nullstelle $\lambda_1 = \lambda_2 = \lambda_3 = 1$ ergibt sich das Gleichungssystem

$$\begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Dieses hat nur eine linear unabhängige Lösung, $\vec{v} = [1, 0, 0]^T$. In Fällen wie diesem können anstelle der fehlenden Eigenvektoren sogenannte *Hauptvektoren* bestimmt werden (in der Literatur zu finden unter dem Stichwort *Jordan-Form*).

5.9 Diagonalisierung: Wenn es zu einer $(n \times n)$ -Matrix A genau n linear unabhängige Eigenvektoren $\vec{v}_1, \dots, \vec{v}_n$ gibt, dann kann man diese zu einer $(n \times n)$ -Matrix V zusammenfassen. Mit der Diagonalmatrix

$$D = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

erhält man dann

$$AV = VD \quad \text{bzw.} \quad D = V^{-1}AV \quad \text{bzw.} \quad A = VDV^{-1}.$$

Man sagt dann, dass A *diagonalisierbar* ist. Bezüglich der Basis aus Eigenvektoren [→ 4.13] ist A also ähnhlich zu einer Diagonalmatrix, $\tilde{A} = D$.

Eine weitere Anwendung der Diagonalisierung ergibt sich bei dem Problem, Potenzen von Matrizen effizient zu berechnen. Man erhält

$$A^k = A \cdots A = (VDV^{-1}) \cdots (VDV^{-1}) = VD^kV^{-1},$$

wobei $D^k = \text{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k)$.

5.10 Beispiel [→ 5.3]: Für

$$A = \begin{bmatrix} 5 & -9 & -3 \\ 4 & -9 & -4 \\ -6 & 15 & 8 \end{bmatrix}$$

erhält man

$$V = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -3 & -1 \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} 2 & -3 & -1 \\ 1 & -2 & -1 \\ -1 & 3 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Damit ist

$$\begin{aligned} A^5 &= VD^5V^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2^5 & 0 & 0 \\ 0 & 3^5 & 0 \\ 0 & 0 & (-1)^5 \end{bmatrix} \cdot \begin{bmatrix} 2 & -3 & -1 \\ 1 & -2 & -1 \\ -1 & 3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 65 & -99 & -33 \\ 244 & -489 & -244 \\ -666 & 1365 & 698 \end{bmatrix}. \end{aligned}$$

6 Folgen

6.1 Folge: Eine Abbildung

$$n \mapsto a_n,$$

die jeder natürlichen Zahl $n \in \mathbb{N}$ eine Zahl $a_n \in \mathbb{R}$ zuweist, heißt *reelle Zahlenfolge* oder auch kurz *Folge*. Man schreibt dafür $(a_n)_{n \in \mathbb{N}}$ oder auch kurz $(a_n)_n$. Man bezeichnet den Wert a_n als n -tes *Folglied* und n als den zugehörigen *Index*. Das Definitionsgebiet kann allgemeiner die Form $\{k, k+1, k+2, \dots\}$ haben, wobei $k \in \mathbb{Z}$ eine beliebige ganze Zahl ist. Man schreibt dann $(a_n)_{n \geq k}$.

Eine Folge kann auf verschiedene Weisen gegeben sein:

- Bei der *expliziten Form* lässt sich jedes Folglied a_n mittels einer Formel unmittelbar aus dem Index n berechnen, z.B.

$$a_n = \sqrt[n]{n!}, \quad n \in \mathbb{N}.$$

- Bei der *rekursiven Form* sind einige Folglieder a_1, a_2, \dots, a_m explizit gegeben. Alle weiteren Folglieder werden dann mittels einer Formel aus vorhergehenden berechnet. Das heißt, für alle $n > m$ wird a_n aus a_1, \dots, a_{n-1} berechnet, z.B.

$$a_1 = 0, \quad a_2 = 1, \quad a_n = a_{n-1} + a_{n-2}, \quad n > 2.$$

6.2 Eigenschaften: Eine Folge a heißt

- *konstant*, wenn alle Folglieder denselben Wert haben.
- *positiv/negativ*, wenn alle Folglieder positiv bzw. negativ sind.
- *alternierend*, wenn aufeinanderfolgende Folglieder verschiedenes Vorzeichen haben, d.h.,

$$a_{n+1}a_n < 0, \quad n \in \mathbb{N}.$$

- *monoton wachsend/fallend*, wenn

$$a_{n+1} \geq a_n \quad \text{bzw.} \quad a_{n+1} \leq a_n, \quad n \in \mathbb{N}.$$

- *streng monoton wachsend/fallend*, wenn

$$a_{n+1} > a_n \quad \text{bzw.} \quad a_{n+1} < a_n, \quad n \in \mathbb{N}.$$

- *beschränkt*, wenn es eine Zahl $c \in \mathbb{R}$ gibt, sodass

$$|a_n| \leq c, \quad n \in \mathbb{N}.$$

Alle Eigenschaften einer Folge können mit dem Zusatz „*fast überall (f.ü.)*“ versehen werden. Dies bedeutet, dass diese Eigenschaft erst ab einem gewissen Index n_0 durchgängig erfüllt ist und für die endlich vielen Indizes unterhalb von n_0 verletzt sein darf.

6.3 Beispiel:

- Die Folge $(a_n)_n$ mit $a_n = n^2 - 15$ ist streng monoton wachsend und f.ü. positiv.
- Die Folge $(a_n)_n$ mit $a_n = (-1)^n/n^2$ ist alternierend und beschränkt, da $|a_n| \leq 1$.
- Eine konstante Folge ist monoton wachsend, monoton fallend und beschränkt.

6.4 Nullfolge: Eine Folge $(a_n)_n$ heißt *Nullfolge*, wenn es für jedes beliebig vorgegebene $\varepsilon > 0$ eine Zahl n_0 gibt, sodass

$$|a_n| < \varepsilon \quad \text{für alle } n \geq n_0.$$

6.5 Beispiel: Gegeben sei die Folge $(a_n)_n$ mit $a_n = 1/n$. Für ein beliebiges $\varepsilon > 0$ wählt man $n_0 = 1/\varepsilon + 1$. Damit gilt für alle $n \geq n_0$

$$|a_n| = \frac{1}{n} \leq \frac{1}{n_0} = \frac{1}{1/\varepsilon + 1} < \frac{1}{1/\varepsilon} = \varepsilon.$$

Also ist $(a_n)_n$ eine Nullfolge.

6.6 Regeln: Sei $(a_n)_n$ eine Nullfolge, dann gilt:

- $(a_n)_n$ ist beschränkt.
- Sei $(b_n)_n$ eine Nullfolge, dann ist auch $(a_n \pm b_n)_n$ eine Nullfolge.
- Sei $c \in \mathbb{R}$ eine Konstante, dann ist $(a_n \cdot c)_n$ eine Nullfolge.
- Sei $(c_n)_n$ eine beschränkte Folge, dann ist $(a_n \cdot c_n)_n$ eine Nullfolge.
- Sei $(d_n)_n$ eine Folge mit $|d_n| \leq |a_n|$ f.ü., dann ist $(d_n)_n$ eine Nullfolge.

6.7 Beispiel: Gegeben sei die Nullfolge $(a_n)_n$ mit $a_n = 1/n$.

- $(6/n)_n$ ist eine Nullfolge.
- Die Folge $(c_n)_n$ mit $c_n = \sin n - \pi$ ist beschränkt durch $|c_n| \leq 5$. Also ist

$$\left(\frac{\sin n - \pi}{n} \right)_n$$

eine Nullfolge.

- Sei

$$d_n = \frac{(-1)^n 5n}{n^2 - 3},$$

dann ist $|d_n| < |6/n|$ für $n \geq 5$. Also ist $(d_n)_n$ eine Nullfolge.

- Seien p und q zwei Polynome, dann ist die Folge $(a_n)_n$ mit $a_n = p(n)/q(n)$ genau dann eine Nullfolge, wenn $\text{grad}(p) < \text{grad}(q)$.

6.8 Grenzwert: Eine Folge $(a_n)_n$ heißt *konvergent mit Grenzwert (oder Limes) a* , wenn $(a_n - a)_n$ eine Nullfolge ist. Man schreibt dann

$$\lim_{n \rightarrow \infty} a_n = a$$

und liest: „Der Grenzwert von a_n für n gegen Unendlich ist a “. Man beachte, dass das ∞ -Symbol hier nur eine Schreibweise ist. In der Definition des Grenzwerts kommen nur *endliche* Werte von n vor. Eine Folge, die nicht konvergent ist, heißt *divergent*.

6.9 Regeln: Seien $(a_n)_n$ und $(b_n)_n$ zwei konvergente Folgen mit Grenzwert a bzw. b , dann gilt:

- Die Folge $(a_n)_n$ ist beschränkt, denn $|a_n - a| \leq c$ [\rightarrow 6.6] und

$$|a_n| = |a_n - a + a| \leq |a_n - a| + |a| \leq c + a.$$

- Die Folge $(\alpha a_n + \beta b_n)_n$ ist konvergent mit

$$\lim_{n \rightarrow \infty} (\alpha a_n + \beta b_n) = \alpha a + \beta b, \quad \alpha, \beta \in \mathbb{R}.$$

- Die Folge $(a_n \cdot b_n)_n$ ist konvergent mit

$$\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b.$$

- Die Folge $(a_n/b_n)_n$ ist konvergent mit

$$\lim_{n \rightarrow \infty} (a_n/b_n) = a/b,$$

sofern $b \neq 0$. Die Folge ist divergent, falls $a \neq 0$ und $b = 0$.

