

# Mathematik für Informatiker

Prof. Dr. Martin Henk

<http://fma2.math.uni-magdeburg.de/~henk/lectures/math4inf>  
<http://shamash.math.uni-magdeburg.de/gilgamesh>

# Inhaltsverzeichnis

<b>I Mathematik I</b>	<b>4</b>
<b>0 Schulstoff</b>	<b>5</b>
<b>1 Aussagen und Mengen</b>	<b>7</b>
<b>2 Relationen und Abbildungen</b>	<b>12</b>
<b>3 Elementares Zählen und komplexe Zahlen</b>	<b>16</b>
<b>4 Lineare Gleichungssysteme und Matrizen</b>	<b>26</b>
<b>5 Vektorräume</b>	<b>35</b>
<b>6 Lineare Abbildungen und Matrizen</b>	<b>40</b>
<b>7 Normierte Vektorräume</b>	<b>43</b>
<b>8 Homogene Koordinaten, Quaternionen und Projektionen</b>	<b>46</b>
<b>9 Determinanten und Eigenwerte</b>	<b>49</b>
<b>II Mathematik II</b>	<b>55</b>
<b>10 Gruppen, Ringe, Körper</b>	<b>56</b>
<b>11 Folgen und Reihen</b>	<b>61</b>
<b>12 Stetigkeit von Funktionen</b>	<b>66</b>
<b>13 Differenzierbarkeit I</b>	<b>72</b>
<b>14 Taylor- und Potenzreihen</b>	<b>77</b>
<b>15 Integralrechnung I</b>	<b>81</b>
<b>16 Fourierreihen</b>	<b>87</b>
<b>17 Differenzierbarkeit II</b>	<b>90</b>
<b>18 Integralrechnung II</b>	<b>96</b>
<b>III Mathematik III</b>	<b>100</b>
<b>19 Elementare Wahrscheinlichkeitsrechnung</b>	<b>101</b>
<b>20 Elementare Statistik</b>	<b>110</b>
<b>21 Numerische Mathematik</b>	<b>117</b>
<b>22 Gewöhnliche Differentialgleichung</b>	<b>124</b>

<b>23 Kurven und Flächen</b>	<b>131</b>
<b>Index</b>	<b>139</b>

**Teil I**

**Mathematik I**

# 0 Schulstoff

**Bemerkung 0.1.** Für reelle Zahlen  $x, y, z$  gilt:

- i)  $x \leq y$  und  $y \leq z$ , dann  $x \leq z$ .
- ii)  $x \leq y$ , dann  $x \pm z \leq y \pm z$ .
- iii)  $x \leq y$  und  $z \geq 0$ , dann  $x \cdot z \leq y \cdot z$ .
- iv)  $x \leq y$ , dann  $-x \geq -y$ .
- v)  $0 < x \leq y$ , dann  $\frac{1}{x} \geq \frac{1}{y} > 0$ .

**Notation 0.2** (Betrag). Der Betrag für eine reelle Zahl  $x$  ist

$$|x| = \begin{cases} x, & \text{falls } x \geq 0, \\ -x, & \text{falls } x \leq 0. \end{cases}$$

**Bemerkung 0.3.** Für reelle Zahlen  $x, y$  gilt:

- i)  $|x| \geq 0$  mit Gleichheit nur für  $x = 0$ .
- ii)  $|x \cdot y| = |x| \cdot |y|$ .
- iii)  $|x + y| \leq |x| + |y|$ , bzw.  $|x + y| \geq |x| - |y|$  (Dreiecksungleichung).

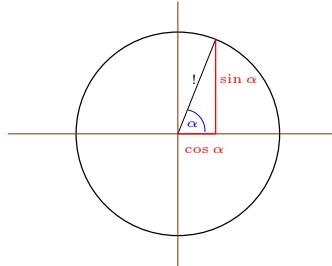


Abbildung 1: Sinus, Cosinus am Einheitskreis

**Bemerkung 0.4.**

- i)  $\sin(\alpha + 2\pi) = \sin \alpha$ ,  $\cos(\alpha + 2\pi) = \cos \alpha$ ,
- ii)  $\sin(-\alpha) = -\sin \alpha$ ,  $\cos(-\alpha) = \cos \alpha$ ,
- iii)  $\sin^2 \alpha + \cos^2 \alpha = 1$ , (Phthagoras)<sup>1</sup>
- iv)  $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$  und  $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$  (Additionstheoreme).

---

<sup>1</sup>Pythagoras von Samos, 582–507

**Notation 0.5.** Zur Abkürzung längerer Summen oder Produkte schreibt man

$$\sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_n,$$

$$\prod_{i=1}^n a_i = a_1 \cdot a_2 \cdot \dots \cdot a_n.$$

**Bemerkung 0.6.** Seien  $a_{ij}$  Zahlen für  $1 \leq i \leq n, 1 \leq j \leq m$ . Dann ist

$$\sum_{i=1}^n \sum_{j=1}^m a_{ij} = \sum_{j=1}^m \sum_{i=1}^n a_{ij}.$$

**Notation 0.7.** Sei  $a$  eine reelle Zahl, und sei  $n$  eine nicht-negative ganze Zahl.

- i)  $a^n = a \cdot a \cdot \dots \cdot a$  ( $n$ -mal),
- ii)  $a^0 = 1$ ,
- iii)  $a^{-n} = \frac{1}{a^n}, a \neq 0$ .

**Notation 0.8.** Sei  $a \geq 0$  und  $n$  eine positive ganze Zahl. Dann gibt es genau eine nicht-negative reelle Zahl  $x$  mit  $x^n = a$ ;  $x$  wird mit  $\sqrt[n]{a}$  bezeichnet. Für  $n = 2$  schreibt man auch nur  $\sqrt{a}$ .

Ist  $n$  ungerade und  $a < 0$  so ist auch  $\sqrt[n]{a} = -\sqrt[n]{-a}$  definiert.

**Notation 0.9.** Für  $a \geq 0$  und positive ganze Zahlen  $m, n$  definiert man  $a^{\frac{m}{n}} = \sqrt[n]{a^m} = (\sqrt[n]{a})^m$ ,

$$a^{-\frac{m}{n}} = \frac{1}{a^{\frac{m}{n}}}, a \neq 0.$$

**Bemerkung 0.10.** Für  $x, y \geq 0$  und reelle Zahlen  $\alpha, \beta$  gilt

- i)  $x^\alpha \cdot x^\beta = x^{\alpha+\beta}$ ,
- ii)  $\frac{x^\alpha}{x^\beta} = x^{\alpha-\beta}, x \neq 0$ ,
- iii)  $x^\alpha \cdot y^\alpha = (x \cdot y)^\alpha$ ,
- iv)  $\left(\frac{x}{y}\right)^\alpha = \frac{x^\alpha}{y^\alpha}, y \neq 0$ ,
- v)  $(x^\alpha)^\beta = x^{\alpha \cdot \beta}$ .

**Notation 0.11** (Logarithmus). Seien  $a, b > 0, b \neq 1$ . Die eindeutige Zahl  $x$  mit  $b^x = a$  heisst *Logarithmus von  $a$  zur Basis  $b$*  und wird  $\log_b a$  bezeichnet. Ist die Basis gegeben durch die Eulersche<sup>2</sup> Zahl  $e = 2,71828182\dots$ , dann heißt

$$\ln a = \log_e a$$

*natürlicher Logarithmus von  $a$ .*

**Bemerkung 0.12.**

- i)  $b^{\log_b a} = a, \log_b b = 1, \log_b 1 = 0$ ,
- ii)  $\log_b(a \cdot c) = \log_b a + \log_b c, \log_b\left(\frac{a}{c}\right) = \log_b a - \log_b c$ .
- iii)  $\log_c a = \frac{\log_b a}{\log_b c}$ ,
- iv)  $\log_b(a^n) = n \cdot \log_b a$ ,
- v)  $\log_b a \cdot \log_a b = 1$ ,
- vi)  $a^{\log_b c} = c^{\log_b a}$ .

---

<sup>2</sup>Leonhard Euler, 1707 – 1783

# 1 Aussagen und Mengen

**Vereinbarung 1.1** (Aussagen<sup>3</sup>). Eine *Aussage* ist ein Satz, der entweder *wahr* (*w*) oder *falsch* (*f*) ist. *w* und *f*, bzw. 1 und 0, heißen *Wahrheitswerte*.

**Definition 1.2** (Negation, Konjunktion, Disjunktion). Seien *A* und *B* Aussagen.

	Symbol	Sprechweise
<i>Negation von A</i>	$\overline{A}$	nicht <i>A</i>
<i>Konjunktion von A und B</i>	$A \wedge B$	<i>A und B</i>
<i>Disjunktion von A und B</i>	$A \vee B$	<i>A oder B</i> .

Die Wahrheitswerte dieser Verknüpfungen sind in den folgenden *Wahrheitstafeln* angegeben:

$A$	$\overline{A}$	$A$	$B$	$A \wedge B$	$A$	$B$	$A \vee B$
$w$	$f$	$w$	$w$	$w$	$w$	$w$	$w$
$f$	$w$	$f$	$f$	$f$	$w$	$f$	$w$
$f$	$w$	$f$	$w$	$f$	$f$	$w$	$w$
$f$	$f$	$f$	$f$	$f$	$f$	$f$	$f$

**Definition 1.3** (Implikation). Seien *A* und *B* Aussagen. Die Aussage „wenn *A* gilt, dann gilt *B*“ wird *Implikation* (oder *logische Folgerung*) genannt und mit  $A \Rightarrow B$  bezeichnet. Ihre Wahrheitswerte sind gegeben durch

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

Weitere Sprechweisen: „aus *A* folgt *B*“, „*A* impliziert *B*“, „*A* ist hinreichend für *B*“, oder „*B* ist notwendig für *A*“.

**Definition 1.4** (Äquivalenz). Seien *A* und *B* Aussagen. Die Aussageform

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

heißt *Äquivalenz* und wird mit  $A \Leftrightarrow B$  bezeichnet. Ihre Wahrheitswerte sind gegeben durch

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

Weitere Sprechweisen: „*A* ist gleichwertig mit *B*“, „*A* gilt genau dann, wenn *B* gilt“, „*A* gilt dann und nur dann, wenn *B* gilt“, „*A* ist äquivalent zu *B*“ oder „*A* ist notwendig und hinreichend für *B*“.

**Bemerkung 1.5.** Zwei (verknüpfte) Aussagen *A* und *B* sind also genau dann äquivalent, d.h. die Aussage  $A \Leftrightarrow B$  ist wahr, wenn sie für jede Wahl von (Eingangs-)Wahrheitswerten die gleichen Wahrheitswerte annehmen.

---

<sup>3</sup>Aristoteles, 384 – 322 v.Chr.

**Bemerkung 1.6.** Seien  $A, B$  Aussagen. Dann sind die folgenden Aussagen wahr:

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

**Bemerkung 1.7.** Für Aussagen  $A, B$  und  $C$  gilt:

- i)  $((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C),$
- ii)  $((A \Leftrightarrow B) \wedge (B \Leftrightarrow C)) \Rightarrow (A \Leftrightarrow C).$

**Bemerkung 1.8.** Ein mathematischer Satz besteht aus einer Voraussetzung  $A$  und einer Behauptung  $B$ . Der Beweis eines solchen Satzes besteht i.A. aus dem Nachweis, dass die Implikation  $A \Rightarrow B$  wahr ist. Man spricht dann auch von einem logischen Schluss. Die gebräuchlichsten Beweismethoden in der Mathematik sind:

- i) *Direkter Beweis einer Implikation  $A \Rightarrow B$ : Aus der Voraussetzung (Prämissen)  $A$  wird die Behauptung (Konklusion)  $B$  abgeleitet, indem gezeigt wird, dass die Implikation  $A \Rightarrow B$  wahr ist.*
- ii) *Indirekter Beweis oder Widerspruchsbeweis einer Implikation  $A \Rightarrow B$ : Anstatt der Implikation  $A \Rightarrow B$  kann man auch die Implikation  $\overline{B} \Rightarrow \overline{A}$  nachweisen, oder zeigen, dass die Konjunktion  $A \wedge \overline{B}$  falsch ist, d.h. zu einem Widerspruch führt (vgl. Bemerkung 1.6 iii)).*
- iii) *Beweis einer Äquivalenz  $A \Leftrightarrow B$ : Man beweist die Aussageformen  $A \Rightarrow B$  und  $B \Rightarrow A$ .*

**Vereinbarung 1.9** (Mengen<sup>4</sup>). Eine Menge ist eine Zusammenfassung von verschiedenen Objekten zu einer Gesamtheit. Die Objekte in einer Menge heißen Elemente der Menge. Weiterhin fordern wir, dass für jedes nur vorstellbare Objekt eindeutig entschieden werden kann, ob es ein Element der Menge ist oder nicht.

**Notation 1.10.** Die Aussage „ $x$  ist ein Element der Menge  $M$ “ wird mit  $x \in M$  bezeichnet, die Negation mit  $x \notin M$ , d.h.  $x$  ist kein Element der Menge  $M$ .

**Notation 1.11.** Mengen lassen sich auf unterschiedliche Arten beschreiben, z.B.

- i) Aufzählung der Elemente, z.B.,

$$M = \{1, 2, 3, 4, 5\}, \quad M = \{\circ, \diamond, \star\}.$$

- ii)  $\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$  ist die Menge der *natürlichen Zahlen*
- iii)  $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$  ist die Menge der *ganzen Zahlen*
- iv) Beschreibung von Eigenschaften, z.B.,

$$\begin{aligned} M &= \{\text{alle geraden positiven natürlichen Zahlen}\} \\ &= \{n \in \mathbb{N} : n \text{ ist Vielfaches von } 2\} \\ &= \{2n : n \in \mathbb{N}\}. \end{aligned}$$

- v)  $\mathbb{Q} = \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\}$  ist die Menge der *rationalen Zahlen*.
- vi)  $\mathbb{R}$  bezeichnet die Menge der *reellen Zahlen*.
- vii)  $\emptyset = \{\}$  bezeichnet die *leere Menge*.

---

<sup>4</sup>Georg Cantor, 1845–1918

**Bemerkung 1.12** (Russellsche Antinomie<sup>5</sup>). Sei  $M$  die "Menge" aller Mengen, die sich nicht selbst enthalten, d.h.,

$$M = \{X : X \text{ Menge und } X \notin X\}.$$

Ist  $M \in M$ ?

Falls "Ja", dann muss  $M$  die Bedingung  $M \notin M$  erfüllen!

Falls "Nein", dann gilt nach Definition von  $M$  aber  $M \in M$ !

Im Sinne unserer Vereinbarung ist also  $M$  keine Menge!

**Definition 1.13** (Allquantor und Existenzquantor). Sei  $M$  eine nichtleere Menge und für jedes  $x \in M$  sei  $A(x)$  eine Aussage.

- i) Die Aussage „für alle  $x \in M$  gilt  $A(x)$ “ wird mit „ $\forall x \in M : A(x)$ “ bezeichnet;  $\forall$  heißt *Allquantor*.
- ii) Die Aussage „es gibt ein  $x \in M$ , für das  $A(x)$  gilt“ wird mit „ $\exists x \in M : A(x)$ “ bezeichnet;  $\exists$  heißt *Existenzquantor*.

**Bemerkung 1.14.** Seien  $A(x)$ ,  $x \in M$ , Aussagen. Es gilt:

$$\forall x : \overline{A(x)} \Leftrightarrow \exists x : \overline{\overline{A(x)}} \quad \text{und} \quad \exists x : \overline{A(x)} \Leftrightarrow \forall x : \overline{\overline{A(x)}}.$$

**Erklärung 1.15** (Vollständige Induktion). Sei  $M = \{k, k+1, k+2, \dots\}$ , wobei  $k$  eine ganze Zahl ist, und seien  $A(n)$ ,  $n \in M$ , Aussagen. Zum Beweis der Aussage  $\forall n \in M : A(n)$  verwendet man häufig das Prinzip der *vollständigen Induktion*:

- i) *Induktionsanfang*: Man zeigt, dass die Aussage für  $n = k$  gilt, also man zeigt  $A(k)$ .
- ii) *Induktionsschritt*: Man zeigt die Implikation  $A(n) \Rightarrow A(n+1)$ .

In Worten: Wir zeigen zunächst, dass für die kleinste Zahl in  $M$  die Aussage gilt, d.h.,  $A(k)$ . Dann beweisen wir für jedes  $n$ , dass aus  $A(n)$  die Aussage  $A(n+1)$  folgt.

**Satz 1.16** (Geometrische Reihe). Sei  $x \in \mathbb{R}$ ,  $x \neq 1$ , und sei  $n \in \mathbb{N}$ . Dann gilt

$$\sum_{i=0}^n x^i = \frac{1 - x^{n+1}}{1 - x}. \quad (1.1)$$

*Beweis.* Für  $n \in \mathbb{N}$  sei  $A(n)$  die Aussage (1.1).

- i) (Induktionsanfang) Wir zeigen, dass  $A(1)$  wahr ist, denn

$$\sum_{i=0}^1 x^i = x^0 + x^1 = 1 + x = (1 + x) \frac{1 - x}{1 - x} = \frac{1 - x^2}{1 - x}.$$

- ii) (Induktionsschritt) Wir folgern  $A(n+1)$  aus  $A(n)$ :

$$\begin{aligned} \sum_{i=0}^{n+1} x^i &= \left( \sum_{i=0}^n x^i \right) + x^{n+1} \stackrel{A(n)}{=} \frac{1 - x^{n+1}}{1 - x} + x^{n+1} \\ &= \frac{1 - x^{n+1} + x^{n+1}(1 - x)}{1 - x} = \frac{1 - x^{n+2}}{1 - x}. \end{aligned}$$

□

---

<sup>5</sup>Bertrand Russell, 1872–1970

*Beispiel 1.17.* Für  $n \in \mathbb{N}$  sei  $A(n)$  die Aussage, dass in jeder Gruppe von  $n$  Studenten, in der einer Mathematik studiert, alle Mathematik studieren.

- Induktionsanfang:  $A(1)$  ist sicherlich wahr.
- Induktionsschritt: Sei  $M$  eine Gruppe von  $n+1$  Studenten, die wir mit  $s_1, \dots, s_{n+1}$  bezeichnen. Sei  $s_2$  ein Mathematikstudent.  $A(n)$  angewendet auf die Gruppe  $\{s_1, \dots, s_n\}$  und  $\{s_2, \dots, s_{n+1}\}$  zeigt, dass  $s_1, \dots, s_n$  und  $s_2, \dots, s_{n+1}$  Mathematik studieren; also alle und somit gilt  $A(n+1)$ .

Also gilt  $A(n)$  für jede natürliche Zahl  $n$ ?

**Definition 1.18** (Teilmenge).

- Eine Menge  $B$  heißt *Teilmenge* einer Menge  $A$  (Schreibweise:  $B \subseteq A$ ), wenn jedes Element von  $B$  auch Element von  $A$  ist, d.h.,

$$B \subseteq A \Leftrightarrow \forall b \in B : b \in A.$$

- Zwei Mengen  $A$  und  $B$  heißen *gleich* (Schreibweise:  $B = A$ ), falls  $B \subseteq A$  und  $A \subseteq B$ . Sind  $A$  und  $B$  nicht gleich, so schreibt man  $A \neq B$ .
- Eine Menge  $B$  heißt *echte Teilmenge* einer Menge  $A$  (Schreibweise:  $B \subset A$ ), wenn  $B \subseteq A$  und  $B \neq A$ .

**Bemerkung 1.19.** Die leere Menge ist Teilmenge einer jeden Menge.

**Definition 1.20** (Durchschnitt, Vereinigung, Differenz). Seien  $A, B$  Mengen.

i)

$$A \cap B = \{x : x \in A \wedge x \in B\}$$

heißt *Durchschnitt* von  $A$  und  $B$ . Ist  $A \cap B = \emptyset$ , so heißen  $A$  und  $B$  *disjunkt*.

ii)

$$A \cup B = \{x : x \in A \vee x \in B\}$$

heißt *Vereinigung* von  $A$  und  $B$ .

iii)

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

heißt *Differenz* von  $A$  und  $B$ .

- Ist  $B \subseteq U$ , so heißt die Differenz  $U \setminus B$  auch *Komplementärmenge von  $B$  in der Grundmenge  $U$*  und wird mit  $\overline{B}$  bezeichnet.

**Satz 1.21** (Verknüpfungsregeln für Mengen). Seien  $A, B$  Teilmengen einer Grundmenge  $U$ , also  $A, B \subseteq U$ .

i) Kommutativität:

$$A \cap B = B \cap A, \quad A \cup B = B \cup A.$$

ii) Assoziativität:

$$(A \cap B) \cap C = A \cap (B \cap C), \quad (A \cup B) \cup C = A \cup (B \cup C).$$

iii) Distributivität:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

iv) Idempotenz:

$$A \cap A = A, \quad A \cup A = A.$$

v) Absorption:

$$(A \cup B) \cap A = A, \quad (A \cap B) \cup A = A.$$

vi) Neutralität von  $\emptyset$  und  $U$ :

$$A \cap \emptyset = \emptyset, \quad A \cup \emptyset = A, \quad A \cup U = U, \quad A \cap U = A.$$

vii) Komplementregeln:

$$A \cap \overline{A} = \emptyset, \quad A \cup \overline{A} = U, \quad \overline{\overline{A}} = A.$$

viii) De Morgan'sche Regeln<sup>6</sup>:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

**Definition 1.22** (Boolesche Algebra<sup>7</sup>). Eine Menge, in der drei Operationen, z.B.  $(\cap, \cup, \neg)$  oder  $(\wedge, \vee, \neg)$ , definiert sind, die den Eigenschaften i)-vii) aus Satz 1.21 genügen, wird als *Boolesche Algebra* bezeichnet.

**Definition 1.23** (Kartesisches Produkt<sup>8</sup>).

i) Das *kartesische Produkt*  $X \times Y$  (Sprechweise:  $X$  kreuz  $Y$ ) zweier Mengen  $X, Y$  ist definiert als

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

$(x, y) \in X \times Y$  heißt *geordnetes Paar*.

ii) Allgemein ist das kartesische Produkt  $X_1 \times X_2 \times \dots \times X_n$  von  $n$ -Mengen  $X_1, \dots, X_n$  definiert als

$$X_1 \times X_2 \times \dots \times X_n = \{(x_1, x_2, \dots, x_n) : x_i \in X_i, 1 \leq i \leq n\}.$$

$(x_1, x_2, \dots, x_n) \in X_1 \times X_2 \times \dots \times X_n$  heißt *geordnetes  $n$ -Tupel*.

Für das kartesische Produkt  $X \times X \times \dots \times X$  ( $n$ -mal) schreibt man auch  $X^n$ .

**Bemerkung 1.24.** Für  $(x_1, \dots, x_n), (x'_1, \dots, x'_n) \in X_1 \times \dots \times X_n$  gilt

$$(x_1, \dots, x_n) = (x'_1, \dots, x'_n) \Leftrightarrow x_i = x'_i, 1 \leq i \leq n.$$

**Definition 1.25** (Mächtigkeit endlicher Mengen). Eine Menge  $A$  heißt *endlich*, wenn sie endlich viele Elemente enthält. Die Anzahl der Elemente von  $A$  heißt *Mächtigkeit* von  $A$  und wird mit  $|A|$  bezeichnet.

**Bemerkung 1.26.** Seien  $A, B$  endliche Mengen. Dann gilt

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

ii)  $|A \times B| = |A| \cdot |B|$

**Definition 1.27** (Potenzmenge). Sei  $A$  Menge. Die Menge  $P(A) = \{X : X \subseteq A\}$  aller Teilmengen von  $A$  heißt *Potenzmenge* von  $A$ .

**Lemma 1.28.** Sei  $A$  eine endliche Menge. Dann gilt

$$|P(A)| = 2^{|A|}.$$

---

<sup>6</sup>Augustus De Morgan, 1806–1871

<sup>7</sup>George Boole, 1815–1864

<sup>8</sup>René Descartes, 1596–1650

## 2 Relationen und Abbildungen

**Definition 2.1** (Relationen).

- i) Eine *binäre Relation*  $R$  zwischen den beiden Mengen  $A_1$  und  $A_2$  ist eine Teilmenge von  $A_1 \times A_2$ , also  $R \subseteq A_1 \times A_2$ . Für  $(x, y) \in R$  sagt man,  $x$  und  $y$  stehen in Relation  $R$  und schreibt dafür auch  $xRy$ . Im Falle  $A_1 = A_2 = A$  heißt  $R$  Relation auf  $A$ .
- ii) Eine *n-stellige* Relation auf den  $n$ -Mengen  $A_1, \dots, A_n$  ist eine Teilmenge von  $A_1 \times A_2 \times \dots \times A_n$ .

**Definition 2.2** (Verkettung/Komposition von Relationen). Sei  $R$  eine Relation zwischen den beiden Mengen  $A_1, A_2$ , und sei  $S$  eine Relation zwischen den Mengen  $A_2$  und  $A_3$ . Unter der *Verkettung (Komposition)*  $S \circ R$  von „ $S$  nach  $R$ “ versteht man die Relation zwischen  $A_1$  und  $A_3$  gegeben durch

$$S \circ R = \{(x, z) \in A_1 \times A_3 : \exists y \in A_2 \text{ mit } (x, y) \in R \text{ und } (y, z) \in S\}.$$

**Definition 2.3** (Inverse Relation). Sei  $R \subseteq A_1 \times A_2$  eine binäre Relation. Dann heißt

$$R^{-1} = \{(y, x) : (x, y) \in R\}$$

die zu  $R$  *inverse Relation*.  $R^{-1} \subseteq A_2 \times A_1$  ist eine Relation zwischen  $A_2$  und  $A_1$ .

**Bemerkung 2.4.** Seien  $A_1, A_2, A_3$  Mengen und  $R \subseteq A_1 \times A_2$ ,  $S \subseteq A_2 \times A_3$ . Dann gilt

$$R = (R^{-1})^{-1} \quad \text{und} \quad (S \circ R)^{-1} = R^{-1} \circ S^{-1}.$$

**Definition 2.5** (Äquivalenzrelation).

- i) Sei  $R$  eine Relation auf der Menge  $A$ .
  - a)  $R$  heißt *reflexiv*, falls für alle  $x \in A$  gilt:  $(x, x) \in R$ ,
  - b)  $R$  heißt *symmetrisch*, falls für alle  $(x, y) \in R$  gilt:  $(y, x) \in R$ ,
  - c)  $R$  heißt *transitiv*, falls für alle  $(x, y), (y, z) \in R$  gilt:  $(x, z) \in R$ .
- ii) Eine binäre Relation  $R$  auf  $A$ , die reflexiv, symmetrisch und transitiv ist, heißt *Äquivalenzrelation* auf  $A$ . In diesem Fall sagt man für  $(x, y) \in R$  auch „ $x$  ist äquivalent zu  $y$ “.

**Definition 2.6** (Äquivalenzklasse). Sei  $R$  eine Äquivalenzrelation auf  $A$ . Für  $x \in A$  heißt

$$[x]_R = \{y \in A : (x, y) \in R\}$$

die *Äquivalenzklasse von  $x$  bzgl.  $R$* .

**Satz 2.7.** Sei  $R$  eine Äquivalenzrelation auf  $A$ . Dann gilt:

- i)  $A = \bigcup_{x \in A} [x]_R$ ,
- ii) Für  $x, y \in A$  gilt entweder  $[x]_R = [y]_R$  oder  $[x]_R \cap [y]_R = \emptyset$ .

In Worten:  $A$  ist die Vereinigung seiner Äquivalenzklassen, und zwei verschiedene Äquivalenzklassen sind disjunkt.

**Beispiel 2.8** (Restklassen). Sei  $m \in \mathbb{N}$  und sei  $R_m$  die Relation auf  $\mathbb{Z}$  gegeben durch

$$R_m = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : x - y \text{ ist durch } m \text{ teilbar}\}.$$

- Dann ist  $R_m$  eine Äquivalenzrelation, und  $\mathbb{Z}$  ist die disjunkte Vereinigung der paarweise verschiedenen Äquivalenzklassen  $[0]_{R_m}, [1]_{R_m}, \dots, [m-1]_{R_m}$ .
- Für  $k \in \{0, \dots, m-1\}$  besteht  $[k]_{R_m}$  aus allen Zahlen, die bei Division durch  $m$  den Rest  $k$  lassen. Daher nennt man  $[k]_{R_m}$  auch *Restklasse*, und man bezeichnet sie einfach mit  $[k]_m$ .

**Definition 2.9** (Totale, partielle, Ordnung(srelation)). Sei  $R$  eine Relation auf der Menge  $A$ .

- $R$  heißt *antisymmetrisch*, falls für alle  $x, y \in A$  mit  $(x, y) \in R$  und  $(y, x) \in R$  gilt:  $x = y$ .
- Eine Relation  $R$  auf  $A$ , die reflexiv, antisymmetrisch und transitiv ist, heißt *Ordnung(srelation)* auf  $A$ .
- Ein Ordnung(srelation)  $R$  auf  $A$  mit der Eigenschaft  $(x, y) \in R$  oder  $(y, x) \in R$  für alle  $x, y \in A$  heißt *totale Ordnung*. Andernfalls heißt  $R$  eine *partielle Ordnung*.

*Beispiel 2.10.*

- $\leq$  ist eine totale Ordnung auf  $\mathbb{R}$ .
- $\subseteq$  ist eine partielle Ordnung z.B. auf der Potenzmenge von  $\{1, 2, 3\}$ .

**Definition 2.11** (Abbildung/Funktion). Eine *Abbildung (Funktion)*  $f$  von einer Menge  $X$  in eine Menge  $Y$  ist eine Vorschrift, die jedem  $x \in X$  genau ein Element  $f(x) \in Y$  zuordnet. Man schreibt dafür:

$$f : X \rightarrow Y \text{ mit } x \mapsto f(x).$$

Man sagt: „ $x$  wird auf  $f(x)$  abgebildet“, bzw. „ $f(x)$  ist das *Bild* oder der *Funktionswert* von  $x$ “.

Die Menge  $X$  heißt *Definitionsbereich* und die Menge  $Y$  heißt *Wertebereich* der Abbildung  $f$ .

**Definition 2.12** (Graph). Sei  $f : X \rightarrow Y$  mit  $x \mapsto f(x)$  eine Abbildung. Die Relation

$$\Gamma_f = \{(x, f(x)) : x \in X\} \subseteq X \times Y$$

heißt der *Graph* der Abbildung.

**Definition 2.13** (Bildmenge, Urbildmenge). Sei  $f : X \rightarrow Y$  mit  $x \mapsto f(x)$  eine Abbildung. Für  $A \subseteq X$  heißt

$$f(A) = \{f(x) : x \in A\}$$

*Bildmenge* von  $A$ , und für  $B \subseteq Y$  heißt

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

*Urbildmenge* von  $B$ . Für eelementige Mengen  $B$ , also  $B = \{y\}$ ,  $y \in Y$ , schreibt man auch einfach  $f^{-1}(y)$ .

*Beispiel 2.14.* Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $x \mapsto x^2$ , d.h.,  $f(x) = x^2$ .

- Definitions- und Wertebereich sind die reellen Zahlen.  $\Gamma_f = \{(x, x^2) : x \in \mathbb{R}\}$ .
- $f(\mathbb{R}) = \mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$ , also die nicht-negativen reellen Zahlen.
- $f(\{0, -1, 1\}) = \{0, 1\}$  und  $f^{-1}(\{2, 4\}) = \{\pm\sqrt{2}, \pm 2\}$ .
- $f^{-1}(\{-1\}) = f^{-1}(-1) = \emptyset$ .

**Definition 2.15** (Injektiv, surjektiv, bijektiv). Sei  $f : X \rightarrow Y$  eine Abbildung.

- $f$  heißt *injektiv*, wenn

$$\forall x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2),$$

d.h., verschiedene Elemente aus  $X$  werden auf verschiedene Elemente in  $Y$  abgebildet.

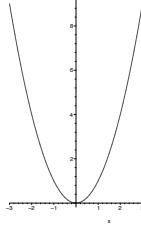


Abbildung 2: Der Graph  $\Gamma_f$

ii)  $f$  heißt *surjektiv*, wenn

$$f(X) = Y,$$

d.h., jedes Element aus  $Y$  ist das Bild eines Elementes aus  $X$ .

iii)  $f$  heißt *bijektiv*, oder *eins-zu-eins Abbildung* bzw. *eineindeutige Abbildung*, wenn  $f$  injektiv und surjektiv ist.