6.10 Wichtige Grenzwerte:

$$\lim_{n \rightarrow \infty} \frac{p_0 + p_1 n + \cdots + p_k n^k}{q_0 + q_1 n + \cdots + q_k n^k} = \frac{p_k}{q_k}, \quad q_k \neq 0$$

$$\lim_{n \rightarrow \infty} \frac{\ln n}{n^\alpha} = 0, \quad \alpha > 0$$

$$\lim_{n \rightarrow \infty} n^\alpha q^n = 0, \quad \alpha \in \mathbb{R}, |q| < 1$$

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e \approx 2.718$$

$$\lim_{n \rightarrow \infty} \frac{n}{\sqrt[n]{n!}} = e$$

6.11 Konvergenzkriterien:

- *Cauchy*: Eine Folge $(a_n)_n$ reeller Zahlen ist genau dann konvergent, wenn es für jedes beliebig vorgegebene $\varepsilon > 0$ eine Zahl n_0 gibt, sodass

$$|a_n - a_m| < \varepsilon \quad \text{für alle } n, m \geq n_0.$$

- *Intervallschachtelung*: Sei $(a_n)_m$ monoton wachsend, $(b_n)_n$ monoton fallend und $(b_n - a_n)_n$ eine Nullfolge. Dann konvergieren beide Folgen gegen denselben Grenzwert und es gilt

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n \in [a_k, b_k], \quad k \in \mathbb{N}.$$

- Sei $(a_n)_n$ beschränkt und monoton f.ü., dann ist $(a_n)_n$ konvergent.

6.12 Bestimmte Divergenz: Eine Folge $(a_n)_n$ heißt *bestimmt divergent*, wenn sie f.ü. positiv oder f.ü. negativ ist und $(1/a_n)_n$ eine Nullfolge ist. Wir schreiben dann

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} a_n = -\infty.$$

6.13 Beispiel:

- Es gilt

$$\lim_{n \rightarrow \infty} e^n/n = \infty, \quad \lim_{n \rightarrow \infty} \ln(1/n) = -\infty.$$

- Die Folge $(1 + n + (-1)^n n)_n$ ist positiv und nicht beschränkt. Dennoch ist sie nicht bestimmt divergent, da alle ungeraden Folgenglieder der Wert 1 haben.

6.14 Vektorfolge: Eine Abbildung

$$n \mapsto \vec{a}_n,$$

die jeder natürlichen Zahl $n \in \mathbb{N}$ einen Vektor $\vec{a}_n \in \mathbb{R}^d$ zuweist, heißt *Vektorfolge* in \mathbb{R}^d . Man schreibt dafür $(\vec{a}_n)_{n \in \mathbb{N}}$ oder auch kurz $(\vec{a}_n)_n$. Die Komponenten $a_{i,n}$ von \vec{a}_n bilden die reellen Zahlenfolgen $(a_i)_n$,

$$\vec{a}_n = \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{d,n} \end{bmatrix}, \quad (\vec{a}_n)_n = \begin{bmatrix} (a_1)_n \\ (a_2)_n \\ \vdots \\ (a_d)_n \end{bmatrix}.$$

Die Vektorfolge $(\vec{a}_n)_n$ heißt *konvergent mit Grenzwert \vec{a}* , wenn alle Komponentenfolgen konvergieren,

$$\lim_{n \rightarrow \infty} \vec{a}_n = \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{bmatrix}, \quad a_i = \lim_{n \rightarrow \infty} a_{i,n}.$$

6.15 Funktionenfolge: Sei $I \subset \mathbb{R}$ ein Intervall. Eine Abbildung

$$n \mapsto f_n,$$

die jeder natürlichen Zahl $n \in \mathbb{N}$ eine Funktion $f_n : I \rightarrow \mathbb{R}$ zuweist, heißt *Funktionsfolge*. Man schreibt dafür $(f_n)_{n \in \mathbb{N}}$ oder auch kurz $(f_n)_n$. Für festes $x \in I$ bilden die Funktionswerte $f_n(x)$ die reelle Zahlenfolge $(f_n(x))_n$. Die Funktionenfolge heißt *punktweise konvergent mit Grenzfunktion* f , wenn die Zahlenfolge $(f_n(x))_n$ für alle $x \in I$ konvergiert,

$$\lim_{n \rightarrow \infty} f_n = f, \quad f(x) := \lim_{n \rightarrow \infty} f_n(x), \quad x \in I.$$

6.16 Beispiel: Für $f_n(x) = x^n$, $x \in [0, 1]$, ist die Grenzfunktion gegeben durch

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x = 1 \end{cases}.$$

6.17 Beispiel: Die durch $f_n(x) = x^n$, $x \in [-1, 1]$, gegebene Funktionenfolge ist divergent, da für $x = -1$ kein Grenzwert existiert.

6.18 Beispiel: Für $f_n(x) = (1 + x/n)^n$, $x \in \mathbb{R}$, ist die Grenzfunktion gegeben durch

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = e^x.$$

Es ist z.B.

$$f(-1) = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^n = e^{-1}, \quad f(2) = \lim_{n \rightarrow \infty} \left(1 + \frac{2}{n}\right)^n = e^2.$$

7 Reihen

7.1 Reihe: Sei $(a_n)_n$ eine Folge. Dann definiert man die zugehörige Folge $(s_m)_m$ gemäß

$$\begin{aligned} s_1 &:= a_1 \\ s_2 &:= a_1 + a_2 \\ s_3 &:= a_1 + a_2 + a_3 \\ &\vdots \\ s_m &:= a_1 + a_2 + \cdots + a_m = \sum_{n=1}^m a_n \\ &\vdots \end{aligned}$$

Man bezeichnet die Folge $(s_m)_m$ als die *Reihe* und die Folgenglieder s_m als die *Partialsummen* zur Folge $(a_n)_n$. Wenn die Reihe, also die Folge der Partialsummen, konvergiert, dann schreibt man für den Grenzwert

$$\sum_{n=1}^{\infty} a_n := \lim_{m \rightarrow \infty} s_m.$$

Vollkommen analog definiert man zur Folge $(a_n)_{n \geq k}$ die die Reihe $(s_m)_{m \geq k}$ durch die Partialsummen

$$s_m := a_k + a_{k+1} + \cdots + a_m = \sum_{n=k}^m a_n, \quad m \geq k,$$

und schreibt im Falle der Konvergenz

$$\sum_{n=k}^{\infty} a_n := \lim_{m \rightarrow \infty} s_m.$$

Gemäß den Regeln für Grenzwerte von Folgen gilt

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n, \quad \alpha, \beta \in \mathbb{R},$$

sofern die Reihen zu den Folgen $(a_n)_n$ und $(b_n)_n$ konvergieren.

7.2 Beispiel: Für ein gegebenes $q \neq 1$ sei die Folge $(a_n)_{n \geq 0}$ gegeben durch $a_n := q^n$. Die zugehörige Folge der Partialsummen

$$s_0 = 1, \quad s_1 = 1 + q, \quad s_2 = 1 + q + q^2, \dots$$

wird als *geometrische Reihe* bezeichnet. In diesem Fall lassen sich die Partialsummen explizit angeben,

$$s_m = \frac{1 - q^{m+1}}{1 - q}.$$

Die geometrische Reihe konvergiert also genau dann mit Grenzwert

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1 - q},$$

wenn $|q| < 1$. Anderenfalls ist die geometrische Reihe divergent.

7.3 Notwendige Bedingung: Sei $\sum_{n=1}^{\infty} a_n$ eine konvergente Reihe mit Grenzwert s . Dann gilt

$$a_n = s_n - s_{n-1} = (s_n - s) - (s_{n-1} - s).$$

Beide Klammern auf der rechten Seite sind Nullfolgen, also muss auch $(a_n)_n$ eine Nullfolge sein [→ 6.6]. Für die Konvergenz einer Reihe ist es also notwendig, dass $(a_n)_n$ eine Nullfolge ist. Diese Bedingung ist aber keineswegs hinreichend, d.h., selbst wenn $(a_n)_n$ eine Nullfolge ist, kann die zugehörige Reihe divergent sein.

7.4 Beispiel:

- Sei $a_n := 1/n$. Die zugehörige Folge der Partialsummen

$$s_1 = 1, \quad s_2 = 1 + 1/2, \quad s_3 = 1 + 1/2 + 1/3, \dots$$

wird als *harmonische Reihe* bezeichnet. Diese ist bestimmt divergent,

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

- Die Reihe zur Folge $a_n = 1/n^\alpha$ konvergiert genau dann, wenn $\alpha > 1$. Es ist z.B.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

7.5 Leibniz-Kriterium: Sei $(a_n)_n$ eine alternierende Nullfolge [→ 6.2] mit der Eigenschaft, dass die Folge der Beträge $(|a_n|)_n$ monoton fällt, dann konvergiert die zugehörige Reihe.