*Beispiel 2.16.*

- Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = x^2$ .  $f$  ist weder injektiv, da  $f(1) = f(-1)$ , noch surjektiv, da  $f(\mathbb{R}) = \mathbb{R}_{\geq 0} \neq \mathbb{R}$ .
- Sei  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  mit  $f(x) = x^2$ . Dann ist  $f$  surjektiv, aber immer noch nicht injektiv.
- Sei  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  mit  $f(x) = x^2$ . Dann ist  $f$  bijektiv.

**Definition 2.17** (Verkettung/Komposition von Abbildungen). Seien  $f : X \rightarrow Y$  und  $g : X' \rightarrow Y'$  Abbildungen mit  $f(X) \subseteq X'$ . Die *Verkettung (Komposition)* von „ $g$  nach  $f$ “ ist die Abbildung  $g \circ f : X \rightarrow Y'$  mit  $x \mapsto (g \circ f)(x) = g(f(x))$ . Die Verkettung von Abbildungen entspricht der Verkettung der Relationen  $\Gamma_{g \circ f} = \Gamma_g \circ \Gamma_f$ .

**Definition 2.18** (Inverse Abbildung). Sei  $f : X \rightarrow Y$  bijektiv. Die Abbildung  $g : Y \rightarrow X$ , die jedem  $y \in Y$  das eindeutig bestimmte  $x \in X$  mit  $y = f(x)$  zuordnet, heißt *inverse Abbildung*, und man schreibt für  $g$  auch  $f^{-1}$ . Die inverse Abbildung entspricht der inversen Relation:  $\Gamma_{f^{-1}} = \Gamma_f^{-1}$ .

**Lemma 2.19.** Seien  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  Abbildungen. Sind  $f$  und  $g$  bijektiv, dann ist auch  $g \circ f : X \rightarrow Z$  bijektiv, und es gilt:

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

**Definition 2.20** (Gleichmächtigkeit von Mengen). Zwei Mengen  $A, B$  heißen *gleichmächtig* oder *von gleicher Kardinalität* wenn es eine bijektive Abbildung von  $A$  nach  $B$  (und somit auch von  $B$  nach  $A$ ) gibt. Man schreibt dann auch  $|A| = |B|$ .

*Beispiel 2.21.*

- Zwei endliche Mengen  $A, B$  mit  $|A| = |B|$  sind stets gleichmächtig.
- $\mathbb{N}$  und  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$  sind gleichmächtig.
- $\mathbb{N}$  und  $\mathbb{Z}$  sind gleichmächtig.
- Man kann auch zeigen, dass  $\mathbb{R}$  und  $\{x \in \mathbb{R} : 0 < x < 1\}$  gleichmächtig sind.

**Definition 2.22** (Abzählbarkeit, Überabzählbarkeit). Sei  $M$  eine unendliche Menge.

- i)  $M$  heißt *abzählbar unendlich* oder kurz *abzählbar* wenn es eine bijektive Abbildung von  $\mathbb{N}$  nach  $M$  gibt, d.h.  $M$  ist von gleicher Mächtigkeit wie  $\mathbb{N}$ .
- ii) Ist  $M$  nicht *abzählbar unendlich*, dann heißt  $M$  *überabzählbar unendlich* oder kurz *überabzählbar*.

**Satz 2.23.**

- i)  $\mathbb{Q}$  ist abzählbar.
- ii)  $\mathbb{R}$  ist überabzählbar.
- iii) Die Vereinigung von abzählbar vielen abzählbaren Mengen ist abzählbar.

### 3 Elementares Zählen und komplexe Zahlen

**Definition 3.1** (Permutationen). Für  $n \in \mathbb{N}$  sei

$$S_n = \{\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\} : \sigma \text{ bijektiv}\}$$

die Menge aller bijektiven Abbildungen von  $\{1, \dots, n\}$  auf sich selbst.  $\sigma \in S_n$  heißt *Permutation*.

**Definition 3.2** (Fakultät). Für  $n \in \mathbb{N}$  heißt

$$n! = \prod_{i=1}^n i = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n$$

*Fakultät* (sprich:  $n$ -Fakultät). Man vereinbart zudem  $0! = 1$ .

**Lemma 3.3.** Für  $n \in \mathbb{N}$  ist  $|S_n| = n!$ .

**Bemerkung 3.4** (Stirling<sup>9</sup>-Formel).

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

d.h. für „große  $n$ “ entspricht  $n!$  etwa dem Ausdruck auf der rechten Seite.

**Bemerkung 3.5.** Die Fakultätsfunktion  $g : \mathbb{N}_0 \rightarrow \mathbb{N}$  mit  $g(n) = n!$  lässt sich auch rekursiv definieren:

$$g(n) = \begin{cases} 1, & n = 0, \\ n \cdot g(n-1), & n \geq 1. \end{cases}$$

**Satz 3.6** (Rekursionssatz<sup>10</sup>). Es sei  $z_0 \in \mathbb{Z}$ ,  $\mathbb{Z}_{\geq z_0} = \{z \in \mathbb{Z} : z \geq z_0\}$  und  $M$  eine beliebige, nicht leere Menge. Weiter sei  $f : \mathbb{Z}_{\geq z_0} \times M \rightarrow M$  eine Abbildung und  $m \in M$ . Dann liefert die folgende rekursive Vorschrift eine eindeutige Abbildung  $g : \mathbb{Z}_{\geq z_0} \rightarrow M$ .

$$g(n) = \begin{cases} m, & n = z_0, \\ f(n, g(n-1)), & n \geq z_0 + 1. \end{cases}$$

*Beispiel 3.7.*

- Sei  $z_0 = 0$  und  $f : \mathbb{Z}_{\geq 0} \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$  mit  $f(n, m) = n + m$ . Dann ist

$$\begin{aligned} g(n) &= \begin{cases} 0, & n = 0, \\ f(n, g(n-1)), & n \geq 1, \end{cases} \\ &= \sum_{i=0}^n i = \frac{(n+1)n}{2}. \end{aligned}$$

- (Verallgemeinerung) *Fibonacci*<sup>11</sup>-Zahlen: Sei  $F : \mathbb{N}_0 \rightarrow \mathbb{N}_0$  mit

$$F(n) = \begin{cases} 0, & n = 0, \\ 1, & n = 1, \\ F(n-1) + F(n-2), & n \geq 2. \end{cases}$$

---

<sup>9</sup>James Stirling, 1692–1770

<sup>10</sup>Richard Dedekind, 1831–1916

<sup>11</sup>Leonardo Pisano, 1170(80)-1250

**Bemerkung 3.8.**

- i ) Sei  $a = (1 + \sqrt{5})/2$  (goldener Schnitt) und  $b = (1 - \sqrt{5})/2$ . Für die  $n$ -te Fibonacci-Zahl  $F(n)$  gilt  $F(n) = (1/\sqrt{5})(a^n - b^n)$ .
- ii ) Think recursively!
- iii ) Türme von Hanoi, siehe [http://en.wikipedia.org/wiki/Tower\\_of\\_Hanoi](http://en.wikipedia.org/wiki/Tower_of_Hanoi)

**Definition 3.9** (Binomialkoeffizient). Für  $n \in \mathbb{N}_0$  und  $k \in \{0, \dots, n\}$  heißt

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

Zudem vereinbart man  $\binom{n}{k} = 0$  falls  $k < 0$  oder  $k > n$ .  $\binom{n}{k}$  liest man „ $n$  über  $k$ “.

**Satz 3.10.**

- i)  $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$  für  $n \geq 1$ .
- ii)  $\binom{n}{k} = \binom{n}{n-k}$ .

**Satz 3.11.** Die Anzahl der  $k$ -elementigen Teilmengen einer  $n$ -elementigen Menge beträgt  $\binom{n}{k}$ .

*Beispiel 3.12.*

- $\binom{4}{0} = 1$ ,  $\binom{4}{1} = 4$ ,  $\binom{4}{2} = 6$ ,  $\binom{4}{3} = 4$ ,  $\binom{4}{4} = 1$ .
- Sei  $M = \{1, 2, 3, 4\}$ .
 

0-elem. Teilmengen:	$\emptyset$
1-elem. Teilmengen:	$\{1\}, \{2\}, \{3\}, \{4\}$
2-elem. Teilmengen:	$\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}$
3-elem. Teilmengen:	$\{2, 3, 4\}, \{1, 3, 4\}, \{1, 2, 4\}, \{1, 2, 3\}$
4-elem. Teilmengen:	$\{1, 2, 3, 4\}$

**Satz 3.13** (Binomischer Lehrsatz). Für  $n \in \mathbb{N}_0$  und  $x, y \in \mathbb{R}$  gilt:

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

**Korollar 3.14.**

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

**Definition 3.15** (Primzahl). Eine natürliche Zahl  $p > 1$  heißt **Primzahl**, wenn sie nur durch sich selbst und 1 teilbar ist. Die Menge aller Primzahlen wird mit  $\mathbb{P}$  bezeichnet.

**Satz 3.16** (Primfaktorzerlegung). Jede natürliche Zahl größer als 1 lässt sich eindeutig (bis auf die Reihenfolge) als Produkt von Primzahlen schreiben.

**Satz 3.17** (Euklid<sup>12</sup>). Es gibt unendlich viele Primzahlen.

**Bemerkung 3.18.** Der Primzahlsatz<sup>13</sup> sagt aus, dass

$$|\{p \in \mathbb{P} : p \leq n\}| \approx \frac{n}{\ln n}.$$

---

<sup>12</sup>Euklid, ca. 300 v.Chr.

<sup>13</sup>Jacques Hadamard, 1865–1963; Charles de la Vallée Poussin, 1866–1962

**Notation 3.19.** Für Zahlen  $a, b \in \mathbb{Z}$  schreibt man  $a|b$ , falls  $a$  ein Teiler von  $b$  ist, d.h. falls es ein  $m \in \mathbb{Z}$  gibt mit  $b = m \cdot a$ .

**Definition 3.20** (ggT). Für  $a, b \in \mathbb{Z}$ ,  $a \neq 0$ , heißt die größte natürliche Zahl  $n \in \mathbb{N}$  mit  $n|a$  und  $n|b$  der *größte gemeinsame Teiler* von  $a$  und  $b$ . Er wird mit  $\text{ggT}(a, b)$  bezeichnet.

**Lemma 3.21.** Seien  $a, b \in \mathbb{Z}$  und  $m \in \mathbb{N}$ .

- i)  $\text{ggT}(a, b) = \text{ggT}(b, a) = \text{ggT}(-a, b) = \text{ggT}(a, -b) = \text{ggT}(-a, -b)$ .
- ii)  $\text{ggT}(a, b) = \text{ggT}(a \pm m \cdot b, b) = \text{ggT}(a, b \pm m \cdot a)$ .

**Lemma 3.22** (Division mit Rest). Sei  $a \in \mathbb{Z}$  und  $b \in \mathbb{N}$ . Dann gibt es eindeutige  $q \in \mathbb{Z}$  und  $r \in \{0, \dots, b-1\}$  mit

$$a = q \cdot b + r.$$

$r$  heißt der Rest von  $a$  bei Division durch  $b$  und wird mit  $a \bmod b$  bezeichnet (gesprochen „ $a$  modulo  $b$ “ oder einfach „ $a \bmod b$ “).

Insbesondere gilt  $\text{ggT}(a, b) = \text{ggT}(a \bmod b, b)$ .

**Satz 3.23** (Euklidischer Algorithmus). Seien  $a, b \in \mathbb{N}$  und sei  $a \geq b$ . Das folgende Verfahren, der sogenannte Euklidische Algorithmus, berechnet (rekursiv)  $\text{ggT}(a, b)$ .

- (1) Berechne  $r = a \bmod b$ .
- (2) Ist  $r = 0$ , dann  $\text{ggT}(a, b) = b$  und STOP.
- (3) Rufe das Verfahren auf für  $a = b$  und  $b = r$ , d.h. berechne  $\text{ggT}(b, a \bmod b)$ .

*Beispiel 3.24.* Berechnen von  $\text{ggT}(29393, 2805)$ .

Aufruf von $\text{ggT}(a, b)$	$r = a \bmod b$
$\text{ggT}(29393, 2805)$	1343
$\text{ggT}(2805, 1343)$	119
$\text{ggT}(1343, 119)$	34
$\text{ggT}(119, 34)$	17
$\text{ggT}(34, 17)$	0

$$29393 = 10 \cdot 2805 + 1343 \tag{3.1}$$

$$2805 = 2 \cdot 1343 + 119 \tag{3.2}$$

$$1343 = 11 \cdot 119 + 34 \tag{3.3}$$

$$119 = 3 \cdot 34 + 17 \tag{3.4}$$

$$34 = 2 \cdot 17 + 0$$

$$\text{ggT}(29393, 2805) = 17$$

**Lemma 3.25** (Lemma von Bézout<sup>14</sup>). Seien  $a, b \in \mathbb{Z}$ . Dann gibt es  $x, y \in \mathbb{Z}$ , so dass

$$\text{ggT}(a, b) = x \cdot a + y \cdot b.$$

---

<sup>14</sup>Étienne Bézout; 1730 – 1783

**Bemerkung 3.26.** Diese Zahlen  $x, y$  lassen sich leicht aus dem Euklidischen Algorithmus berechnen, wenn man in jedem Schritt bei der Berechnung von  $r = a \bmod b$  auch die Zahl  $m \in \mathbb{N}$  mit  $a = m \cdot b + r$  abspeichert.

Beispiel 3.27 ( $\text{ggT}(29393, 2805)$ ) (fortgesetzt)). Aus (3.4) folgt

$$17 = 119 - 3 \cdot 34.$$

Mit (3.3) erhält man

$$17 = 119 - 3 \cdot (1343 - 11 \cdot 119) = 34 \cdot 119 - 3 \cdot 1343.$$

Mit (3.2) erhält man

$$17 = 34 \cdot (2805 - 2 \cdot 1343) - 3 \cdot 1343 = 34 \cdot 2805 - 71 \cdot 1343.$$

Und abschließend mit (3.1)

$$17 = 34 * 2805 - 71 \cdot (29393 - 10 \cdot 2805) = -71 \cdot 29393 + 744 \cdot 2805.$$

**Definition 3.28** (Kongruente Zahlen). Seien  $a, b \in \mathbb{Z}$  und  $m \in \mathbb{N}$ .  $a$  und  $b$  heißen *kongruent modulo m*, falls sie bei Division durch  $m$  den gleichen Rest lassen, d.h. falls  $a \bmod m = b \bmod m$ . Man schreibt dafür  $a \equiv b \bmod m$ , gesprochen „ $a$  kongruent  $b$  mod(ulo)  $m$ “.

Mit der Notation aus Beispiel 2.8 gilt also:  $a \equiv b \bmod m$  genau dann, wenn  $a, b$  in der gleichen Restklasse  $R_m$  enthalten sind.

**Satz 3.29.** Sei  $a \equiv b \bmod m$  und  $c \equiv d \bmod m$ . Dann gilt

$$\begin{aligned} a + c &\equiv (b + d) \bmod m, \\ a \cdot c &\equiv (b \cdot d) \bmod m. \end{aligned}$$

**Satz 3.30** (Kleiner Satz von Fermat<sup>15</sup>). Sei  $p$  eine Primzahl und sei  $a \in \mathbb{N}$  mit  $\text{ggT}(a, p) = 1$ . Dann ist  $a^{p-1} \equiv 1 \bmod p$ .

Folgerung: Zum Testen, ob eine Zahl  $\bar{p} \in \mathbb{N}$  keine Primzahl ist, kann man wie folgt vorgehen: Man suche sich eine Zahl  $a$  mit  $\text{ggT}(a, \bar{p}) = 1$ . Ist dann  $a^{\bar{p}-1} \not\equiv 1 \bmod \bar{p}$ , dann ist nach dem Kleinen Satz von Fermat  $\bar{p}$  keine Primzahl. Ist zum Beispiel  $\bar{p}$  ungerade, so kann man  $a = 2$  wählen.

**Notation 3.31.** Sei  $m \in \mathbb{N}$ ,  $m \geq 2$ .

i) Sei

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

die Menge der Restklassen bei Division durch  $m$  (siehe Beispiel 2.8).

ii) Auf der Menge  $\mathbb{Z}_m$  wird nun eine Addition  $\oplus$  und Multiplikation  $\odot$  wie folgt definiert:

$$[a]_m \oplus [b]_m = [a + b]_m \text{ und } [a]_m \odot [b]_m = [a \cdot b]_m$$

**Bemerkung 3.32.** Aufgrund von Satz 3.29 sind  $\oplus$  und  $\odot$  wohldefiniert, d.h. für  $[a]_m = [c]_m$ ,  $[b]_m = [d]_m$  gilt

$$[a]_m \oplus [b]_m = [c]_m \oplus [d]_m \text{ und } [a]_m \odot [b]_m = [c]_m \odot [d]_m.$$

Also sind  $\oplus$  und  $\odot$  Abbildungen von  $\mathbb{Z}_m \times \mathbb{Z}_m$  nach  $\mathbb{Z}_m$ .

---

<sup>15</sup>Pierre de Fermat, ca. 1605 –1665

*Beispiel 3.33* (Verknüpfungstafeln für  $m = 4$ ).

$\oplus$	[0]_4	[1]_4	[2]_4	[3]_4
[0]_4	[0]_4	[1]_4	[2]_4	[3]_4
[1]_4	[1]_4	[2]_4	[3]_4	[0]_4
[2]_4	[2]_4	[3]_4	[0]_4	[1]_4
[3]_4	[3]_4	[0]_4	[1]_4	[2]_4

$\odot$	[0]_4	[1]_4	[2]_4	[3]_4
[0]_4	[0]_4	[0]_4	[0]_4	[0]_4
[1]_4	[0]_4	[1]_4	[2]_4	[3]_4
[2]_4	[0]_4	[2]_4	[0]_4	[2]_4
[3]_4	[0]_4	[3]_4	[2]_4	[1]_4

*Beispiel 3.34* (Verknüpfungstafeln für  $m = 5$ ).

$\oplus$	[0]_5	[1]_5	[2]_5	[3]_5	[4]_5
[0]_5	[0]_5	[1]_5	[2]_5	[3]_5	[4]_5
[1]_5	[1]_5	[2]_5	[3]_5	[4]_5	[0]_5
[2]_5	[2]_5	[3]_5	[4]_5	[0]_5	[1]_5
[3]_5	[3]_5	[4]_5	[0]_5	[1]_5	[2]_5
[4]_5	[4]_5	[0]_5	[1]_5	[2]_5	[3]_5

$\odot$	[0]_5	[1]_5	[2]_5	[3]_5	[4]_5
[0]_5	[0]_5	[0]_5	[0]_5	[0]_5	[0]_5
[1]_5	[0]_5	[1]_5	[2]_5	[3]_5	[4]_5
[2]_5	[0]_5	[2]_5	[4]_5	[1]_5	[3]_5
[3]_5	[0]_5	[3]_5	[1]_5	[4]_5	[2]_5
[4]_5	[0]_5	[4]_5	[3]_5	[2]_5	[1]_5

### Bemerkung 3.35.

i) Sowohl bei  $(\mathbb{Z}_4, \oplus)$  als auch bei  $(\mathbb{Z}_5, \oplus)$  gilt

$$[a]_m \oplus [0]_m = [a]_m,$$

und für jedes  $[a]_m \in \mathbb{Z}_m$  existiert ein  $[a']_m \in \mathbb{Z}_m$  mit

$$[a]_m \oplus [a']_m = [0]_m.$$

ii) Für  $(\mathbb{Z}_4, \odot)$  und  $(\mathbb{Z}_5, \odot)$  gilt analog

$$[a]_m \odot [1]_m = [a]_m,$$

aber für  $[2]_4$  gibt es kein Element in  $\mathbb{Z}_4$  mit  $[2]_4 \cdot [a']_4 = [1]_4$ . Für  $m = 5$  gibt es hingegen für jedes  $[a]_m \in \mathbb{Z}_m \setminus \{[0]_m\}$  ein  $[a']_m \in \mathbb{Z}_m$  mit

$$[a]_m \odot [a']_m = [1]_m.$$

**Definition 3.36** (Gruppe). Sei  $G$  eine nichtleere Menge, und sei  $\otimes : G \times G \rightarrow G$  mit  $(x, y) \mapsto x \otimes y$  eine Abbildung (Verknüpfung).

$(G, \otimes)$  heißt *Gruppe*, wenn die folgenden Bedingungen 1.–3. erfüllt sind:

1. Für alle  $x, y, z \in G$  gilt:  $(x \otimes y) \otimes z = x \otimes (y \otimes z)$  (Assoziativgesetz).
2. Es gibt ein *neutrales* Element  $e \in G$ , so dass für alle  $x \in G$  gilt:  $e \otimes x = x \otimes e = x$ .

3. Zu jedem  $x \in G$  gibt es ein *inverses Element*  $x' \in G$ , so dass  $x \otimes x' = x' \otimes x = e$ .  $x'$  wird auch mit  $x^{-1}$  bezeichnet.

Gilt zusätzlich  $x \otimes y = y \otimes x$  für alle  $x, y \in G$ , dann heißt  $(G, \otimes)$  *kommutative (abelsche<sup>16</sup>) Gruppe*.

*Beispiel 3.37.*

1.  $(\mathbb{Z}, +)$  ist kommutative Gruppe mit neutralem Element 0, und für  $a \in \mathbb{Z}$  ist  $-a$  das inverse Element. Ebenso sind  $(\mathbb{Q}, +), (\mathbb{R}, +)$  kommutative Gruppen.
2.  $(\mathbb{Q} \setminus \{0\}, \cdot)$  ist kommutative Gruppe mit neutralem Element 1, und für  $a \in \mathbb{Q} \setminus \{0\}$  ist  $1/a$  das inverse Element. Ebenso ist  $(\mathbb{R} \setminus \{0\}, \cdot)$  kommutative Gruppe.
3.  $(S_n, \circ)$  ist Gruppe mit neutralem Element  $\text{id}_{\{1, \dots, n\}}$ , und das inverse Element von  $\sigma \in S_n$  ist durch die Umkehrabbildung gegeben.  $S_n$  ist nicht kommutativ für  $n \geq 3$ .
4.  $(\mathbb{Z}_m, \oplus)$  ist kommutative Gruppe mit neutralem Element  $[0]_m$ , und für  $[a]_m \in \mathbb{Z}_m$  ist  $[-a]_m = [m-a]_m$  das inverse Element.
5.  $(\mathbb{Z}_4 \setminus \{0\}, \odot)$  ist keine Gruppe, aber  $(\mathbb{Z}_5 \setminus \{0\}, \odot)$  ist kommutative Gruppe.

**Definition 3.38** (Körper). Sei  $\mathbb{K}$  eine nichtleere Mengen, und seien  $+, \cdot : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$  mit  $(x, y) \mapsto x + y$  (Addition), bzw.  $(x, y) \mapsto x \cdot y$  (Multiplikation) Abbildungen.  $(\mathbb{K}, +, \cdot)$  heißt *Körper*, wenn die folgenden Bedingungen 1.–3. erfüllt sind:

1.  $(\mathbb{K}, +)$  ist kommutative Gruppe.  
Das neutrale Element von  $(\mathbb{K}, +)$  wird mit 0 bezeichnet.
2.  $(\mathbb{K} \setminus \{0\}, \cdot)$  ist kommutative Gruppe.  
Das neutrale Element von  $(\mathbb{K} \setminus \{0\}, \cdot)$  wird mit 1 bezeichnet und *Einselement* genannt.
3. Für alle  $x, y, z \in \mathbb{K}$  gilt:  $x \cdot y + x \cdot z = x \cdot (y + z)$  (Distributivgesetz).

*Beispiel 3.39.*

1.  $(\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$  sind Körper.
2.  $(\mathbb{Z}, +, \cdot)$  ist kein Körper.

**Satz 3.40.**  $(\mathbb{Z}_m, \oplus, \odot)$  ist genau dann ein Körper, wenn  $m$  Primzahl ist.

**Definition 3.41** (Komplexe Zahlen). Die Menge

$$\mathbb{C} = \{x + y \cdot i : x, y \in \mathbb{R}\}$$

heißt Menge der *komplexen Zahlen*. Dabei ist  $i \in \mathbb{C}$  definiert durch

$$i^2 = i \cdot i = -1$$

und heißt *imaginäre Einheit*.

Für  $z = x + y \cdot i \in \mathbb{C}$  heißt  $x$  *Realteil von z* und wird mit  $\text{Re}(z)$  bezeichnet.  $y$  heißt *Imaginärteil von z* und wird mit  $\text{Im}(z)$  bezeichnet.

Für  $x+0i$  bzw.  $0+yi$  schreibt man auch nur  $x$  bzw.  $y$  i. Zwei komplexe Zahlen  $z = x+yi$  und  $z' = x'+y'i$  sind gleich, d.h.  $z = z'$ , genau dann, wenn  $x = x'$  und  $y = y'$ .

---

<sup>16</sup>Niels Abel, 1802–1829

**Bemerkung 3.42.**

i ) Es ist  $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ .

ii) Betrachtet man die imaginäre Einheit als eine Variable mit der Eigenschaft  $i^2 = -1$ , dann ergeben sich Addition und Multiplikation von komplexen Zahlen unmittelbar aus den entsprechenden Operationen für reelle Zahlen.

Seien  $z = x + y i, z' = x' + y' i \in \mathbb{C}$ :

$$\begin{aligned} z + z' &= (x + y i) + (x' + y' i) \\ &= (x + x') + (y + y') i. \\ z \cdot z' &= (x + y i) \cdot (x' + y' i) = x \cdot x' + x \cdot y' i + y i \cdot x' + y i \cdot y' i \\ &= x x' + y y' i^2 + x y' i + y x' i \\ &= x x' - y y' + (x y' + y x') i. \end{aligned}$$

iii) Weiterhin ist für  $z = x + y i \neq 0$  ( $= 0 + 0 i$ )

$$z^{-1} = \frac{1}{x + y i} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} i.$$

**Satz 3.43.**  $(\mathbb{C}, +, \cdot)$  ist ein Körper mit neutralem Element 0 ( $= 0 + 0 i$ ) und Einselement 1 ( $= 1 + 0 i$ ).

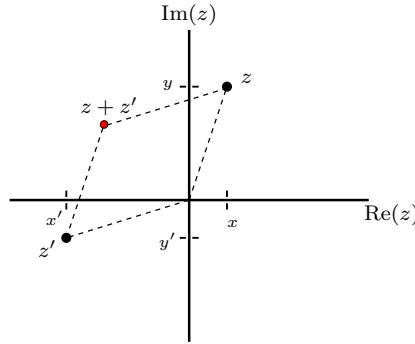


Abbildung 3: Addition von komplexen Zahlen in der Gaußschen Zahlenebene

**Notation 3.44** (Gaußsche<sup>17</sup> Zahlenebene).

**Definition 3.45** (Konjugierte Zahl, Betrag). Sei  $z = x + y i \in \mathbb{C}$ .

i)  $\bar{z} = x - y i \in \mathbb{C}$  heißt die zu  $z$  konjugierte (komplexe) Zahl.

ii)  $|z| = \sqrt{x^2 + y^2}$  heißt der Betrag der komplexen Zahl  $z$ .

**Satz 3.46.** Seien  $z, z' \in \mathbb{C}$ .

i)  $\overline{z + z'} = \bar{z} + \bar{z}'$  und  $\overline{z \cdot z'} = \bar{z} \cdot \bar{z}'$ .

ii)  $z = \bar{z}$  genau dann, wenn  $z \in \mathbb{R}$ .

iii)  $|z|^2 = z \cdot \bar{z}$ .

iv)  $z^{-1} = \bar{z}/|z|^2$ ,  $z \neq 0$ .

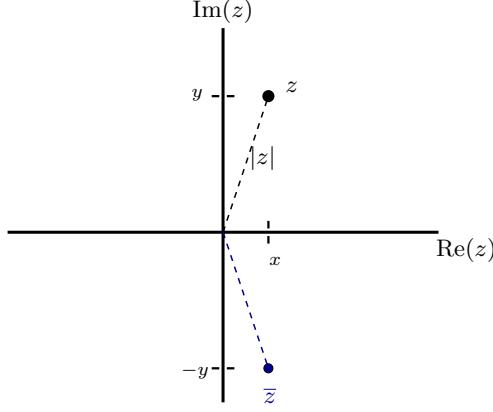


Abbildung 4: Konjugierte komplexe Zahl und Betrag von komplexen Zahlen in der Gaußsche Zahlenebene

**Definition 3.47** (Polarendarstellung). Jede komplexe Zahl  $z = x + y i \in \mathbb{C}$ ,  $z \neq 0$ , lässt sich eindeutig in der Form

$$z = r (\cos \phi + i \sin \phi) \quad (3.5)$$

darstellen. Dabei ist  $r = |z|$ , und  $\phi$  ist der Winkel zwischen den Vektoren  $(x, y)$  und  $(1, 0)$  entgegen dem Uhrzeigersinn (Drehwinkel), also

$$\cos \phi = \frac{x}{|z|} \text{ und } \sin \phi = \frac{y}{|z|}.$$

(3.5) heißt *Polarendarstellung* von  $z$ , und  $(r, \phi)$  nennt man *Polarkoordinaten* von  $z$ . Für die Polardarstellung benutzt man oft auch abkürzend die Eulersche Exponentialdarstellung

$$z = r (\cos \phi + i \sin \phi) = r e^{i\phi},$$

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

**Bemerkung 3.48.** Für  $z = r(\cos \phi + i \sin \phi)$ ,  $z' = r'(\cos \phi' + i \sin \phi') \in \mathbb{C}$  ist (siehe Bemerkung 0.4).

$$\begin{aligned} z \cdot z' &= r(\cos \phi + i \sin \phi) \cdot r'(\cos \phi' + i \sin \phi') \\ &= r r' (\cos(\phi + \phi') + i \sin(\phi + \phi')). \end{aligned}$$

bzw. mit  $z = r e^{i\phi}$ ,  $z' = r' e^{i\phi'}$  ist  $z \cdot z' = r e^{i\phi} r' e^{i\phi'} = r r' e^{i(\phi+\phi')}$  (s. Abb. 5).

**Lemma 3.49** (Formel von Moivre<sup>18</sup>). Für  $z = r (\cos \phi + i \sin \phi) = r e^{i\phi}$  und  $n \in \mathbb{N}$  gilt:

$$z^n = r^n (\cos(n \cdot \phi) + i \sin(n \cdot \phi)) = r^n e^{in\phi}.$$

**Definition 3.50** (Einheitswurzeln). Für  $n \in \mathbb{N}$  heißen

$$\omega_k^n = \cos \left( \frac{2k\pi}{n} \right) + i \sin \left( \frac{2k\pi}{n} \right) = e^{i \frac{2k\pi}{n}}, \quad k = 0, \dots, n-1,$$

$n$ -te Einheitswurzeln.

---

<sup>17</sup>Carl Friedrich Gauß (Gauss) 1777–1855

<sup>18</sup>Abraham de Moivre, 1667–1754

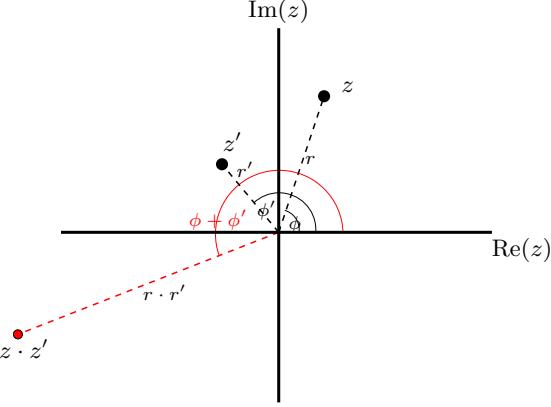


Abbildung 5: Multiplikation von komplexen Zahlen in Polardarstellung

Es ist  $(\omega_k^n)^n = 1$  für  $k = 0, \dots, n - 1$ .

*Beispiel 3.51.*

$$n = 1: \omega_0^1 = 1.$$

$$n = 2: \omega_0^2 = 1, \omega_1^2 = -1.$$

$$n = 3: \omega_0^3 = 1, \omega_1^3 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \omega_2^3 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i.$$

$$n = 4: \omega_0^4 = 1, \omega_1^4 = i, \omega_2^4 = -1, \omega_3^4 = -i.$$

**Bemerkung 3.52.** Die  $n$ -ten Einheitswurzeln  $\{\omega_k^n : k = 0, \dots, n - 1\}$  bilden mit der Multiplikation eine Gruppe.

**Bemerkung 3.53.** Sei  $z = r(\cos \phi + i \sin \phi) = r e^{i\phi} \in \mathbb{C}$  und  $n \in \mathbb{N}$ . Für

$$v_k = \sqrt[n]{r} \left( \cos \frac{\phi}{n} + i \sin \frac{\phi}{n} \right) \cdot \omega_k^n = \sqrt[n]{r} e^{i \frac{\phi+2k\pi}{n}}, \quad k = 0, \dots, n - 1,$$

gilt  $(v_k)^n = z$ ,  $k = 0, \dots, n - 1$ . Dies sind die  $n$ -ten Wurzeln von  $z$ .