7.6 Beispiel: Sei $a_n := (-1)^{n+1}/n$. Die zugehörige Folge der Partialsummen

$$s_1 = 1, \quad s_2 = 1 - 1/2, \quad s_3 = 1 - 1/2 + 1/3, \dots$$

wird als *alternierende harmonische Reihe* bezeichnet. Sie konvergiert nach dem Leibniz-Kriterium und hat den Grenzwert

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln 2.$$

7.7 Sprachgebrauch: In der Praxis wird das Symbol $\sum_{n=1}^{\infty} a_n$, das eigentlich nur für den Grenzwert einer Reihe steht, häufig auch für die Reihe selbst verwendet. Man sagt also:

Die Reihe $\sum_{n=1}^{\infty} a_n$...

und meint damit eigentlich

Die Reihe zur Folge $(a_n)_n$...

Außerdem wird der Grenzwert einer Reihe gelegentlich auch nur als *Wert* der Reihe bezeichnet. Der Satz

Die Reihe $\sum_{n=0}^{\infty} (-1)^n/(2n+1)$ hat den Wert $\pi/4$.

ist also zu interpretieren als

Die zur Folge $(a_n)_{n \geq 0}$, $a_n = (-1)^n/(2n+1)$ gehörende Reihe hat den Grenzwert $\pi/4$.

7.8 Vergleichskriterien: Gegeben seien die Folgen $(a_n)_n, (b_n)_n$.

- *Konvergente Majorante:* Wenn $\sum_{n=1}^{\infty} a_n$ konvergiert und

$$\lim_{n \rightarrow \infty} \frac{b_n}{a_n} \text{ existiert}$$

oder

$$|b_n| \leq a_n \quad \text{f.ü.},$$

dann konvergiert auch $\sum_{n=1}^{\infty} b_n$.

- *Divergente Minorante:* Wenn $\sum_{n=1}^{\infty} a_n$ divergiert und

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} \text{ existiert}$$

oder

$$0 \leq a_n \leq b_n \quad \text{f.ü.},$$

dann divergiert auch $\sum_{n=1}^{\infty} b_n$.

7.9 Absolute Konvergenz: Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, wenn die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert. Eine absolut konvergente Reihe ist nach dem Majorantenkriterium konvergent. Eine Reihe, die konvergent, aber nicht absolut konvergent ist, heißt *bedingt konvergent*. Bei absolut konvergenten Reihen können die Folgenglieder a_n beliebig umgeordnet werden, ohne dass sich der Wert der Reihe ändert. Hier gilt also in einem gewissen Sinne das Kommutativgesetz. Bei bedingt konvergenten Reihen bewirkt eine Umordnung der Folgenglieder a_n hingegen in der Regel eine Änderung des Reihenwertes oder sogar einen Verlust der Konvergenz.

7.10 Beispiel:

- [→ 7.4] Die Reihe $\sum_{n=1}^{\infty} (-1)^n/n^2$ ist absolut konvergent.
- [→ 7.6] Die Reihe $\sum_{n=1}^{\infty} (-1)^n/n$ ist bedingt konvergent.

7.11 Quotientenkriterium: Sei

$$q_n := \left| \frac{a_{n+1}}{a_n} \right| \quad \text{und, falls konvergent,} \quad q := \lim_{n \rightarrow \infty} q_n.$$

Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, wenn $q < 1$ und divergent, wenn $q > 1$. Im Fall $q = 1$ ist keine Entscheidung möglich.

7.12 Wurzelkriterium: Sei

$$w_n := \sqrt[n]{|a_n|} \quad \text{und, falls konvergent,} \quad w := \lim_{n \rightarrow \infty} w_n.$$

Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, wenn $w < 1$ und divergent, wenn $w > 1$. Im Fall $w = 1$ ist keine Entscheidung möglich. Das Wurzelkriterium ist im folgenden Sinne stärker als das Quotientenkriterium: Wenn $\lim_{n \rightarrow \infty} q_n$ existiert, dann existiert auch $\lim_{n \rightarrow \infty} w_n$ und es gilt

$$q = \lim_{n \rightarrow \infty} q_n = \lim_{n \rightarrow \infty} w_n = w.$$

7.13 Beispiel: Sei $a_n = (-1)^n n^3 2^{-n}$. Das Quotientenkriterium liefert

$$q_n = \left| \frac{a_{n+1}}{a_n} \right| = \frac{(n+1)^3 2^{-n-1}}{n^3 2^{-n}} = \frac{(1+1/n)^3}{2}, \quad q = \lim_{n \rightarrow \infty} q_n = \frac{1}{2}.$$

Genauso liefert das Wurzelkriterium

$$w_n = \sqrt[n]{|a_n|} = \frac{\sqrt[n]{n^3}}{2}, \quad w = \lim_{n \rightarrow \infty} w_n = \frac{1}{2}.$$

Die Reihe

$$\sum_{n=0}^{\infty} \frac{(-1)^n n^3}{2^n}$$

ist also absolut konvergent.

7.14 Beispiel: Sei

$$a_n = \begin{cases} 2^{-n} & \text{für } n \text{ gerade} \\ 2^{-n+2} & \text{für } n \text{ ungerade} \end{cases}$$

Es ist also

$$\sum_{n=0}^{\infty} a_n = 1 + 2 + \frac{1}{4} + \frac{1}{2} + \frac{1}{16} + \frac{1}{8} + \dots$$

Das Quotientenkriterium liefert die divergente Folge

$$q_n = \begin{cases} 2 & \text{für } n \text{ gerade} \\ 1/8 & \text{für } n \text{ ungerade,} \end{cases}$$

während das Wurzelkriterium mit

$$w_n = \begin{cases} 1/2 & \text{für } n \text{ gerade} \\ \sqrt[n]{4}/2 & \text{für } n \text{ ungerade} \end{cases} \Rightarrow w = \lim_{n \rightarrow \infty} w_n = 1/2$$

zeigt, dass die Reihe absolut konvergent ist.

7.15 Beispiel: Für $a_n = n^n/n!$ ist

$$q_n = \frac{(n+1)^{n+1} n!}{n^n (n+1)!} = \left(1 + \frac{1}{n}\right)^n, \quad w_n = n/\sqrt[n]{n!}.$$

Zum einen folgt aus $\lim_{n \rightarrow \infty} q_n = e$ die Divergenz der zugehörigen Reihe. Zum anderen liefert die Gleichheit der Grenzwerte $w = q$ die *Stirlingsche Formel* [\rightarrow 6.10]

$$\lim_{n \rightarrow \infty} \frac{n}{\sqrt[n]{n!}} = e.$$

7.16 Beispiel: Für $a_n = x^n/n!$ ist

$$q_n = \left| \frac{x^{n+1} n!}{(n+1)! x^n} \right| = \frac{|x|}{n+1}, \quad q = \lim_{n \rightarrow \infty} q_n = 0.$$

Die zugehörige Reihe ist also für alle $x \in \mathbb{R}$ konvergent und definiert die *e-Funktion*,

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}.$$

8 Reelle Funktionen

8.1 Reelle Funktion: Eine *reelle Funktion* $f : D_f \rightarrow \mathbb{R}$ ordnet jedem Element $x \in D_f$ der Menge $D_f \subset \mathbb{R}$ eine reelle Zahl $y \in \mathbb{R}$ zu, und man schreibt

$$y = f(x), \quad x \in D.$$

- Die Menge D_f heißt *Definitionsbereich* von f .
- x heißt *Urbild von y* , und y heißt *Bild von x* .
- Die *Bildmenge* von f ist die Menge aller Bilder,

$$B_f := \{y = f(x) : x \in D_f\}.$$

8.2 Eigenschaften: Die Funktion f heißt

- *injektiv*, wenn $f(x_1) \neq f(x_2)$ für alle $x_1, x_2 \in D_f$ mit $x_1 \neq x_2$.
- *monoton wachsend/fallend*, wenn $f(x_1) \leq f(x_2)$ bzw. $f(x_1) \geq f(x_2)$ für alle $x_1, x_2 \in D_f$ mit $x_1 < x_2$.
- *streng monoton wachsend/fallend*, wenn $f(x_1) < f(x_2)$ bzw. $f(x_1) > f(x_2)$ für alle $x_1, x_2 \in D_f$ mit $x_1 < x_2$.
- *beschränkt*, wenn es eine Zahl $c \in \mathbb{R}$ gibt, sodass $|f(x)| \leq c$ für alle $x \in D_f$.
- *gerade/ungerade*, wenn $f(x) = f(-x)$ bzw. $f(x) = -f(-x)$ für alle $x \in D_f$.

8.3 Beispiel: Sei $f(x) := 1/(1 + x^2)$, $x \in D_f$.

- Für $D_f = \mathbb{R}_{\geq 0}$ ist f injektiv und $B_f = (0, 1]$.
- Für $D_f = [-1, 2)$ ist f nicht injektiv und $B_f = (1/5, 1]$.

8.4 Beispiel: Sei $f(x) := x + |x|$, $x \in D_f$.

- Für $D_f = \mathbb{R}$ ist f monoton wachsend und unbeschränkt.
- Für $D_f = [0, 3]$ ist f streng monoton wachsend und beschränkt.

8.5 Beispiel:

- Das Monom x^n ist eine gerade Funktion, wenn n gerade ist und eine ungerade Funktion, wenn n ungerade ist.

- Die Funktionen $\cos x$ und

$$\cosh x := \frac{e^x + e^{-x}}{2}$$

sind gerade.

- Die Funktionen $\sin x$ und

$$\sinh x := \frac{e^x - e^{-x}}{2}$$

sind ungerade.

8.6 Regeln:

- Aus strenger Monotonie folgt Injektivität.
- Summe und Produkt zweier gerader Funktionen sind gerade.
- Die Summe zweier ungerader Funktionen ist ungerade, aber ihr Produkt ist gerade.

8.7 Verkettung: Seien $f : D_f \rightarrow \mathbb{R}$ und $g : D_g \rightarrow \mathbb{R}$ zwei reelle Funktionen. Wenn $B_g \subset D_f$, dann ist die verkettete Funktion $h := f \circ g : D_g \rightarrow \mathbb{R}$ (lies „ f nach g “) definiert durch

$$h(x) := f(g(x)), \quad x \in D_g.$$

8.8 Beispiel: Sei $f(x) := \ln(1 - x)$, $D_f := (-\infty, 1)$ und $g(x) := \cos x$.

- Für $D_g = \mathbb{R}$ ist $B_g = [-1, 1] \not\subset D_f$. Die Verkettung $f \circ g$ ist also nicht definiert.
- Für $D_g = (0, 6)$ ist $B_g = [-1, 1) \subset D_f$ und damit

$$h(x) = f(g(x)) = \ln(1 - \cos x), \quad x \in (0, 6).$$

8.9 Regeln:

- Die Verkettung zweier monoton wachsender Funktionen ist monoton wachsend.
- Die Verkettung zweier monoton fallender Funktionen ist monoton wachsend.
- Die Verkettung zweier injektiver Funktionen ist injektiv.
- Die Verkettung zweier gerader Funktionen ist gerade.
- Die Verkettung zweier ungerader Funktionen ist ungerade.

8.10 Umkehrfunktion: Sei $f : D_f \rightarrow \mathbb{R}$ injektiv, dann gibt es zu jedem Bild $y \in B_f$ ein eindeutig bestimmtes Urbild x . Die Funktion, die jedem Bild das zugehörige Urbild zuordnet, wird *Umkehrfunktion* genannt und mit f^{-1} bezeichnet,

$$f^{-1} : B_f \rightarrow \mathbb{R}, \quad f^{-1}(y) = x \text{ für dasjenige } x \in D_f \text{ mit } f(x) = y.$$

Es gilt

- $D_{f^{-1}} = B_f$ und $B_{f^{-1}} = D_f$.
- $f^{-1}(f(x)) = x$ für alle $x \in D_f$.
- $f(f^{-1}(y)) = y$ für alle $y \in B_f$.
- $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.
- Das Schaubild von f^{-1} erhält man aus dem Schaubild von f durch Spiegelung an der ersten Winkelhalbierenden.

8.11 Beispiel:

- Die Funktion $f(x) = \sin(x)$, $D_f = [-\pi/2, \pi/2]$ ist injektiv und es gilt $B_f = [-1, 1]$. Die Umkehrfunktion f^{-1} hat also das Definitionsgebiet $D_{f^{-1}} = B_f = [-1, 1]$ und die Bildmenge $B_{f^{-1}} = D_f = [-\pi/2, \pi/2]$. Die so definierte Umkehrfunktion wird als *Arcussinus* bezeichnet und man schreibt dafür

$$\arcsin : [-1, 1] \rightarrow [-\pi/2, \pi/2].$$

Der Arcussinus ordnet also jeder Zahl zwischen -1 und 1 einen Winkel aus dem Intervall $[-\pi/2, \pi/2]$ zu, dessen Sinus dem gegebenen Argument entspricht. Es gilt beispielsweise

$$\arcsin(0) = 0, \quad \arcsin(1/2) = \pi/6, \quad \arcsin(-1) = -\pi/2,$$

denn

$$\sin(0) = 0, \quad \sin(\pi/6) = 1/2, \quad \sin(-\pi/2) = -1.$$

- Die Funktion $f(x) = \cos(x)$, $D_f = [0, \pi]$ ist injektiv und es gilt $B_f = [-1, 1]$. Die Umkehrfunktion f^{-1} hat also das Definitionsgebiet $D_{f^{-1}} = B_f = [-1, 1]$ und die Bildmenge $B_{f^{-1}} = D_f = [0, \pi]$. Die so definierte Umkehrfunktion wird als *Arcuscosinus* bezeichnet und man schreibt dafür

$$\arccos : [-1, 1] \rightarrow [0, \pi].$$

Es gilt beispielsweise

$$\arccos(1) = 0, \quad \arccos(1/2) = \pi/3, \quad \arccos(-1) = \pi,$$

denn

$$\cos(0) = 1, \quad \cos(\pi/3) = 1/2, \quad \cos(\pi) = -1.$$

- Die Funktion $f(x) = \tan(x)$, $D_f = (-\pi/2, \pi/2)$ ist injektiv und es gilt $B_f = \mathbb{R}$. Die Umkehrfunktion f^{-1} hat also das Definitionsgebiet $D_{f^{-1}} = B_f = \mathbb{R}$ und die Bildmenge $B_{f^{-1}} = D_f = (-\pi/2, \pi/2)$. Die so definierte Umkehrfunktion wird als *Arcustangens* bezeichnet und man schreibt dafür

$$\arctan : \mathbb{R} \rightarrow (-\pi/2, \pi/2).$$

Es gilt beispielsweise

$$\arctan(0) = 0, \quad \arctan(1) = \pi/4, \quad \arctan(-\sqrt{3}) = -\pi/3,$$

denn

$$\tan(0) = 0, \quad \tan(\pi/4) = 1, \quad \tan(-\pi/3) = -\sqrt{3}.$$

- Die Funktion $f(x) = e^x$, $D_f = \mathbb{R}$ ist injektiv und es gilt $B_f = \mathbb{R}_{>0}$. Die Umkehrfunktion f^{-1} hat also das Definitionsgebiet $D_{f^{-1}} = B_f = \mathbb{R}_{>0}$ und die Bildmenge $B_{f^{-1}} = D_f = \mathbb{R}$. Die so definierte Umkehrfunktion wird als *natürlicher Logarithmus* bezeichnet und man schreibt dafür

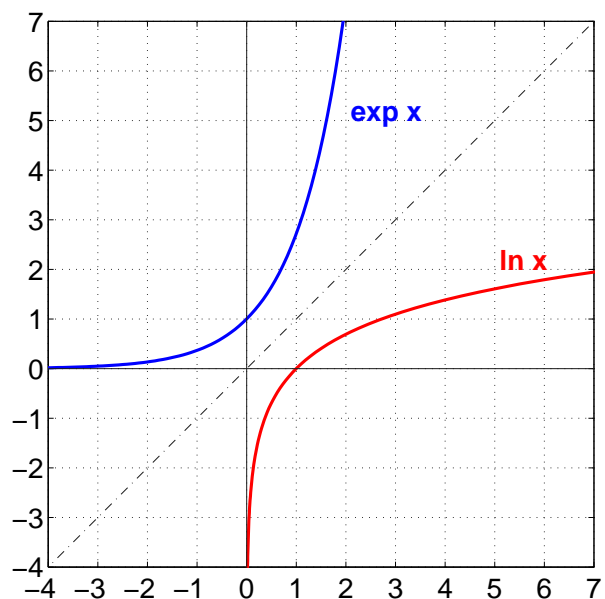
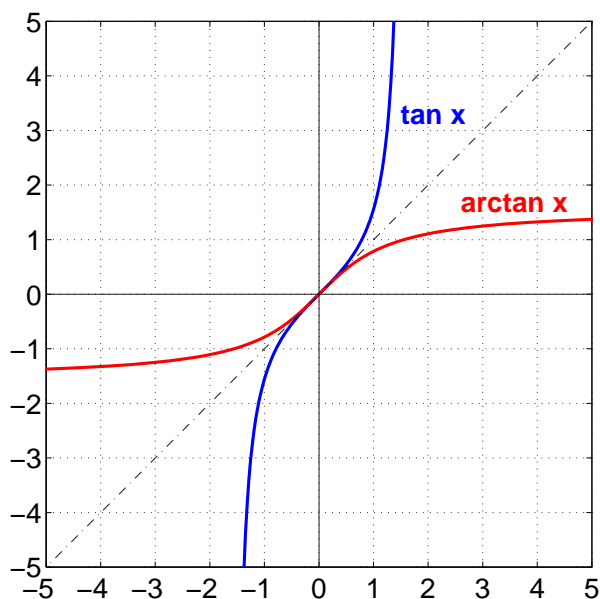
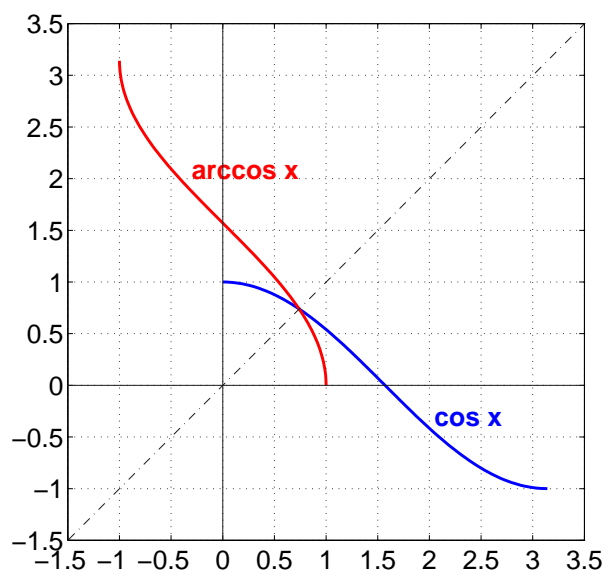
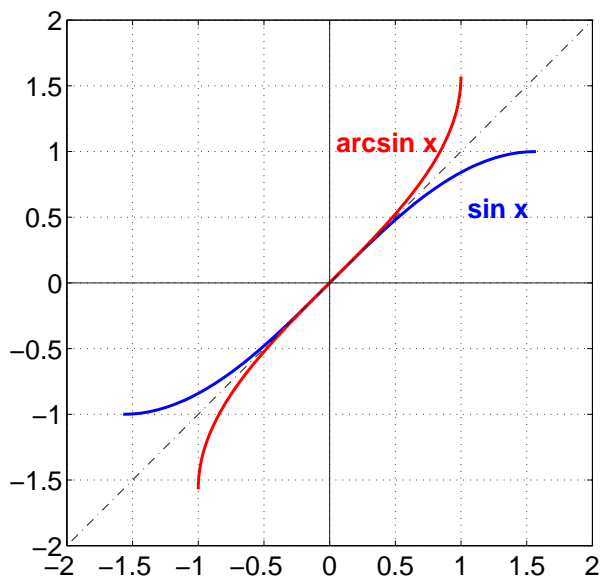
$$\ln : \mathbb{R}_{>0} \rightarrow \mathbb{R}.$$