**Notation 3.54.** Sei  $z = r(\cos \phi + i \sin \phi)$ . Als „(Quadrat)-Wurzel“ von  $z$ , im Zeichen  $\sqrt{z}$ , versteht man

$$\sqrt{z} = \sqrt{r} \left( \cos \frac{\phi}{2} + i \sin \frac{\phi}{2} \right) = \sqrt{r} e^{i \frac{\phi}{2}}.$$

*Beispiel 3.55.*

$$1. \sqrt{-1} = i, \sqrt{i} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i.$$

$$2. \text{ Für } x \in \mathbb{R}, x \leq 0, \text{ ist } \sqrt{x} = \sqrt{|x|}i.$$

**Satz 3.56** (Fundamentalsatz der Algebra). Jede Polynomfunktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  der Form

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0, \quad a_i \in \mathbb{C},$$

mit  $a_n \neq 0$  besitzt eine Darstellung

$$f(z) = a_n (z - c_1) \cdot (z - c_2) \cdot \dots \cdot (z - c_n),$$

mit  $c_i \in \mathbb{C}$ ,  $1 \leq i \leq n$ .  $c_1, \dots, c_n$  sind die Nullstellen des Polynoms  $f(z)$ .

*Beispiel 3.57.*

1. Sei  $f(z) = z^2 + a_1 z + a_0$ . Mit

$$c_1 = -\frac{a_1}{2} + \sqrt{\left(\frac{a_1}{2}\right)^2 - a_0}, \quad c_2 = -\frac{a_1}{2} - \sqrt{\left(\frac{a_1}{2}\right)^2 - a_0}$$

ist  $f(z) = (z - c_1)(z - c_2)$ .

2.  $z^2 + 1 = (z - i)(z + i)$ ,  $z^4 - 1 = (z - 1)(z - i)(z + 1)(z + i)$ .

3.  $z^n - 1 = \prod_{k=0}^{n-1} (z - \omega_k^n)$ .

**Satz 3.58.** Sei  $f : \mathbb{C} \rightarrow \mathbb{C}$  mit  $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ , wobei  $a_i \in \mathbb{R}$  und  $a_n \neq 0$ .

i) Ist  $z \in \mathbb{C}$  mit  $f(z) = 0$ , dann ist auch  $f(\bar{z}) = 0$ .

ii) Ist  $n$  ungerade, dann gibt es ein  $x^* \in \mathbb{R}$  mit  $f(x^*) = 0$ .

## 4 Lineare Gleichungssysteme und Matrizen

**Notation 4.1.** Im Folgenden sei  $(\mathbb{K}, +, \cdot)$  ein Körper (z.B.  $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_2, \mathbb{Z}_3$ , etc.). Die Elemente von  $\mathbb{K}$  heißen *Skalare*.

**Definition 4.2** (Lineares Gleichungssystem). Ein *lineares Gleichungssystem* (über dem Körper  $\mathbb{K}$ ) aus  $m$  Gleichungen mit  $n$  Unbekannten  $x_1, \dots, x_n$  hat die Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots &\quad \vdots \quad \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{4.1}$$

Dabei sind alle Koeffizienten  $a_{ij}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , und  $b_i$ ,  $1 \leq i \leq m$ , aus  $\mathbb{K}$ . Die Menge

$$L = \{(x_1, \dots, x_n) \in \mathbb{K}^n : x_1, \dots, x_n \text{ erfüllen (4.1)}\}$$

heißt *Lösungsmenge* des Gleichungssystems.

*Beispiel 4.3.*

1. Das lineare Gleichungssystem über  $\mathbb{Q}$

$$\begin{aligned} 2x_1 + 3x_2 &= 0 \\ 2x_1 + 3x_2 &= 1 \end{aligned}$$

besitzt keine Lösung.

2. Das lineare Gleichungssystem über  $\mathbb{R}$

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$$

besitzt genau eine Lösung  $(x_1, x_2) = (0, 0)$ .

3. Das lineare Gleichungssystem über  $\mathbb{Z}_2$

$$\begin{aligned} [1]_2 x_1 + [1]_2 x_2 &= [0]_2 \\ [1]_2 x_1 - [1]_2 x_2 &= [0]_2 \end{aligned}$$

hat die Lösungen  $([0]_2, [0]_2)$  und  $([1]_2, [1]_2)$ .

4. Das lineare Gleichungssystem über  $\mathbb{R}$

$$\begin{array}{rclclclclcl} x_1 &+& 2x_2 & & + & x_4 &=& 1 \\ x_1 &+& 2x_2 &+& 2x_3 &+& 3x_4 &=& 5 \\ 2x_1 &+& 4x_2 & & &+& 3x_4 &=& 5 \\ &&&&& 3x_3 &+& 2x_4 &=& 3 \end{array}$$

besitzt etwa die Lösung  $(x_1, x_2, x_3, x_4) = (-2, 0, -1, 3)$  oder  $(x_1, x_2, x_3, x_4) = (0, -1, -1, 3)$ .

**Lemma 4.4.** Die Lösungsmenge eines linearen Gleichungssystems über  $\mathbb{K}$  (4.1) ändert sich nicht, wenn man

- i) zwei Gleichungen vertauscht,

- ii) das  $\lambda$ -fache einer Gleichung zu einer anderen addiert,  $\lambda \in \mathbb{K}$ ,
- iii) eine Gleichung mit  $\lambda \in \mathbb{K} \setminus \{0\}$  multipliziert.

Der im folgenden beschriebene *Gauß'sche Algorithmus* (auch *Gauß'sches Eliminationsverfahren* genannt) benutzt die Regeln aus Lemma 4.4 sukzessive, um ein gegebenes lineares Gleichungssystem in ein anderes zu überführen, bei dem man die Lösungsmenge direkt ablesen kann.

*Beispiel 4.5.*

0. Gegeben sei folgendes Gleichungssystem über  $\mathbb{R}$

$$\begin{array}{rcl} x_1 + 2x_2 & + & x_4 = 1 \\ x_1 + 2x_2 + 2x_3 & + & 3x_4 = 5 \\ 2x_1 + 4x_2 & + & 3x_4 = 5 \\ 3x_3 + 2x_4 & = & 3 \end{array}$$

1. Addition des  $(-1)$ -fachen der 1.ten Gleichung zur 2.ten Gleichung, und des  $(-2)$ -fachen der 1.ten Gleichung zur 3.ten führt zu

$$\begin{array}{rcl} x_1 + 2x_2 & + & x_4 = 1 \\ 2x_3 + 2x_4 & = & 4 \\ x_4 & = & 3 \\ 3x_3 + 2x_4 & = & 3 \end{array}$$

2. Addition des  $(-3/2)$ -fachen der 2.ten Gleichung zur 4.ten Gleichung führt zu

$$\begin{array}{rcl} x_1 + 2x_2 & + & x_4 = 1 \\ 2x_3 + 2x_4 & = & 4 \\ x_4 & = & 3 \\ -x_4 & = & -3 \end{array}$$

3. Addition der 3.ten Gleichung zur 4.ten Gleichung führt zu

$$\begin{array}{rcl} x_1 + 2x_2 & + & x_4 = 1 \\ 2x_3 + 2x_4 & = & 4 \\ x_4 & = & 3 \\ 0 & = & 0 \end{array}$$

Hieraus lassen sich die Lösungen sofort ablesen:

4. Die 3.te Gleichung besagt  $x_4 = 3$ .

Diesen Wert von  $x_4$  eingesetzt in die 2.te Gleichung ergibt  $x_3 = -1$ .

Setzt man nun die Werte für  $x_3$  und  $x_4$  in die 1.te Gleichung ein, so erhält man  $x_1 + 2x_2 = -2$ .

Also ist

$$\begin{aligned} L &= \{(x_1, x_2, -1, 3) : x_1 + 2x_2 = -2\} \\ &= \{(-2 - 2x_2, x_2, -1, 3) : x_2 \in \mathbb{R}\}. \end{aligned}$$

**Definition 4.6** (Matrix).

i) Eine Anordnung von Skalaren  $a_{ij} \in \mathbb{K}$  in  $m$  Zeilen und  $n$  Spalten der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \text{ kurz } A = (a_{ij})_{i,j=1}^{m,n} \text{ oder } A = (a_{ij}),$$

wird  $m \times n$ -Matrix oder  $(m, n)$ -Matrix genannt.  $a_{ij}$  heißen die *Elemente* oder *Koeffizienten* der Matrix. Der erste Index,  $i$ , gibt die Zeilennummer (Zeilenindex), der zweite die Spaltennummer (den Spaltenindex) an, in der  $a_{ij}$  steht.

ii) Die Menge aller  $m \times n$ -Matrizen mit Koeffizienten aus  $\mathbb{K}$  wird mit  $\mathbb{K}^{m \times n}$  bezeichnet.

iii) Sind alle Koeffizienten von  $A$  gleich 0, dann heißt  $A$  Nullmatrix und wird mit  $\mathbf{0}$  bezeichnet.

**Definition 4.7** (Koeffizientenmatrix eines linearen Gleichungssystems). Für ein lineares Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots &\quad \vdots \quad \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{4.2}$$

über  $\mathbb{K}$  heißt

i)

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{m \times n}$$

die *Koeffizientenmatrix* des Gleichungssystems.

ii) Die Matrix

$$(A, b) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \in \mathbb{K}^{m \times (n+1)}$$

heißt *erweiterte Matrix* des Gleichungssystems.

*Beispiel 4.8.* Für das lineare Gleichungssystem aus Beispiel 4.5 ergeben sich die Koeffizientenmatrix und erweiterte Matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 2 & 2 & 3 \\ 2 & 4 & 0 & 3 \\ 0 & 0 & 3 & 2 \end{pmatrix} \in \mathbb{R}^{4 \times 4}, \quad (A, b) = \begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 3 & 5 \\ 2 & 4 & 0 & 3 & 5 \\ 0 & 0 & 3 & 2 & 3 \end{pmatrix} \in \mathbb{R}^{4 \times 5}.$$

**Definition 4.9** (Elementare Zeilenumformungen einer Matrix). Sei  $A \in \mathbb{K}^{m \times n}$ . Unter *elementaren Zeilenumformungen* versteht man die folgenden drei Operationen:

i)  $V_{k,l}$ : Das Vertauschen der  $k$ -ten mit der  $l$ -ten Zeile aus  $A$ .

- ii)  $A_{k,l}(\lambda)$ : Addition des  $\lambda$ -fachen der  $k$ -ten Zeile von  $A$  zu der  $l$ -ten Zeile von  $A$ ,  $\lambda \in \mathbb{K}$ ,
  - iii)  $M_k(\lambda)$ : Multiplikation der  $k$ -ten Zeile von  $A$  mit  $\lambda \in \mathbb{K} \setminus \{0\}$ .

*Beispiel 4.10.* Überträgt man die Umformungen aus Beispiel 4.5 auf die zugehörige erweiterte Matrix des Gleichungssystems (siehe Beispiel 4.8), dann ergeben sich die folgenden elementaren Zeilenumformungen:

$$\left( \begin{array}{ccccc} 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 3 & 5 \\ 2 & 4 & 0 & 3 & 5 \\ 0 & 0 & 3 & 2 & 3 \end{array} \right) \xrightarrow{A_{1,2}(-1), A_{1,3}(-2)} \left( \begin{array}{ccccc} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 2 & 3 \end{array} \right)$$
  

$$\xrightarrow{A_{2,4}(-3/2)} \left( \begin{array}{ccccc} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 & -3 \end{array} \right) \xrightarrow{A_{3,4}(1), M_2(1/2)} \left( \begin{array}{ccccc} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Die Gestalt der letzten Matrix nennt man auch (Zeilen)Stufenform.

**Definition 4.11** ((Zeilen)Stufenform). Eine Matrix  $A \in \mathbb{K}^{m \times n}$  hat *(Zeilen)Stufenform* wenn sie die folgende Darstellung hat:

Hierbei sind \* irgendwelche Elemente aus  $\mathbb{K}$ .

In „Worten“:  $A$  ist in (Zeilen)Stufenform wenn  $A = \mathbf{0}$ , oder es gibt ein  $r \in \{1, \dots, m\}$  und eine Folge von Zahlen  $1 \leq k_1 < k_2 < \dots < k_r \leq n$  mit folgenden Eigenschaften:

- i) Für  $i > r$ ,  $1 \leq j \leq n$ , ist  $a_{ij} = 0$ ,
  - ii) Für  $1 \leq i \leq r$  ist  $a_{ik_i} = 1$  und  $a_{ij} = 0$  für  $j < k_i$ .

$k_1, \dots, k_r$  sind die Spaltenindizes der Stufen.

### Bemerkung 4.12.

- i) Ist  $A \neq \mathbf{0}$  in Stufenform und ist  $i \leq r$ , dann steht das erste von Null verschiedene Element (eine 1) der  $i$ -ten Zeile der Matrix in der  $k_i$ -ten Spalte. Mit wachsendem  $i$  wandern diese ersten von Null verschiedenen Glieder wegen  $k_i < k_{i+1}$  nach rechts.

ii) Ist  $A \neq \mathbf{0}$  in Stufenform, dann hat  $A$  genau  $r$  Zeilen, in der nicht nur Nullen stehen.

**Satz 4.13.** Jede Matrix  $A \in \mathbb{K}^{m \times n}$  kann durch elementare Zeilenumformungen in Stufenform gebracht werden.

**Bemerkung 4.14** (Gauß-Algorithmus). Mittels elementarer Zeilenumformungen kann die erweiterte Matrix  $(A, b) \in \mathbb{K}^{m \times (n+1)}$  eines linearen Gleichungssystems in Stufenform  $(A', b')$  überführt werden. Sei  $A' =$

$(a'_{ij})_{i,j=1}^{m,n}$  und  $b' = (b'_1, \dots, b'_m)$ . Dann hat  $(A', b')$  die Gestalt

$$(A', b') = \begin{pmatrix} 0 & \dots & 0 & \boxed{1} & * & \dots & \dots & \dots & \dots & \dots & b'_1 \\ 0 & \dots & 0 & 0 & \dots & 0 & 1 & * & \dots & \dots & b'_2 \\ 0 & \dots & & & & & 1 & & \dots & \dots & \vdots \\ 0 & \dots & & & & & & \ddots & \dots & \dots & \vdots \\ 0 & \dots & & & & & & & \ddots & \dots & b'_r \\ 0 & \dots & & & & & & & & 0 & b'_{r+1} \\ 0 & \dots & & & & & & & & 0 & 0 \\ \vdots & & & & & & & & & \vdots & \vdots \\ 0 & \dots & & & & & & & & 0 & 0 \end{pmatrix}$$

Dabei ist  $b'_{r+1} \in \{0, 1\}$ .

Das zu  $(A', b')$  zugehörige Gleichungssystem hat die gleiche Lösungsmenge wie das gegebene Gleichungssystem (siehe Lemma 4.4).

i) Ist  $b'_{r+1} = 1$ , dann lautet die  $(r+1)$ -te Gleichung des Gleichungssystems  $(A', b')$

$$0x_1 + 0x_2 + \dots + 0x_n = 1.$$

Also hat das Gleichungssystem keine Lösung.

ii) Sei  $b'_{r+1} = 0$  und seien  $1 \leq k_1 < k_2 < \dots < k_r \leq n$  die Spaltenindizes der Stufen. Dann ist das Gleichungssystem  $(A', b')$  gegeben durch

$$\begin{aligned} x_{k_1} + \sum_{j=k_1+1}^n a'_{1j} x_j &= b'_1, \\ x_{k_2} + \sum_{j=k_2+1}^n a'_{2j} x_j &= b'_2, \\ &\vdots && \vdots \\ x_{k_r} + \sum_{j=k_r+1}^n a'_{rj} x_j &= b'_r. \end{aligned}$$

In diesem Fall sind alle  $x_j$ , außer  $x_{k_1}, \dots, x_{k_r}$  frei wählbar; aus der Wahl von  $x_j$ ,  $j \notin \{k_1, \dots, k_r\}$ , lassen sich dann die restlichen  $x_{k_1}, \dots, x_{k_r}$  „rückwärts“ berechnen:

$$\begin{aligned} x_{k_r} &= b'_r - \sum_{j=k_r+1}^n a'_{rj} x_j, \\ x_{k_{r-1}} &= b'_{r-1} - \sum_{j=k_{r-1}+1}^n a'_{r-1j} x_j, \\ &\vdots && \vdots \\ x_{k_1} &= b'_1 - \sum_{j=k_1+1}^n a'_{1j} x_j. \end{aligned}$$

**Satz 4.15.** Mit der obigen Notation gilt:

i) Ist  $b'_{r+1} = 1$  dann ist das Gleichungssystem nicht lösbar.

ii) Ist  $b'_{r+1} = 0$  dann ist das Gleichungssystem lösbar.