Es gilt beispielsweise

$$\ln(1) = 0, \quad \ln(e) = 1, \quad \ln(1/e^2) = -2,$$

denn

$$e^0 = 1, \quad e^1 = e, \quad e^{-2} = 1/e^2.$$



8.12 Häufungspunkt: $x_* \in \mathbb{R}$ heißt *Häufungspunkt* der Menge $D_f \subset \mathbb{R}$, wenn es eine Folge $(x_n)_n$ gibt mit

$$\lim_{n \rightarrow \infty} x_n = x_*, \quad x_n \in D_f, \quad x_n \neq x_*, \quad \text{f.ü.}$$

8.13 Beispiel:

- $x_* = 1$ ist ein Häufungspunkt der Menge $(0, 1)$.
- $x_* = 1$ ist kein Häufungspunkt der Menge \mathbb{N} .
- $x_* = \pi$ ist ein Häufungspunkt der Menge \mathbb{Q} der rationalen Zahlen.

8.14 Grenzwert: Sei $f : D_f \rightarrow \mathbb{R}$ eine reelle Funktion. Man sagt, dass f an der Stelle x_* den *Grenzwert* f_* hat, wenn x_* ein Häufungspunkt von D_f ist und wenn für jede gegen x_* konvergente Folge $(x_n)_n$ gemäß 8.12 gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f_*.$$

Man schreibt dann

$$\lim_{x \rightarrow x_*} f(x) = f_*.$$

Wenn die Folge $(f(x_n))_n$ stets bestimmt divergent ist, dann schreibt man

$$\lim_{x \rightarrow x_*} f(x) = \infty \quad \text{bzw.} \quad \lim_{x \rightarrow x_*} f(x) = -\infty.$$

Wenn $\lim_{n \rightarrow \infty} f(x_n) = f_*$ für jede bestimmt divergente Folge $(x_n)_n$ mit $\lim_{n \rightarrow \infty} x_n = \infty$, dann schreibt man

$$\lim_{x \rightarrow \infty} f(x) = f_*.$$

Analog sind die Ausdrücke

$$\lim_{x \rightarrow -\infty} f(x) = f_*, \quad \lim_{x \rightarrow \infty} f(x) = \pm\infty, \quad \lim_{x \rightarrow -\infty} f(x) = \pm\infty$$

erklärt.

8.15 Beispiel:

•

$$\lim_{x \rightarrow 0} x \sin(1/x) = 0, \quad \lim_{x \rightarrow \infty} \arctan(x) = \pi/2, \quad \lim_{x \rightarrow 0} \ln|x| = -\infty.$$

- Die Funktion $f(x) = \sin(1/x)$, $x \in \mathbb{R}_{\neq 0}$, hat an der Stelle $x_* = 0$ keinen Grenzwert.
- Die Funktion $f(x) = x$, $x \in \mathbb{N}$, hat an der Stelle $x_* = 1$ keinen Grenzwert, da $x_* = 1$ kein Häufungspunkt von \mathbb{N} ist [\rightarrow 8.13].
- Die *Heaviside-Funktion* $H : \mathbb{R} \rightarrow \mathbb{R}$ ist definiert durch

$$H(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0. \end{cases}$$

Sie hat an der Stelle $x_* = 0$ keinen Grenzwert. Sei $f(x) := H(x) + H(-x)$, dann gilt

$$\lim_{x \rightarrow 0} f(x) = 1,$$

aber $f(0) = 2$.

8.16 Einseitiger Grenzwert: Man sagt, dass f an der Stelle x_* den *rechtsseitigen bzw. linksseitigen Grenzwert* f_* hat, wenn x_* ein Häufungspunkt von D_f ist und wenn für jede Folge $(x_n)_n$ gemäß 8.12 mit der zusätzlichen Eigenschaft

$$x_n > x_* \quad \text{bzw.} \quad x_n < x_*, \quad n \in \mathbb{N},$$

gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f_*.$$

Man schreibt dann

$$\lim_{x \downarrow x_*} f(x) = f_* \quad \text{bzw.} \quad \lim_{x \uparrow x_*} f(x) = f_*.$$

8.17 Beispiel: Für die Heaviside-Funktion [\rightarrow 8.15] gilt

$$\lim_{x \uparrow 0} H(x) = 0, \quad \lim_{x \downarrow 0} H(x) = 1.$$

8.18 Stetigkeit: Eine Funktion $f : D_f \rightarrow \mathbb{R}$ heißt *stetig an der Stelle* $x_* \in D_f$, wenn

$$\lim_{x \rightarrow x_*} f(x) = f(x_*).$$

f heißt *stetig*, wenn f an allen Stellen $x \in D_f$ stetig ist.

8.19 Regeln:

- Alle elementaren Funktionen (Polynome, exp, sin, cos, tan und deren Umkehrfunktionen) sowie Betrag-, Potenz- und Wurzelfunktionen sind stetig auf ihrem gesamten Definitionsgebiet.
- Summe, Differenz, Produkt, Quotient und Verkettung stetiger Funktionen sind stetig auf ihrem gesamten Definitionsgebiet.
- Die Umkehrfunktion einer stetigen Funktion ist stetig, sofern das Definitionsgebiet ein Intervall ist.
- *Zwischenwertsatz:* Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann ist B_f ein Intervall. Das heißt insbesondere, dass die Funktion alle Werte zwischen $f(a)$ und $f(b)$ annimmt.
- *Satz von Weierstraß:* Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann gibt es Stellen $\underline{x}, \bar{x} \in [a, b]$ mit

$$f(\underline{x}) \leq f(x) \leq f(\bar{x}), \quad x \in [a, b].$$

Eine derartige Aussage gilt in der Regel nicht, wenn das Definitionsgebiet kein abgeschlossenes Intervall ist.

8.20 Bemerkungen:

- f ist stetig an der Stelle x_* genau dann, wenn es für jedes $\varepsilon > 0$ es ein $\delta > 0$ gibt, sodass

$$|f(x) - f(x_*)| < \varepsilon \quad \text{falls} \quad |x - x_*| < \delta.$$

- Bei einer stetigen Funktion bewirken kleine Änderungen des Arguments kleine Änderungen des Funktionswerts.
- Das Schaubild einer stetigen Funktion besitzt keine Sprünge, sofern das Definitionsgebiet ein Intervall ist.

Nur die erste Aussage ist streng mathematischer Natur. Die beiden anderen sind unpräzise, aber gelegentlich hilfreich.

9 Differenziation

9.1 Differenzierbarkeit: Im Folgenden sei $f : I \rightarrow \mathbb{R}$ stets eine reelle Funktion, deren Definitionsgebiet $D_f = I$ ein offenes Intervall ist. Für zwei Stellen $x \neq x_0$ in I definieren wir den *Differenzenquotienten*

$$\frac{f(x) - f(x_0)}{x - x_0}.$$

Er gibt die Steigung der Sekante an, die durch die Punkte $(x, f(x))$ und $(x_0, f(x_0))$ gegeben ist. Die Funktion f heißt *differenzierbar an der Stelle* $x_0 \in I$, wenn der Grenzwert

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert.

- Der Grenzwert heißt *Ableitung* von f an der Stelle x_0 und man schreibt dafür

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Andere gebräuchliche Bezeichnungen sind

$$f'(x_0) = Df(x_0) = \frac{df}{dx}(x_0).$$

- Setzt man $h := x - x_0$, dann erhält man die äquivalente Darstellung

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

- Die Funktion f heißt *differenzierbar*, wenn sie an allen Stellen $x_0 \in I$ differenzierbar ist. Die Funktion $f' : I \rightarrow \mathbb{R}$, die jedem Punkt $x_0 \in I$ den Wert der Ableitung zuweist, heißt *Ableitungsfunktion*.

9.2 Beispiel:

- Für die lineare Funktion $f(x) = ax + b$ ist

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{(ax + b) - (ax_0 + b)}{x - x_0} = a.$$

Also ist f differenzierbar und es gilt $f'(x_0) = a$ für alle $x_0 \in \mathbb{R}$.

- Für die quadratische Funktion $f(x) = x^2$ ist

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = x + x_0.$$

Der Grenzwert für x gegen x_0 ist

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} x + x_0 = 2x_0.$$

Also ist f differenzierbar und es gilt $f'(x_0) = 2x_0$ für alle $x_0 \in \mathbb{R}$.

- Für die Betragsfunktion $f(x) = |x|$ und $x_0 = 0$ ist

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{|x|}{x} = \begin{cases} +1 & \text{für } x > 0 \\ -1 & \text{für } x < 0. \end{cases}$$

Es existiert also kein Grenzwert für $x \rightarrow 0$ und die Betragsfunktion ist deshalb nicht differenzierbar an der Stelle $x_0 = 0$.

9.3 Wichtige Ableitungen:

f	f'	
x^n	nx^{n-1}	$x \in \mathbb{R}, n \in \mathbb{N}$
x^n	nx^{n-1}	$x \neq 0, -n \in \mathbb{N}$
x^α	$\alpha x^{\alpha-1}$	$x > 0, \alpha \in \mathbb{R}$
e^x	e^x	$x \in \mathbb{R}$
$\ln x$	$1/x$	$x > 0$
$\sin x$	$\cos x$	$x \in \mathbb{R}$
$\cos x$	$-\sin x$	$x \in \mathbb{R}$
$\arctan x$	$1/(1+x^2)$	$x \in \mathbb{R}$
$\sinh x$	$\cosh x$	$x \in \mathbb{R}$
$\cosh x$	$\sinh x$	$x \in \mathbb{R}$
$\operatorname{artanh} x$	$1/(1-x^2)$	$x \in (-1, 1)$

9.4 Regeln: Seien f und g zwei differenzierbare Funktionen, dann gilt

- *Linearität:* Für $a, b \in \mathbb{R}$ ist die Funktion $af + bg$ differenzierbar,

$$(af + bg)' = af' + bg'.$$

- *Produktregel:* Die Funktion $f \cdot g$ ist differenzierbar,

$$(f \cdot g)' = fg' + f'g.$$

- *Quotientenregel:* Die Funktion f/g ist differenzierbar, sofern $g \neq 0$,

$$\left(\frac{f}{g}\right)' = \frac{f'g - g'f}{g^2}.$$

- *Kettenregel:* Die Funktion $f \circ g$ ist differenzierbar,

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

- *Umkehrfunktion:* Wenn $f' \neq 0$, dann ist f streng monoton [\rightarrow 9.8] und damit injektiv. Es existiert also die Umkehrfunktion [\rightarrow 8.10] $g := f^{-1}$, und es gilt

$$g'(y) = \frac{1}{f'(x)}, \quad \text{wobei } f(x) = y, \quad g(y) = x.$$

9.5 Beispiel:

- Linearität:

$$(3 \sin x - 2x^3)' = 3 \cos x - 6x^2$$

- Produktregel:

$$(x^2 \ln x)' = 2x \ln x + x, \quad x > 0$$

- Quotientenregel:

$$(\tan x)' = \left(\frac{\sin x}{\cos x} \right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x, \quad x \in (-\pi/2, \pi/2)$$

- Kettenregel:

$$\begin{aligned} (\exp(x^2))' &= \exp(x^2) \cdot (2x) \\ (\arctan(1/x))' &= \frac{-1/x^2}{1 + (1/x)^2} = \frac{-1}{1 + x^2}, \quad x \neq 0 \end{aligned}$$

- Umkehrfunktion:

$$\begin{aligned} y = f(x) = \tan x, \quad f'(x) = 1 + \tan^2 x > 0 \\ x = g(y) = \arctan y, \quad g'(y) = \frac{1}{1 + \tan^2 x} = \frac{1}{1 + y^2} \end{aligned}$$

9.6 Tangente: Sei f an der Stelle x_0 differenzierbar. Die Funktion

$$t(x) := f(x_0) + f'(x_0)(x - x_0)$$

bzw.

$$t(x_0 + h) = f(x_0) + f'(x_0)h$$

beschreibt eine Gerade, die man als *Tangente* von f an der Stelle x_0 bezeichnet. Sei $r(h) := f(x_0 + h) - t(x_0 + h)$ die Abweichung zwischen Funktion und Tangente, dann gilt

$$f(x_0 + h) = f(x_0) + hf'(x_0) + r(h), \quad \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

Ersetzt man die Ableitung $f'(x_0)$ durch eine beliebige andere reelle Zahl, so ist der Grenzwert des Quotienten $r(h)/h$ von Null verschieden. In diesem Sinne liefert die Ableitung $f'(x_0)$ die beste lineare Approximation der Funktion f in der Nähe des Punktes x_0 .

9.7 Mittelwertsatz: Sei f differenzierbar auf dem Intervall I , dann gibt eine Zahl $\vartheta \in (0, 1)$, sodass

$$f(x_0 + h) = f(x_0) + hf'(x_0 + \vartheta h),$$

sofern $x_0, x_0 + h \in I$. Setzt man $a = x_0, b = x_0 + h$ und $\xi = x_0 + \vartheta h$ für $h > 0$, dann erhält man

$$\frac{f(b) - f(a)}{b - a} = f'(\xi), \quad \xi \in (a, b).$$

Es gibt also an einer Zwischenstelle ξ eine Tangente, die parallel zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ ist.

9.8 Folgerungen: Sei $f : I \rightarrow \mathbb{R}$ eine differenzierbare Funktion.

- Wenn $f'(x) = 0$ für alle $x \in I$, dann gilt

$$f(x_0 + h) = f(x_0) + h \cdot 0 = f(x_0)$$

für alle h . Die Funktion f ist also konstant.

- Wenn $f'(x) > 0$ für alle $x \in I$, dann gilt

$$f(x_0 + h) = f(x_0) + hf'(x_0 + \vartheta h) > f(x_0)$$

für alle $h > 0$. Die Funktion f ist also streng monoton wachsend. Genauso ist f streng monoton fallend, wenn $f'(x) < 0$.

- Für die Abweichung r zwischen Funktion und Tangente [\rightarrow 9.6] gilt

$$r(h) = h(f'(x_0 + \vartheta h) - f'(x_0)).$$

9.9 Beispiel: Sei $f(x) = \sin x$. Mit $f'(x) = \cos x$ erhält man im Punkt $x_0 = 0$ die Tangente

$$t(h) = h \quad \text{und damit} \quad r(h) = \sin h - h = h(\cos(\vartheta h) - 1).$$

Für $|h| \leq 1/10$ gilt dann die Abschätzung

$$|\sin h - h| \leq |h| \cdot (1 - \cos h) \leq |h| \cdot (1 - \cos \frac{1}{10}) < \frac{|h|}{200}.$$

9.10 Regel von l'Hospital: Wenn f und g stetige Funktionen sind, dann gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{f(x_0)}{g(x_0)}, \quad \text{falls} \quad g(x_0) \neq 0.$$

Im Fall $f(x_0) \neq 0, g(x_0) = 0$ existiert kein Grenzwert. Wenn aber f und g an der Stelle x_0 differenzierbar sind, dann erhält man für $f(x_0) = g(x_0) = 0$ gemäß 9.6 mit $x = x_0 + h$

$$\begin{aligned} f(x) &= hf'(x_0) + r_1(h) \\ g(x) &= hg'(x_0) + r_2(h) \end{aligned}$$

und damit

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{h \rightarrow 0} \frac{f'(x_0) + r_1(h)/h}{g'(x_0) + r_2(h)/h} = \frac{f'(x_0)}{g'(x_0)}, \quad \text{falls} \quad g'(x_0) \neq 0.$$

Allgemeiner gilt: Wenn f und g differenzierbare Funktionen sind und

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$$

oder

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = \infty,$$

dann ist

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Diese Regeln gelten auch vollkommen analog für Grenzwerte der Form $x \rightarrow \infty$.