- a) Ist  $r = n$ , dann ist das Gleichungssystem eindeutig lösbar.
- b) Ist  $r < n$ , dann gibt es mehrere Lösungen.

**Definition 4.16** (Addition von Matrizen). Seien  $A, B \in \mathbb{K}^{m \times n}$  mit  $A = (a_{ij})$  und  $B = (b_{ij})$ . Dann ist die *Addition von A und B*, geschrieben  $A + B$ , definiert als

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

$A + B$  ist wieder eine  $m \times n$ -Matrix, d.h.  $A + B \in \mathbb{K}^{m \times n}$ .

Zwei Matrizen  $A, B \in \mathbb{K}^{m \times n}$ ,  $A = (a_{ij})$ ,  $B = (b_{ij})$  sind gleich, geschrieben  $A = B$ , wenn sie die gleichen Koeffizienten haben, d.h. wenn  $a_{ij} = b_{ij}$ , für  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ .

**Definition 4.17** (Multiplikation mit einem Skalar). Sei  $A \in \mathbb{K}^{m \times n}$  und  $\lambda \in \mathbb{K}$ . Die *Multiplikation von A mit  $\lambda$* , geschrieben  $\lambda \cdot A$  (bzw.  $\lambda A$ ), ist definiert als

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

$\lambda \cdot A$  ist wieder eine  $m \times n$ -Matrix, d.h.  $\lambda \cdot A \in \mathbb{K}^{m \times n}$ .

**Satz 4.18.** Seien  $A, B, C \in \mathbb{K}^{m \times n}$ ,  $\lambda, \mu \in \mathbb{K}$ . Dann gilt

- i)  $A + B = B + A$ , (Kommutativität),
- ii)  $A + (B + C) = (A + B) + C$ , (Assoziativität),
- iii)  $A + \mathbf{0} = A$  und  $A + (-A) = \mathbf{0}$ ,
- iv)  $\lambda(\mu A) = (\lambda\mu)A$ ,
- v)  $1A = A$ ,
- vi)  $\lambda(A + B) = \lambda A + \lambda B$  und  $(\lambda + \mu)A = \lambda A + \mu A$ .

Insbesondere ist  $(\mathbb{K}^{m \times n}, +)$  eine kommutative Gruppe mit neutralem Element  $\mathbf{0}$ , und für  $A \in \mathbb{K}^{m \times n}$  ist  $-A$  das inverse Element.

**Definition 4.19** (Transponierte einer Matrix). Sei  $A \in \mathbb{K}^{m \times n}$ . Die *transponierte Matrix* von  $A$ , bezeichnet mit  $A^\top$  (gesprochen  $A$  transponiert), ist definiert als die  $(n \times m)$ -Matrix, deren Spalten gleich den Zeilen von  $A$ , und deren Zeilen gleich den Spalten von  $A$  sind, d.h.

$$A^\top = (a_{ij}^\top)_{i,j=1}^{n,m} \text{ mit } a_{ij}^\top = a_{ji}, \quad 1 \leq i \leq n, 1 \leq j \leq m.$$

**Bemerkung 4.20.** Seien  $A, B \in \mathbb{K}^{m \times n}$  und  $\lambda \in \mathbb{K}$ . Dann gilt

- i)  $(A + B)^\top = A^\top + B^\top$ ,
- ii)  $(\lambda A)^\top = \lambda A^\top$ ,
- iii)  $(A^\top)^\top = A$ .

**Definition 4.21** (Quadratische Matrizen).

- i) Eine  $(n, n)$ -Matrix  $A = (a_{ij})$  heißt *quadratisch*. Die Elemente  $a_{ii}$ ,  $1 \leq i \leq n$ , heißen *Diagonalelemente* von  $A$ .
- ii) Eine quadratische Matrix, bei der alle Elemente außer den Diagonalelementen Null sind, heißt *Diagonalmatrix*.
- iii) Die Diagonalmatrix, bei der alle Diagonalelemente 1 sind, heißt *Einheitsmatrix*. Bezeichnung:  $I_n$ .
- iv) Eine quadratische Matrix  $A = (a_{ij})$ , bei der alle Elemente „unterhalb (oberhalb) der Diagonalen“ Null sind, d.h. für die

$$a_{ij} = 0 \text{ für } 1 \leq j < i \leq n \quad \left( a_{ij} = 0 \text{ für } 1 \leq i < j \leq n \right)$$

gilt, heißt obere (untere) *Dreiecksmatrix*.

- v) Eine quadratische  $(n, n)$ -Matrix  $A = (a_{ij})$  mit  $a_{ij} = a_{ji}$  für alle  $1 \leq i, j \leq n$  heißt *symmetrisch*.

**Bemerkung 4.22.** Ist  $A \in \mathbb{K}^{n \times n}$  symmetrisch, dann gilt also  $A^\top = A$ .

**Definition 4.23** (Produkt (Multiplikation) von Matrizen). Sei  $A \in \mathbb{K}^{m \times n}$  und sei  $B \in \mathbb{K}^{n \times r}$ . Das *Produkt von A und B*, geschrieben  $A \cdot B$  (bzw.  $A B$ ), ist die  $m \times r$ -Matrix  $C = (c_{ij})$ , deren Koeffizienten definiert sind als

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj},$$

für  $1 \leq i \leq m$ ,  $1 \leq j \leq r$ .

Man beachte: Das Produkt  $A \cdot B$  ist nur definiert, wenn die Anzahl der Spalten von  $A$  gleich der Anzahl der Zeilen von  $B$  ist!

„ $m \times n$  mal  $n \times r$  ergibt  $m \times r$ “.

**Satz 4.24.**

- i)  $A(BC) = (AB)C$ , für  $A \in \mathbb{K}^{m \times n}$ ,  $B \in \mathbb{K}^{n \times r}$ ,  $C \in \mathbb{K}^{r \times q}$  (Assoziativität),
- ii)  $(A+B)C = AC + BC$  für  $A, B \in \mathbb{K}^{m \times n}$ ,  $C \in \mathbb{K}^{n \times r}$ , bzw.  $A(B+C) = AB + AC$  für  $A \in \mathbb{K}^{m \times n}$ ,  $B, C \in \mathbb{K}^{n \times r}$  (Distributivität)
- iii)  $(\lambda A)B = \lambda(AB) = A(\lambda B)$ , für  $A \in \mathbb{K}^{m \times n}$ ,  $B \in \mathbb{K}^{n \times r}$ ,  $\lambda \in \mathbb{K}$ ,
- iv)  $(AB)^\top = B^\top A^\top$  für  $A \in \mathbb{K}^{m \times n}$ ,  $B \in \mathbb{K}^{n \times r}$ .

**Bemerkung 4.25.**

- i) Im Allgemeinen ist die Matrixmultiplikation nicht kommutativ:

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

- ii) Für  $A \in \mathbb{K}^{n \times n}$  und  $I_n \in \mathbb{K}^{n \times n}$  ist

$$A \cdot I_n = I_n \cdot A = A.$$

- iii) Ist  $A \in \mathbb{K}^{n \times n}$  eine quadratische Matrix, so ist  $A \cdot A$ ,  $A \cdot A \cdot A$  usw. immer definiert und man schreibt abkürzend

$$A^1 = A, \quad A^2 = AA, \quad \dots, \quad A^n = AA^{n-1}, \quad n \in \mathbb{N}.$$

Zudem setzt man  $A^0 = I_n$ .

iv) Sei  $A \in \mathbb{K}^{m \times n}$ ,  $A = (a_{ij})$ , und sei  $x$  die  $n \times 1$ -Matrix mit den Unbekannten  $x_1, \dots, x_n$ , d.h.

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

Dann ist

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

v) Ist nun  $b \in \mathbb{K}^{m \times 1}$  mit

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

dann lässt sich das lineare Gleichungssystem (4.1) schreiben als

$$Ax = b.$$

**Definition 4.26** ((In)homogenes Gleichungssystem). Sei  $A \in \mathbb{K}^{m \times n}$  und  $b \in \mathbb{K}^{m \times 1}$ . Ist  $b \neq \mathbf{0}$ , dann heißt  $Ax = b$  *inhomogenes Gleichungssystem* und  $Ax = 0$  heißt (das zugehörige) *homogene Gleichungssystem*.

**Satz 4.27.** Sei  $A \in \mathbb{K}^{m \times n}$  und  $b \in \mathbb{K}^{m \times 1}$ .

- i) Sind  $v, w$  Lösungen von  $Ax = b$ , dann ist  $v - w$  Lösung des zugehörigen homogenen Gleichungssystems  $Ax = \mathbf{0}$ .
- ii) Ist  $v$  Lösung von  $Ax = b$  und  $z$  Lösung von  $Ax = \mathbf{0}$ , dann ist  $v + z$  Lösung von  $Ax = b$ .
- iii)  $Ax = b$  ist eindeutig lösbar, genau dann wenn das zugehörige homogene Gleichungssystem  $Ax = 0$  nur die triviale Lösung, d.h.  $x_1 = x_2 = \dots = x_n = 0$ , hat.

**Definition 4.28** (Inverse Matrix). Sei  $A \in \mathbb{K}^{n \times n}$  eine quadratische Matrix. Gibt es eine Matrix  $B \in \mathbb{K}^{n \times n}$  mit

$$A \cdot B = B \cdot A = I_n,$$

dann heißt die Matrix  $A$  *invertierbar* oder *reguläre Matrix*. Die Matrix  $B$  wird mit  $A^{-1}$  bezeichnet und die zu  $A$  *inverse Matrix* genannt.

**Bemerkung 4.29.** Zum Bestimmen der inversen Matrix  $A^{-1}$  reicht es eine Matrix  $B$  zu bestimmen, so dass eine der beiden Forderungen  $A \cdot B = I_n$  oder  $B \cdot A = I_n$  erfüllt. Die jeweils andere ist automatisch mit erfüllt.

*Beispiel 4.30.*

Sei  $A = \begin{pmatrix} 2 & 4 \\ -1 & 3 \end{pmatrix}$ . Gesucht ist eine Matrix  $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$  mit  $AB = I_2$ , d.h.

$$AB = \begin{pmatrix} 2b_{11} + 4b_{21} & 2b_{12} + 4b_{22} \\ -b_{11} + 3b_{21} & -b_{12} + 3b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

bzw.

$$\begin{aligned} 2b_{11} + 4b_{21} &= 1, & 2b_{12} + 4b_{22} &= 0, \\ -b_{11} + 3b_{21} &= 0, & -b_{12} + 3b_{22} &= 1. \end{aligned}$$

Für jede Spalte der Matrix  $B$  erhält man also ein lineares Gleichungssystem. Lösen der Gleichungssysteme führt zu  $b_{11} = 3/10$ ,  $b_{21} = 1/10$ ,  $b_{12} = -2/5$  und  $b_{22} = 1/5$ . Somit ist

$$A^{-1} = \frac{1}{10} \begin{pmatrix} 3 & -4 \\ 1 & 2 \end{pmatrix}.$$

**Satz 4.31.** Sei  $A \in \mathbb{K}^{n \times n}$  und seien  $e_i \in \mathbb{K}^{n \times 1}$  mit

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

die Spalten der Einheitsmatrix  $I_n$ . Zum Bestimmen der inversen Matrix  $A^{-1}$  (falls sie existiert) müssen  $n$  lineare Gleichungssysteme in  $n$  Unbekannten gelöst werden. Die  $j$ -te Spalte von  $A^{-1}$  ist die Lösung des Systems

$$Ax = e_j.$$

**Satz 4.32.** Sei  $A \in \mathbb{K}^{n \times n}$  regulär. Das Gleichungssystem  $Ax = b$  hat die eindeutige Lösung  $A^{-1}b$ .

**Satz 4.33.** Seien  $A, B \in \mathbb{K}^{n \times n}$  regulär.

- i) Dann ist  $AB$  regulär und es gilt  $(AB)^{-1} = B^{-1}A^{-1}$ ,
- ii)  $(A^{-1})^\top = (A^\top)^{-1}$ ,
- iii)  $(A^{-1})^{-1} = A$ ,
- iv)  $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$ ,  $\lambda \in \mathbb{K} \setminus \{0\}$ .

## 5 Vektorräume

**Definition 5.1** (Vektorraum). Sei  $(\mathbb{K}, \oplus, \odot)$  ein Körper mit Einselement 1. Ein  $\mathbb{K}$ -Vektorraum ist eine nicht-leere Menge  $V$  mit zwei Verknüpfungen

$$\begin{aligned} + : V \times V &\rightarrow V, \quad (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w} \quad (\text{Vektoraddition}), \\ \cdot : \mathbb{K} \times V &\rightarrow V, \quad (\lambda, \mathbf{v}) \mapsto \lambda \cdot \mathbf{v} \quad (\text{Skalarmultiplikation}) \end{aligned}$$

so dass gilt:

- i)  $(V, +)$  ist kommutative Gruppe mit neutralem Element  $\mathbf{0}$ , und für  $\mathbf{a} \in V$  wird das inverse Element mit  $-\mathbf{a}$  bezeichnet,
- ii)  $1 \cdot \mathbf{v} = \mathbf{v}$  für alle  $\mathbf{v} \in V$ ,
- iii)  $(\lambda \odot \mu) \cdot \mathbf{v} = \lambda \cdot (\mu \cdot \mathbf{v})$  für alle  $\lambda, \mu \in \mathbb{K}$  und  $\mathbf{v} \in V$ ,
- iv)  $(\lambda \oplus \mu) \cdot \mathbf{v} = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v}$  für alle  $\lambda, \mu \in \mathbb{K}$  und  $\mathbf{v} \in V$ ,
- v)  $\lambda \cdot (\mathbf{v} + \mathbf{w}) = \lambda \cdot \mathbf{v} + \lambda \cdot \mathbf{w}$  für alle  $\lambda \in \mathbb{K}$  und  $\mathbf{v}, \mathbf{w} \in V$ .

Die Elemente von  $V$  heißen *Vektoren*.

**Bemerkung 5.2.** Aufgrund der Eigenschaften ii)-v), und da die entsprechenden Verknüpfungen immer aus dem Zusammenhang ersichtlich sind, schreibt man auch einfach  $\lambda \mathbf{v}$  statt  $\lambda \cdot \mathbf{v}$ ,  $(\lambda \mu) \mathbf{v}$  statt  $(\lambda \odot \mu) \cdot \mathbf{v}$  und  $(\lambda + \mu) \mathbf{v}$  statt  $(\lambda \oplus \mu) \cdot \mathbf{v}$ .

*Beispiel 5.3.*

- i)  $\mathbb{K}^{n \times m}$ , d.h. die Menge aller  $n \times m$ -Matrizen mit Koeffizienten aus  $\mathbb{K}$ , bilden einen  $\mathbb{K}$ -Vektorraum (siehe Satz 4.18).
- ii) Ist in i)  $m = 1$ , dann schreibt man  $\mathbb{K}^n$  anstatt  $\mathbb{K}^{n \times 1}$ . Die Koeffizienten von  $\mathbf{x} \in \mathbb{K}^n$  werden mit  $x_1, \dots, x_n$  bezeichnet, d.h.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T.$$

Die Vektoraddition und Skalarmultiplikation sind in diesem Fall

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \mathbf{x} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

- iii) Für  $n = 2$  kann man sich den Vektorraum  $\mathbb{R}^2$  als Ebene mit der bekannten Vektoraddition und Skalarmultiplikation vorstellen.
- iv) Analog kann man den  $\mathbb{R}^3$  mit dem drei-dimensionalen Anschauungsraum identifizieren.