9.11 Beispiel:

•

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1$$

•

$$\lim_{x \rightarrow \infty} \frac{x^2}{e^x} = \lim_{x \rightarrow \infty} \frac{2x}{e^x} = \lim_{x \rightarrow \infty} \frac{2}{e^x} = 0$$

•

$$\lim_{x \downarrow 0} x \ln x = \lim_{x \downarrow 0} \frac{\ln x}{1/x} = \lim_{x \downarrow 0} \frac{1/x}{-1/x^2} = \lim_{x \downarrow 0} (-x) = 0$$

•

$$\lim_{x \rightarrow 1} \frac{\ln x - x + 1}{\cos(2 - 2x) - 1} = \lim_{x \rightarrow 1} \frac{1/x - 1}{2 \sin(2 - 2x)} = \lim_{x \rightarrow 1} \frac{-1/x^2}{-4 \cos(2 - 2x)} = \frac{1}{4}$$

9.12 Stetige Differenzierbarkeit: Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt *stetig differenzierbar*, wenn sie differenzierbar ist und die Ableitungsfunktion f' zudem stetig ist.

9.13 Beispiel:

- Die Funktion $f(x) = \sqrt{|x|^3}$, $x \in \mathbb{R}$, ist stetig differenzierbar, denn die Ableitungsfunktion

$$f'(x) = \begin{cases} \frac{3}{2}\sqrt{x} & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ -\frac{3}{2}\sqrt{x} & \text{für } x < 0 \end{cases}$$

ist stetig.

- Die Funktion

$$g(x) = \begin{cases} x^2 \sin(1/x) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

ist differenzierbar mit

$$g'(x) = \begin{cases} 2x \sin(1/x) - \cos(1/x) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Sie ist aber nicht stetig differenzierbar, da der Grenzwert $\lim_{x \rightarrow 0} g'(x)$ nicht existiert.

Das hier zu beobachtende Verhalten ist typisch. Ableitungsfunktionen differenzierbarer Funktionen weisen niemals Sprungstellen auf. Unstetigkeiten sind vielmehr stets durch ein oszillierendes Verhalten der Ableitungsfunktionen bedingt, das zu einer Nichtexistenz von Grenzwerten führt.

9.14 Höhere Ableitungen: Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt *k-mal differenzierbar*, wenn die Ableitungen in der Rekursion

$$\begin{aligned} f^{(0)} &:= f \\ f^{(n)} &:= (f^{(n-1)})' \end{aligned}$$

für alle $n = 1, \dots, k$ existieren. Die Funktion f heißt *unendlich oft differenzierbar*, wenn diese Ableitungen für alle $n \in \mathbb{N}$ existieren. Man nennt $f^{(k)}$ die *k-te Ableitung* von f und schreibt auch

$$f' = f^{(1)}, \quad f'' = f^{(2)}, \quad f''' = f^{(3)}$$

für Ableitungen niedriger Ordnung. Andere gebräuchliche Schreibweisen sind

$$f^{(k)} = D^k f = \frac{d^k f}{dx^k}.$$

Die Funktion f heißt *k-mal stetig differenzierbar*, wenn die Funktion $f^{(k)}$ stetig ist. Die Menge aller dieser Funktionen wird mit $C^k(I)$ bezeichnet. Insbesondere ist $C^0(I)$ die Menge der stetigen Funktionen auf dem Intervall I . Die Menge aller unendlich oft differenzierbaren Funktionen wird mit $C^\infty(I)$ bezeichnet.

9.15 Beispiel [\rightarrow 9.13]:

- Polynome und alle elementaren Funktionen (exp, sin, cos, ln) sind auf ihrem Definitionsgebiet unendlich oft differenzierbar. Es gilt z.B.

$$f(x) = x^n \quad \Rightarrow \quad f^{(k)}(x) = \begin{cases} \frac{n!}{(n-k)!} x^{n-k} & \text{für } 0 \leq k \leq n \\ 0 & \text{für } k > n \end{cases}$$

$$f(x) = e^{2x} \quad \Rightarrow \quad f^{(k)}(x) = 2^k e^{2x}$$

$$f(x) = \sin x \quad \Rightarrow \quad f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f'''(x) = -\cos x$$

- Für die *Stutzfunktion*

$$H_n(x) = \begin{cases} x^n & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

vom Grad n gilt $H_n \in C^{n-1}(\mathbb{R})$. Insbesondere erhält man für $n = 0$ die Heaviside-Funktion [\rightarrow 8.15]. Diese ist unstetig und man schreibt $H_0 \in C^{-1}(\mathbb{R})$.

- Die Funktion f gemäß Beispiel 9.13 ist stetig differenzierbar, aber nicht zweimal differenzierbar, $f \in C^1(\mathbb{R})$.
- Die Funktion g gemäß Beispiel 9.13 ist stetig, aber nicht stetig differenzierbar, $g \in C^0(\mathbb{R})$.

10 Komplexe Zahlen

10.1 Komplexe Multiplikation: Für zwei Vektoren

$$\vec{z}_1 = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \quad \vec{z}_2 = \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$

in \mathbb{R}^2 wird neben der üblichen Addition die *komplexe Multiplikation*

$$\vec{z}_1 * \vec{z}_2 := \begin{bmatrix} a_1 a_2 - b_1 b_2 \\ a_1 b_2 + b_1 a_2 \end{bmatrix}$$

definiert. Sie hat folgende Eigenschaften:

- *Kommutativgesetz:*

$$\vec{z}_1 * \vec{z}_2 = \vec{z}_2 * \vec{z}_1$$

- *Assoziativgesetz:*

$$\vec{z}_1 * (\vec{z}_2 * \vec{z}_3) = (\vec{z}_1 * \vec{z}_2) * \vec{z}_3$$

- *Distributivgesetz:*

$$\vec{z}_1 * (\vec{z}_2 + \vec{z}_3) = \vec{z}_1 * \vec{z}_2 + \vec{z}_1 * \vec{z}_3$$

Der Raum \mathbb{R}^2 versehen mit der komplexen Multiplikation wird *komplexe Zahlenebene* genannt und mit \mathbb{C} bezeichnet. Die Elemente von \mathbb{C} heißen *komplexe Zahlen*.

10.2 Beispiel:

$$\vec{z}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \vec{z}_2 = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad \Rightarrow \quad \vec{z}_1 * \vec{z}_2 = \vec{z}_2 * \vec{z}_1 = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

10.3 Einheiten: Die Einheitsvektoren \vec{e}_1 und \vec{e}_2 in der komplexen Zahlenebene werden mit

$$\vec{1} := \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{und} \quad \vec{i} := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

bezeichnet und *reelle Einheit* bzw. *imaginäre Einheit* genannt. Es gilt also

$$\vec{z} = \begin{bmatrix} a \\ b \end{bmatrix} = a\vec{1} + b\vec{i}.$$

Man bezeichnet a als *Realteil* und b als *Imaginärteil* der komplexen Zahl \vec{z} und schreibt dafür

$$a = \operatorname{Re} \vec{z}, \quad b = \operatorname{Im} \vec{z}.$$

Für eine beliebige komplexe Zahl $\vec{z} = [a, b]^T$ gilt

$$\vec{1} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \vec{i} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ a \end{bmatrix}.$$

$\vec{1}$ ist also das *neutrale Element* der komplexen Multiplikation, während Multiplikation mit \vec{i} den gegebenen Vektor \vec{z} um den Winkel $\pi/2$ gegen den Uhrzeigersinn dreht. Insbesondere gilt

$$\vec{i} * \vec{i} = -\vec{1}.$$

Damit erhält die Multiplikationsregel nach dem Distributivgesetz die Form

$$\begin{aligned} (a_1 \vec{1} + b_1 \vec{i}) * (a_2 \vec{1} + b_2 \vec{i}) &= a_1 a_2 (\vec{1} * \vec{1}) + a_1 b_2 (\vec{1} * \vec{i}) + b_1 a_2 (\vec{i} * \vec{1}) + b_1 b_2 (\vec{i} * \vec{i}) \\ &= (a_1 a_2 - b_1 b_2) \vec{1} + (a_1 b_2 + b_1 a_2) \vec{i}. \end{aligned}$$

10.4 Notation: Beim Rechnen mit komplexen Zahlen ist es üblich, die Vektorpfeile wegzulassen. Man schreibt also

$$z = a1 + bi \quad \text{für} \quad \vec{z} = a\vec{1} + b\vec{i}.$$

Weiterhin wird die Notation der reellen Einheit weggelassen. Man schreibt also

$$z = a + bi \quad \text{für} \quad z = a1 + bi.$$

Schließlich wird auch der Mal-Punkt nicht mit einem speziellen Symbol notiert. Man schreibt also

$$z_1 z_2 \text{ oder } z_1 \cdot z_2 \quad \text{für} \quad z_1 * z_2.$$

Die Rechenregeln lauten nun

$$\begin{aligned} (a_1 + b_1 i) + (a_2 + b_2 i) &= (a_1 + a_2) + (b_1 + b_2) i \\ (a_1 + b_1 i) \cdot (a_2 + b_2 i) &= (a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2) i. \end{aligned}$$

Das Rechnen mit komplexen Zahlen folgt also denselben Gesetzen wie das Rechnen mit reellen Zahlen. Es ist lediglich die Regel

$$i \cdot i = i^2 = -1$$

zu beachten.

10.5 Beispiel [\rightarrow 10.2]:

$$(2 + i) + (4 + 3i) = 6 + 4i, \quad (2 + i) \cdot (4 + 3i) = 5 + 10i$$

10.6 Polarkoordinaten: Der Punkt $z = a + bi$ kann entweder durch seine kartesischen Koordinaten (a, b) oder durch seine *Polarkoordinaten* (r, φ) definiert werden. Dabei ist r der Abstand vom Ursprung und φ der Winkel zur reellen Einheit 1,

$$\begin{bmatrix} a \\ b \end{bmatrix} = r \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} \quad \text{bzw.} \quad a + bi = r(\cos \varphi + i \sin \varphi).$$

Man nennt r den *Betrag* und φ das *Argument* von z und schreibt

$$r = |z|, \quad \varphi = \arg z.$$

Das Argument φ wird im *mathematisch positiven Sinn*, also gegen den Uhrzeigersinn, gemessen. Weiterhin ist zu beachten, dass φ nur bis auf Vielfache von 2π bestimmt ist. Typischerweise wählt man $\varphi \in [0, 2\pi)$.

10.7 Beispiel: Für $z = 1 - i\sqrt{3}$ ist

$$\begin{aligned} |z| = r &= \sqrt{1^2 + (-\sqrt{3})^2} = 2 \\ \cos \varphi = a/r &= 1/2, \quad \sin \varphi = b/r = -\sqrt{3}/2 \Rightarrow \arg z = \varphi = 5\pi/3. \end{aligned}$$

Es gilt also

$$1 - i\sqrt{3} = 2(\cos 5\pi/3 + i \sin 5\pi/3).$$

Man könnte anstelle von $\arg z = 5\pi/3$ auch $\arg z = -\pi/3$ oder $\arg z = 11\pi/3$ wählen.

10.8 Konjugation: Die Spiegelung einer komplexen Zahl an der reellen Achse bezeichnet man als *Konjugation* und schreibt dafür

$$\bar{z} = \overline{a + bi} = a - bi.$$

\bar{z} wird die zu z *konjugiert komplexe Zahl* genannt. Es gilt

•

$$\overline{\bar{z}} = z$$

•

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$$

•

$$z \cdot \bar{z} = (a + bi) \cdot (a - bi) = a^2 + b^2 = |z|^2$$

•

$$\operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}$$

10.9 Division: Man berechnet den Quotienten zweier komplexer Zahlen, indem man mit dem konjugiert Komplexen des Nenners erweitert,

$$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{|z_2|^2}, \quad z_2 \neq 0.$$

Damit ist der Nenner reell und die Division problemlos möglich. Division durch Null ist wie üblich ausgeschlossen.

10.10 Beispiel:

$$\frac{4 + i}{2 - 3i} = \frac{(4 + i)(2 + 3i)}{(2 - 3i)(2 + 3i)} = \frac{5 + 14i}{13} = \frac{5}{13} + \frac{14}{13}i$$

10.11 Geometrische Deutung der Multiplikation: Gegeben seien zwei komplexe Zahlen mit Polarkoordinaten (r_1, φ_1) und (r_2, φ_2) , also

$$z_1 = r_1(\cos \varphi_1 + i \sin \varphi_1)$$

$$z_2 = r_2(\cos \varphi_2 + i \sin \varphi_2)$$

Berechnet man das Produkt, so erhält man

$$z_1 z_2 = r_1 r_2 ((\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2))$$

Gemäß der Additionstheoreme für Winkelfunktionen lässt sich dies einfacher schreiben als

$$z_1 z_2 = r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)).$$

Die Multiplikation zweier komplexer Zahlen entspricht also einer *Multiplikation der Beträge* und einer *Addition der Argumente*. Genauso entspricht die Division zweier komplexer Zahlen einer *Division der Beträge* und einer *Subtraktion der Argumente*,

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} (\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)).$$

10.12 Komplexe e -Funktion: Man definiert für rein imaginäre Argumente die e -Funktion durch

$$e^{i\varphi} := \cos \varphi + i \sin \varphi, \quad \varphi \in \mathbb{R}.$$

Dieser Zusammenhang wird auch *Eulersche Formel* genannt. Es gilt $|e^{i\varphi}| = 1$. Das heißt, $e^{i\varphi}$ ist eine Zahl auf dem komplexen Einheitskreis, die durch den Winkel φ bestimmt ist. Damit hat eine komplexe Zahl mit Polarkoordinaten (r, φ) die Darstellung

$$z = r e^{i\varphi}$$

und die Multiplikation bekommt die einfache Form

$$z_1 \cdot z_2 = (r_1 e^{i\varphi_1}) \cdot (r_2 e^{i\varphi_2}) = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}.$$

Definiert man die e -Funktion für beliebige komplexe Argumente $z = a + ib$ durch

$$e^{a+ib} := e^a \cdot e^{ib} = e^a (\cos b + i \sin b),$$

dann gilt allgemein

$$e^{z_1} \cdot e^{z_2} = e^{z_1 + z_2}, \quad e^{z_1} / e^{z_2} = e^{z_1 - z_2}.$$

Die Bildmenge umfasst alle komplexen Zahlen mit Ausnahme der 0.

10.13 Beispiel:

$$e^0 = e^{2\pi i} = 1, \quad e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{-i\pi/2} = -i, \quad e^{2+i\pi} = -e^2$$

10.14 Komplexer Logarithmus: Die Umkehrfunktion der e -Funktion wird als *natürlicher Logarithmus* bezeichnet. Sei $w \neq 0$ gegeben, dann muss für $\ln w = z = a + bi$ gelten

$$e^z = e^a e^{ib} = w.$$

Hieraus folgt

$$|w| = e^a, \quad \arg w = b$$

und

$$\ln w = \ln |w| + i \arg w.$$

Es gelten die bekannten Regeln

$$\ln(w_1 \cdot w_2) = \ln w_1 + \ln w_2, \quad \ln(w_1/w_2) = \ln w_1 - \ln w_2.$$

Man beachte, dass der Logarithmus der Zahl 0 nicht definiert ist und dass sich die Mehrdeutigkeit des Arguments auf die Logarithmus-Funktion überträgt. Eindeutigkeit erhält man, indem man wieder $\arg w \in [0, 2\pi)$ fordert.

10.15 Beispiel:

$$\ln(-1) = i\pi, \quad \ln(-e) = 1 + i\pi, \quad \ln(i) = i\pi/2, \quad \ln(3 + 4i) = \ln 5 + i \arctan(4/3)$$

10.16 Komplexe Potenz-Funktion: Für komplexe Zahlen $x \neq 0$ und y definiert man die *Potenz* x^y durch

$$x^y := e^{y \ln x}.$$

Die komplexe *Wurzelfunktion* ist definiert durch

$$\sqrt[n]{z} := z^{1/n} = e^{(\ln z)/n}, \quad n \in \mathbb{N}.$$

10.17 Beispiel:

$$i^i = e^{i \ln i} = e^{-\pi/2}, \quad \sqrt{-1} = e^{(\ln(-1))/2} = e^{i\pi/2} = i, \quad \sqrt{-16} = 4i$$

10.18 Nullstellen von Polynomen: Die Nullstellen eines quadratischen Polynoms

$$p(z) = az^2 + bz + c, \quad a \neq 0,$$

sind durch die Formel

$$z_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

gegeben. Dabei ist die Wurzel im komplexen Sinn zu verstehen und deshalb stets definiert. Es gibt also immer zwei (unter Umständen zusammenfallende) Lösungen einer quadratischen Gleichung im Komplexen.

Allgemein gilt der *Fundamentalsatz der Algebra*: Das Polynom

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0, \quad a_n \neq 0$$

hat stets n komplexe Nullstellen.

10.19 Beispiel:

$$z^2 - 6z + 13 = 0 \quad \Rightarrow \quad z_{1,2} = \frac{6 \pm \sqrt{-16}}{2} = 3 \pm 2i$$

10.20 Beispiel: Zur Lösung der Gleichung

$$z^3 - 8i = 0$$

setzt man $z = r e^{i\varphi}$ und erhält

$$r^3 e^{3i\varphi} = 8e^{i\pi/2}.$$

Ein Vergleich der Beträge liefert $r = 2$, und für das Argument erhält man

$$3\varphi = \pi/2 + 2k\pi, \quad k \in \mathbb{Z}.$$

Dabei wurde verwendet, dass die e -Funktionen übereinstimmen, wenn sich die Argumente um ein ganzzahliges Vielfaches von 2π unterscheiden. Man erhält eine Folge von Argumenten

$$\varphi_k = \frac{\pi}{6} + \frac{2k\pi}{3}, \quad k \in \mathbb{Z},$$

aber nur drei davon führen auf verschiedene Lösungen,

$$\varphi_0 = \frac{\pi}{6}, \quad \varphi_1 = \frac{5\pi}{6}, \quad \varphi_2 = \frac{3\pi}{2}.$$

Alle anderen Werte unterscheiden sich von den gegebenen drei um ein ganzzahliges Vielfaches von 2π und liefern deshalb keine weiteren Lösungen. Man erhält schließlich

$$z_0 = 2e^{i\pi/6} = \sqrt{3} + i, \quad z_1 = 2e^{5i\pi/6} = -\sqrt{3} + i, \quad z_2 = 2e^{3i\pi/2} = -2i.$$