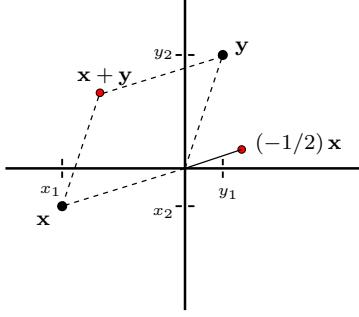


Abbildung 6: Vektoraddition und Skalarmultiplikation im  $\mathbb{R}^2$

- v) Sei  $M$  eine Menge und  $(\mathbb{K}, +, \cdot)$  ein Körper. Die Menge aller Abbildungen von  $M$  nach  $\mathbb{K}$ , bezeichnet mit  $\mathbb{K}^M$ , ist ein  $\mathbb{K}$ -Vektorraum mit den Verknüpfungen

$$+ : \mathbb{K}^M \times \mathbb{K}^M \rightarrow \mathbb{K}^M, (f, g) \mapsto f + g \text{ mit } (f + g)(x) = f(x) + g(x),$$

$$\cdot : \mathbb{K} \times \mathbb{K}^M \rightarrow \mathbb{K}^M, (\lambda, f) \mapsto \lambda \cdot f \text{ mit } (\lambda \cdot f)(x) = \lambda \cdot f(x).$$

- vi) Eine Funktion  $f : \mathbb{K} \rightarrow \mathbb{K}$  der Form  $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ ,  $a_i \in \mathbb{K}$ , heißt Polynom über  $\mathbb{K}$  vom Grad  $\leq n$ . Die Menge aller Polynome über  $\mathbb{K}$  vom Grad  $\leq n$  bilden einen  $\mathbb{K}$ -Vektorraum mit den obigen Verknüpfungen. Aber auch die Menge aller Polynome bildet einen  $\mathbb{K}$ -Vektorraum.

**Bemerkung 5.4.** Ist  $V$  ein  $\mathbb{K}$ -Vektorraum, dann gilt insbesondere

$$0\mathbf{a} = \mathbf{0}, \quad \lambda\mathbf{0} = \mathbf{0}, \quad (-\lambda)\mathbf{a} = -(\lambda\mathbf{a}),$$

für alle  $\lambda \in \mathbb{K}$ ,  $\mathbf{a} \in V$ . Hierbei ist  $0$  das neutrale Element von  $\mathbb{K}$ . Weiterhin ist

$$\lambda\mathbf{v} = \mathbf{0} \Leftrightarrow \lambda = 0 \text{ oder } \mathbf{v} = \mathbf{0}.$$

**Definition 5.5** (Teilraum). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Eine nicht-leere Teilmenge  $U \subseteq V$  heißt *Teilraum* oder *Untervektorraum* von  $V$ , falls  $U$  selbst ein  $\mathbb{K}$ -Vektorraum im Sinne der Definition 5.1 ist.

**Satz 5.6** (Unterraumkriterium). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $U \subseteq V$  eine nicht-leere Teilmenge.  $U$  ist genau dann ein Teilraum von  $V$ , wenn gilt

- i)  $\mathbf{0} \in U$ ,
- ii)  $\lambda\mathbf{v} + \mathbf{w} \in U$  für alle  $\mathbf{v}, \mathbf{w} \in U$  und  $\lambda \in \mathbb{K}$ .

*Beispiel 5.7.*

- i)  $\{\mathbf{0}\}$  und  $V$  sind „triviale“ Teilräume eines Vektorraums  $V$ .
- ii) Jede Gerade im  $\mathbb{R}^2$ , die den Nullpunkt enthält, ist ein Teilraum des  $\mathbb{R}^2$ . Geraden, die nicht den Nullpunkt enthalten, sind keine Teilräume.

**Satz 5.8.** Sei  $A \in \mathbb{K}^{m \times n}$ . Die Lösungsmenge des homogenen Gleichungssystems  $A\mathbf{x} = \mathbf{0}$ , d.h. die Menge  $U = \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}$  ist ein Teilraum von  $\mathbb{K}^n$ .

**Bemerkung 5.9.** Für  $\mathbf{a} = (a_1, \dots, a_n)^\top \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ ,  $\alpha \in \mathbb{R}$ , heißt  $H(\mathbf{a}, \alpha) = \{(x_1, \dots, x_n)^\top \in \mathbb{R}^n : a_1 x_1 + a_2 x_2 + \dots + a_n x_n = \alpha\}$  Hyperebene. Die  $\mathbf{0}$  enthaltenen Hyperebenen, d.h.  $H(\mathbf{a}, 0)$ , sind Teilräume, und für  $\alpha \neq 0$  ist  $H(\mathbf{a}, \alpha)$  kein Teilraum.

**Definition 5.10** (Summe von Teilräumen). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und seien  $U, U'$  Teilräume von  $V$ . Unter der *Summe von  $U$  und  $U'$* , geschrieben  $U + U'$ , versteht man die Menge

$$U + U' = \{\mathbf{u} + \mathbf{u}' : \mathbf{u} \in U, \mathbf{u}' \in U'\}.$$

**Lemma 5.11.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und seien  $U, U' \subseteq V$  Teilräume von  $V$ . Dann sind auch  $U \cap U'$  und  $U + U'$  Teilräume von  $V$ .

**Definition 5.12** (Linearkombination). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und seien  $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ . Ein Ausdruck der Form

$$\sum_{i=1}^m \lambda_i \mathbf{v}_i = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_m \mathbf{v}_m$$

heißt *Linearkombination* von  $\mathbf{v}_1, \dots, \mathbf{v}_m$ .

**Definition 5.13** (Lineare Hülle). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und sei  $M \subseteq V$  eine nicht-leere Teilmenge von  $V$ . Dann heißt

$$\text{lin } M = \left\{ \sum_{i=1}^m \lambda_i \mathbf{v}_i : m \in \mathbb{N}, \lambda_i \in \mathbb{K}, \mathbf{v}_i \in M, 1 \leq i \leq m \right\}$$

die *lineare Hülle von  $M$* . In Worten:  $\text{lin } M$  ist die Menge aller endlichen Linearkombinationen von Elementen aus  $M$ .

**Lemma 5.14.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und sei  $M \subseteq V$  eine nicht-leere Teilmenge von  $V$ . Dann ist  $\text{lin } M$  Teilraum von  $V$ .

**Bemerkung 5.15.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und seien  $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ . Dann ist

$$\text{lin } \{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \left\{ \sum_{i=1}^k \lambda_i \mathbf{v}_i : \lambda_i \in \mathbb{K}, 1 \leq i \leq k \right\}.$$

*Beispiel 5.16.*

- i)  $\text{lin } \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbb{R}^2$ .
- ii)  $\text{lin } \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \mathbb{R}^2$ .
- iii)  $\text{lin } \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 3 \end{pmatrix} = \{(x_1, x_2)^T \in \mathbb{R}^2 : x_1 = -x_2\}$ .

**Definition 5.17** (Erzeugendensystem). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $M \subseteq V$ ,  $M \neq \emptyset$ . Ist  $\text{lin } M = V$ , dann heißt  $M$  ein *Erzeugendensystem* von  $V$ .

Man vereinbart, dass die leere Menge ein Erzeugendensystem des trivialen Vektorraumes  $\{\mathbf{0}\}$  ist.

*Beispiel 5.18.*

- i)  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  ist ein Erzeugendensystem von  $\mathbb{R}^2$ .
- ii)  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \end{pmatrix}$  ist ein Erzeugendensystem von  $\mathbb{R}^2$ .
- iii)  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 3 \end{pmatrix}$  ist kein Erzeugendensystem von  $\mathbb{R}^2$ .

**Definition 5.19** (Linear (un)abhängig). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum, und seien  $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ . Die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_m$  heißen *linear unabhängig*, wenn aus

$$\sum_{i=1}^m \lambda_i \mathbf{v}_i = \mathbf{0}, \quad (5.1)$$

mit  $\lambda_i \in \mathbb{K}$ , stets folgt  $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$ .

Sind die Vektoren nicht linear unabhängig, dann heißen sie *linear abhängig*.

**Bemerkung 5.20.**

- i)  $\mathbf{v}_1, \dots, \mathbf{v}_m$  sind also genau dann linear unabhängig, wenn die einzige Möglichkeit den Nullvektor  $\mathbf{0}$  als Linearkombination der Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_m$  zu schreiben, die triviale ist, d.h. alle Skalare sind 0.
- ii)  $\mathbf{v}_1, \dots, \mathbf{v}_m$  sind also genau dann linear abhängig, wenn es Skalare  $\lambda_1, \dots, \lambda_m$  gibt, die nicht alle 0 sind, so dass (5.1) erfüllt ist.

*Beispiel 5.21.*

- i)  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{R}^2$  sind linear unabhängig.
- ii)  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \end{pmatrix}$  sind linear abhängig.
- iii)  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 3 \end{pmatrix}$  sind linear abhängig.
- iv)  $\mathbf{0}$  ist linear abhängig.
- v)  $1, i \in \mathbb{C}$  sind linear unabhängig falls  $\mathbb{C}$  als  $\mathbb{R}$ -Vektorraum betrachtet wird, aber linear abhängig falls  $\mathbb{C}$  als  $\mathbb{C}$ -Vektorraum betrachtet wird.

**Lemma 5.22.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum und seien  $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ . Dann sind die folgenden Aussagen äquivalent:

- i)  $\mathbf{v}_1, \dots, \mathbf{v}_m$  sind linear abhängig.
- ii) Es gibt ein  $\mathbf{v}_j \in \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ , das sich als Linearkombination der übrigen Vektoren darstellen lässt, d.h.

$$\mathbf{v}_j \in \text{lin}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_m\}.$$

**Definition 5.23.** Die Vektoren  $\mathbf{e}_i \in \mathbb{K}^n$  mit

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

d.h. die Spalten der Einheitsmatrix  $I_n$ , heißen *Einheitsvektoren*.

**Bemerkung 5.24.** Die Einheitsvektoren  $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{K}^n$  sind linear unabhängig und bilden ein Erzeugendensystem von  $\mathbb{K}^n$ , d.h.  $\text{lin}\{\mathbf{e}_1, \dots, \mathbf{e}_n\} = \mathbb{K}^n$ .

**Definition 5.25** (Basis). Ein Erzeugendensystem  $M$ , welches aus linear unabhängigen Vektoren eines  $\mathbb{K}$ -Vektorraums besteht, heißt *Basis*.

**Definition 5.26** (Dimension). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum mit Basis  $M \subseteq V$ . Dann heißt die Mächtigkeit von  $M$  die *Dimension von  $V$*  und wird mit  $\dim_{\mathbb{K}}(V)$  (oder einfach  $\dim(V)$ ) bezeichnet.

Ist  $\dim_{\mathbb{K}}(V)$  endlich, dann heißt  $V$  *endlichdimensionaler Vektorraum*; ansonsten *unendlichdimensionaler Vektorraum*.

*Beispiel 5.27.*

- i)  $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$ .
- ii)  $\dim_{\mathbb{K}}(\mathbb{K}^{m \times n}) = m \cdot n$ .
- iii)  $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ , aber betrachtet man  $\mathbb{C}$  als Vektorraum über  $\mathbb{R}$ , dann ist  $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ .

iv) Der Vektorraum aller Polynome über einem Körper ist nicht endlichdimensional.

Im folgenden betrachten wir nur endlichdimensionale Vektorräume!

**Satz 5.28.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum mit Basis  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ . Jeder Vektor  $\mathbf{v} \in V$  lässt sich eindeutig als Linearkombination der Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  darstellen, d.h. für jedes  $\mathbf{v} \in V$  gibt es eindeutig bestimmte  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  mit

$$\mathbf{v} = \sum_{i=1}^n \lambda_i \mathbf{v}_i.$$

**Satz 5.29.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum mit  $\dim_{\mathbb{K}}(V) = n$ . Dann ist jede Menge von  $n$  linear unabhängigen Vektoren eine Basis von  $V$ , und jede Basis besteht aus genau  $n$  Vektoren. Des Weiteren kann jede Menge von linear unabhängigen Vektoren zu einer Basis ergänzt werden.

**Bemerkung 5.30.** Da jeder Teilraum  $U$  eines  $\mathbb{K}$ -Vektorraumes  $V$  selbst ein Vektorraum ist, sind auch für Teilräume Begriffe wie Basis, Dimension etc. wohldefiniert.

Insbesondere ist  $\dim_{\mathbb{K}}\{\mathbf{0}\} = 0$ .

*Beispiel 5.31.*

- i) Die den Nullpunkt enthaltenen Geraden im  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  haben die Dimension 1.
- ii) Die Ebenen im  $\mathbb{R}^3$ , die den Nullpunkt enthalten, haben Dimension 2.

**Bemerkung 5.32.** Für  $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_n)^T \in \mathbb{R}^n \setminus \{\mathbf{0}\}$  ist die Dimension der Hyperebene  $H(\mathbf{a}, 0)$  gleich  $n - 1$ , also  $\dim_{\mathbb{R}}(H(\mathbf{a}, 0)) = n - 1$  (siehe Bemerkung 5.9).

**Satz 5.33.** Seien  $U, U'$  Teilräume eines  $\mathbb{K}$ -Vektorraumes  $V$ . Dann gilt

$$\dim_{\mathbb{K}}(U + U') = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(U') - \dim_{\mathbb{K}}(U \cap U').$$

## 6 Lineare Abbildungen und Matrizen

**Definition 6.1** (Lineare Abbildung). Seien  $V, W$  zwei  $\mathbb{K}$ -Vektorräume. Eine Abbildung  $f : V \rightarrow W$  heißt *lineare Abbildung*, wenn gilt:

- i)  $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$  für alle  $\mathbf{x}, \mathbf{y} \in V$ ,
- ii)  $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$  für alle  $\mathbf{x} \in V$  und  $\lambda \in \mathbb{K}$ .

**Bemerkung 6.2.** Insbesondere gilt für eine lineare Abbildung  $f(\mathbf{0}) = \mathbf{0}$ .

*Beispiel 6.3.*

- i)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f(\mathbf{x}) = -\mathbf{x}$  ist eine lineare Abbildung (Spiegelung am Ursprung).
- ii)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f((x_1, x_2)^\top) = (x_2, x_1)^\top$  ist eine lineare Abbildung (Spiegelung an der 45 Grad-Achse).
- iii)  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  mit  $f((x_1, x_2, x_3)^\top) = (x_1, x_2)^\top$  ist eine lineare Abbildung (Orthogonale Projektion auf die  $(x_1, x_2)$ -Ebene).
- iv)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit  $f((x_1, x_2)^\top) = (x_1, x_2, 0)^\top$  ist eine lineare Abbildung (Einbettung des  $\mathbb{R}^2$  im  $\mathbb{R}^3$ ).
- v) Sei  $\mathbf{t} \in \mathbb{R}^2$ ,  $\mathbf{t} \neq \mathbf{0}$ . Die Abbildung  $T_{\mathbf{t}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $T_{\mathbf{t}}(\mathbf{x}) = \mathbf{x} + \mathbf{t}$  ist keine lineare Abbildung (Translation um  $\mathbf{t}$ ).

**Bemerkung 6.4.** Seien  $V, W$  zwei  $\mathbb{K}$ -Vektorräume, und sei  $f : V \rightarrow W$  eine lineare Abbildung. Seien  $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$  und  $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ ,  $m \in \mathbb{N}$ . Dann gilt

$$f \left( \sum_{i=1}^m \lambda_i \mathbf{v}_i \right) = \sum_{i=1}^m \lambda_i f(\mathbf{v}_i).$$

**Satz 6.5.** Eine Abbildung  $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$  ist genau dann linear, wenn es eine Matrix  $A \in \mathbb{K}^{m \times n}$  gibt mit

$$f(\mathbf{x}) = A \mathbf{x} \quad \text{für alle } \mathbf{x} \in \mathbb{K}^n.$$

Die Spalten  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{K}^m$  der Matrix  $A$  sind die Bilder der Einheitsvektoren  $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{K}^n$ , d.h.

$$\mathbf{a}_i = f(\mathbf{e}_i), \quad 1 \leq i \leq n.$$

*Beispiel 6.6.*

- i)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f(\mathbf{x}) = -\mathbf{x}$ :

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

- ii)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f((x_1, x_2)^\top) = (x_2, x_1)^\top$ :

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

- iii)  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  mit  $f((x_1, x_2, x_3)^\top) = (x_1, x_2)^\top$ :

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

iv)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit  $f((x_1, x_2)^\top) = (x_1, x_2, 0)^\top$ :

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

v)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^4$  mit  $f((x_1, x_2)^\top) = (2x_1 - 3x_2, x_2, x_1 - x_2, -5x_1)^\top$ :

$$A = \begin{pmatrix} 2 & -3 \\ 0 & 1 \\ 1 & -1 \\ -5 & 0 \end{pmatrix}.$$

**Satz 6.7.** Seien  $V, W$  zwei  $\mathbb{K}$ -Vektorräume, und sei  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  eine Basis von  $V$ . Seien  $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$ . Dann gibt es eine eindeutig bestimmte lineare Abbildung  $f : V \rightarrow W$  mit

$$f(\mathbf{v}_i) = \mathbf{w}_i, \quad 1 \leq i \leq n.$$

Also: Lineare Abbildungen sind eindeutig durch die Bilder einer Basis festgelegt.

**Definition 6.8** (Affine Abbildung). Sei  $A \in \mathbb{K}^{m \times n}$  und  $\mathbf{t} \in \mathbb{K}^m$ . Die Abbildung  $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$  mit  $f(\mathbf{x}) = A\mathbf{x} + \mathbf{t}$  für alle  $\mathbf{x} \in \mathbb{K}^n$  heißt *affine Abbildung*.

**Definition 6.9** (Drehung im  $\mathbb{R}^2$ ). Unter der *Drehung (Rotation)* im  $\mathbb{R}^2$  um den Ursprung  $(\mathbf{0})$  mit Winkel  $\theta$  versteht man die lineare Abbildung  $\text{rot}_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$\text{rot}_\theta \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

**Definition 6.10** (Translationen). Sei  $\mathbf{t} \in \mathbb{R}^n$ . Die Abbildung  $T_{\mathbf{t}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit  $T_{\mathbf{t}}(\mathbf{x}) = \mathbf{x} + \mathbf{t}$  heißt *Translation (Verschiebung) um  $\mathbf{t}$* .

**Bemerkung 6.11.** Die Drehung im  $\mathbb{R}^2$  um den Punkt  $\mathbf{v}$  mit Winkel  $\theta$  ist gegeben durch die Abbildung  $f = T_{\mathbf{v}} \circ \text{rot}_\theta \circ T_{-\mathbf{v}}$ , d.h.

$$f(\mathbf{x}) = \mathbf{v} + \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \cdot (\mathbf{x} - \mathbf{v}).$$

**Satz 6.12.** Seien  $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$  und  $g : \mathbb{K}^m \rightarrow \mathbb{K}^r$  lineare Abbildungen mit  $f(\mathbf{x}) = A\mathbf{x}$  und  $g(\mathbf{y}) = B\mathbf{y}$  mit  $A \in \mathbb{K}^{m \times n}$ ,  $B \in \mathbb{K}^{r \times m}$ . Dann ist  $g \circ f : \mathbb{K}^n \rightarrow \mathbb{K}^r$  wieder eine lineare Abbildung mit

$$(g \circ f)(\mathbf{x}) = (BA)\mathbf{x}.$$

**Satz 6.13.** Eine lineare Abbildung  $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$  mit  $f(\mathbf{x}) = A\mathbf{x}$  ist bijektiv genau dann, wenn  $A$  regulär ist. Die zugehörige Umkehrabbildung  $f^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n$  ist dann gegeben durch  $f^{-1}(\mathbf{y}) = A^{-1}\mathbf{y}$ .

**Definition 6.14** (Rotation(en) im  $\mathbb{R}^3$ ). Unter der Rotation im  $\mathbb{R}^3$  um die Koordinaten-Achsen  $\text{lin}\{\mathbf{e}_3\}$ ,  $\text{lin}\{\mathbf{e}_2\}$  oder  $\text{lin}\{\mathbf{e}_1\}$  mit Winkel  $\theta$  versteht man die Abbildung(en)

i)  $x_3$ -Achse;  $\text{rot}_{\mathbf{e}_3, \theta} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$\text{rot}_{\mathbf{e}_3, \theta}(\mathbf{x}) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

ii)  $x_2$ -Achse;  $\text{rot}_{\mathbf{e}_2, \theta} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$\text{rot}_{\mathbf{e}_2, \theta}(\mathbf{x}) = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

iii)  $x_1$ -Achse;  $\text{rot}_{e_1, \theta} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$\text{rot}_{e_1, \theta}(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

**Definition 6.15** (Isomorphe Vektorräume). Zwei  $\mathbb{K}$ -Vektorräume  $V, W$  heißen *isomorph zueinander*, falls es eine bijektive lineare Abbildung  $f : V \rightarrow W$  gibt.

**Satz 6.16.** Je zwei endlich dimensionale  $\mathbb{K}$ -Vektorräume der gleichen Dimension  $n$  sind zueinander isomorph.

**Definition 6.17.** Sei  $f : V \rightarrow W$  eine lineare Abbildung zwischen zwei  $\mathbb{K}$ -Vektorräumen  $V, W$ . Die Menge

$$\text{Kern}(f) = \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\} \subseteq V$$

heißt *Kern* der Abbildung  $f$ , und die Menge

$$\text{Bild}(f) = f(V) \subseteq W$$

heißt *Bild* der Abbildung  $f$ .

**Bemerkung 6.18.**  $\text{Kern}(f)$  ist Teilraum von  $V$ , und  $\text{Bild}(f)$  ist Teilraum von  $W$ .

**Satz 6.19.** Eine lineare Abbildung ist genau dann injektiv, wenn  $\text{Kern}(f) = \{\mathbf{0}\}$ .

**Satz 6.20** (Dimensionsformel). Sei  $f : V \rightarrow W$  eine lineare Abbildung zwischen zwei  $\mathbb{K}$ -Vektorräumen  $V, W$ , und sei  $V$  endlich-dimensional. Dann gilt:

$$\dim V = \dim \text{Bild}(f) + \dim \text{Kern}(f).$$

**Definition 6.21** (Rang einer Matrix). Sei  $A \in \mathbb{K}^{m \times n}$ . Die maximale Anzahl von linear unabhängigen Spalten in  $A$  heißt *Rang der Matrix*, und wird mit  $\text{rg}(A)$  bezeichnet.

**Satz 6.22.** Sei  $A \in \mathbb{K}^{m \times n}$ . Dann gilt  $\text{rg}(A) = \text{rg}(A^\top)$ , d.h. die maximale Anzahl linear unabhängiger Spalten einer Matrix ist gleich der maximalen Anzahl linear unabhängiger Zeilen.

**Satz 6.23.** Sei  $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$  mit  $f(\mathbf{x}) = A\mathbf{x}$ ,  $A \in \mathbb{K}^{m \times n}$ . Dann ist  $\text{rg}(A) = \dim_{\mathbb{K}} \text{Bild}(f)$ , und es ist

$$\text{Kern}(f) = \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}.$$

**Satz 6.24.** Sei  $A \in \mathbb{K}^{m \times n}$ , und sei  $L = \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}$ . Dann ist

$$\dim_{\mathbb{K}} L = n - \text{rg}(A).$$

## 7 Normierte Vektorräume

**Definition 7.1** (Skalarprodukt). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Eine Abbildung  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$  heißt *Skalarprodukt*, falls für alle  $\mathbf{x}, \mathbf{y}, \mathbf{v} \in V$  und  $\lambda, \mu \in \mathbb{K}$  gilt

- i)  $\langle \mathbf{x}, \mathbf{x} \rangle > 0$  falls  $\mathbf{x} \neq \mathbf{0}$  (Positivität),
- ii)  $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$  (Symmetrie),
- iii)  $\langle \lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{v} \rangle = \lambda \langle \mathbf{x}, \mathbf{v} \rangle + \mu \langle \mathbf{y}, \mathbf{v} \rangle$  (Linearität im ersten Argument).

**Definition 7.2.**

- i) Unter dem Standardskalarprodukt im  $\mathbb{R}^n$  versteht man

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y} = \sum_{i=1}^n x_i y_i,$$

für  $\mathbf{x} = (x_1, \dots, x_n)^\top, \mathbf{y} = (y_1, \dots, y_n)^\top \in \mathbb{R}^n$ .

- ii) Unter dem Standardskalarprodukt im  $\mathbb{C}^n$  versteht man

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \bar{\mathbf{y}} = \sum_{i=1}^n x_i \bar{y}_i,$$

für  $\mathbf{x} = (x_1, \dots, x_n)^\top, \mathbf{y} = (y_1, \dots, y_n)^\top \in \mathbb{C}^n$ .

**Bemerkung 7.3.** Im  $\mathbb{R}^n$  ist das Skalarprodukt auch linear im zweiten Argument, d.h. es gilt

$$\langle \mathbf{x}, \lambda \mathbf{y} + \mu \mathbf{v} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle + \mu \langle \mathbf{x}, \mathbf{v} \rangle, \text{ für alle } \mathbf{x}, \mathbf{y}, \mathbf{v} \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R}.$$

Hingegen gilt im  $\mathbb{C}^n$

$$\langle \mathbf{x}, \lambda \mathbf{y} + \mu \mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle + \bar{\mu} \langle \mathbf{x}, \mathbf{v} \rangle, \text{ für alle } \mathbf{x}, \mathbf{y}, \mathbf{v} \in \mathbb{C}^n, \lambda, \mu \in \mathbb{C}.$$

**Definition 7.4** (Norm). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Die *Länge* oder die *Norm* eines Vektors  $\mathbf{v} \in V$  ist definiert als

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

**Satz 7.5.** Für  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  gilt

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \|\mathbf{w}\| \cos \phi,$$

wobei  $\phi \in [0, \pi]$  der von  $\mathbf{v}$  und  $\mathbf{w}$  eingeschlossene Winkel ist.

**Satz 7.6** (Cauchy<sup>19</sup>-Schwarz<sup>20</sup>-Ungleichung). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$  und zugehöriger Norm  $\|\cdot\|$ . Dann gilt

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

mit Gleichheit dann und nur dann, wenn  $\mathbf{x}$  und  $\mathbf{y}$  linear abhängig sind.

---

<sup>19</sup>Augustin Louis Cauchy, 1789–1857

<sup>20</sup>Hermann Amandus Schwarz, 1843–1921

**Bemerkung 7.7.** Insbesondere gilt für  $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$ :

$$|x_1 y_1 + \dots + x_n y_n| \leq \sqrt{x_1^2 + \dots + x_n^2} \sqrt{y_1^2 + \dots + y_n^2}$$

mit Gleichheit dann und nur dann, wenn die Vektoren  $(x_1, \dots, x_n)^\top, (y_1, \dots, y_n)^\top \in \mathbb{R}^n$  linear abhängig sind.

**Bemerkung 7.8.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$  und zugehöriger Norm  $\|\cdot\|$ . Dann gilt

- i)  $\|\mathbf{x}\| \geq 0$  für alle  $\mathbf{x} \in V$ , wobei Gleichheit dann und nur dann gilt, wenn  $\mathbf{x} = \mathbf{0}$ .
- ii)  $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$  für alle  $\mathbf{x} \in V$ ,  $\lambda \in \mathbb{K}$ .
- iii)  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  für alle  $\mathbf{x}, \mathbf{y} \in V$  (Dreiecksungleichung).

**Definition 7.9** (Polarkoordinaten).

- i) Jedes  $\mathbf{x} = (x_1, x_2)^\top \in \mathbb{R}^2 \setminus \{0\}$  lässt sich eindeutig darstellen als (s. Definition 3.47)

$$\mathbf{x} = \|\mathbf{x}\| \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix}$$

mit  $\cos \phi = x_1 / \|\mathbf{x}\|$  und  $\sin \phi = x_2 / \|\mathbf{x}\|$ .

- ii) Jedes  $\mathbf{x} = (x_1, x_2, x_3)^\top \in \mathbb{R}^3 \setminus \{0\}$  lässt sich eindeutig darstellen als

$$\mathbf{x} = \|\mathbf{x}\| \begin{pmatrix} \sin \psi \cos \phi \\ \sin \psi \sin \phi \\ \cos \psi \end{pmatrix}$$

mit  $\cos \phi = x_1 / \sqrt{x_1^2 + x_2^2}$ ,  $\sin \phi = x_2 / \sqrt{x_1^2 + x_2^2}$ , und  $\cos \psi = x_3 / \|\mathbf{x}\|$  und  $\psi \in [0, \pi]$ .

Diese Darstellungen eines Vektors nennt man Darstellung in *Polarkoordinaten*.

**Beispiel 7.10** (Rotation im  $\mathbb{R}^3$ ). Sei  $\mathbf{v} \in \mathbb{R}^3 \setminus \{0\}$  mit

$$\mathbf{v} = \begin{pmatrix} \sin \psi \cos \phi \\ \sin \psi \sin \phi \\ \cos \psi \end{pmatrix}.$$

Die Rotation um die Achse  $\{\lambda v : \lambda \in \mathbb{R}\}$  mit Winkel  $\theta$  lässt sich durch die folgenden Rotationen beschreiben:

- i) Man drehe die Rotationsachse so, dass sie mit der  $x_3$ -Koordinaten-Achse übereinstimmt,
- ii) Man wende die Rotation mit Winkel  $\theta$  um die  $x_3$ -Achse an,
- iii) Man drehe die Rotationsachse wieder zurück in die Ausgangsposition.

Man erhält so die Abbildung:

$$(\text{rot}_{\mathbf{e}_3, \phi} \circ \text{rot}_{\mathbf{e}_2, -\psi}) \circ \text{rot}_{\mathbf{e}_3, \theta} \circ (\text{rot}_{\mathbf{e}_2, \psi} \circ \text{rot}_{\mathbf{e}_3, -\phi}).$$

**Definition 7.11** (Orthogonale Vektoren). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Zwei Vektoren  $\mathbf{v}, \mathbf{w} \in V$  heißen *orthogonal*, wenn  $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ .

**Definition 7.12** (Orthogonale Projektion). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ , und sei  $\mathbf{u} \in V$ ,  $\mathbf{u} \neq \mathbf{0}$ . Unter der orthogonalen Projektion auf den Teilraum  $U = \{\mathbf{w} \in V : \langle \mathbf{u}, \mathbf{w} \rangle = 0\}$  versteht man die lineare Abbildung  $f : V \rightarrow U$  mit

$$f(\mathbf{v}) = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \mathbf{u}.$$

Ihre Matrixdarstellung ist gegeben durch

$$I_n - \frac{1}{\|\mathbf{u}\|^2} \mathbf{u} \mathbf{u}^\top.$$

**Definition 7.13** (Orthonormalsystem, -basis). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Die Vektoren  $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$  bilden ein Orthonormalsystem, falls für  $1 \leq i, j \leq n$

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{i,j}, \quad \text{mit} \quad \delta_{i,j} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

$\delta_{i,j}$  heißt Kronecker-Symbol. Bilden  $\mathbf{u}_1, \dots, \mathbf{u}_n$  zudem eine Basis von  $V$ , dann heißt diese Basis *Orthonormalbasis* von  $V$ .

**Bemerkung 7.14.**

- i) Vektoren in einem Orthonormalsystem sind linear unabhängig.
- ii)  $\mathbf{e}_1, \dots, \mathbf{e}_n$  sind eine Orthonormalbasis in  $\mathbb{R}^n$ , bzw.  $\mathbb{C}^n$ .
- iii) Sei  $\mathbf{u}_1, \dots, \mathbf{u}_m$  ein Orthonormalsystem, und seien  $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ . Dann ist

$$\left\| \sum_{i=1}^m \alpha_i \mathbf{u}_i \right\| = \sqrt{\sum_{i=1}^m \alpha_i^2}.$$

**Satz 7.15** (Gram<sup>21</sup>-Schmidt<sup>22</sup>-Verfahren). Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ , und seien  $\mathbf{a}_1, \dots, \mathbf{a}_m \in V$  linear unabhängig. Für  $k = 1, \dots, m$  sei

$$\mathbf{u}_k = \mathbf{a}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{u}_j, \mathbf{a}_k \rangle}{\|\mathbf{u}_j\|^2} \mathbf{u}_j.$$

Dann bilden die Vektoren  $\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \dots, \frac{\mathbf{u}_m}{\|\mathbf{u}_m\|}$  ein Orthonormalsystem, und es gilt für  $1 \leq k \leq m$

$$\begin{aligned} \text{lin} \left\{ \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \dots, \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|} \right\} &= \text{lin} \{ \mathbf{u}_1, \dots, \mathbf{u}_k \} \\ &= \text{lin} \{ \mathbf{a}_1, \dots, \mathbf{a}_k \}. \end{aligned}$$

**Definition 7.16** (Orthogonale Matrix/Transformation). Eine Matrix  $U \in \mathbb{R}^{n \times n}$  heißt *orthogonale Matrix*, falls  $U^\top = U^{-1}$ , d.h.

$$U^\top U = U U^\top = I_n.$$

Die zugehörige lineare Abbildung  $f : V \rightarrow V$  mit  $f(\mathbf{x}) = U \mathbf{x}$  heißt *orthogonale Transformation*.

**Satz 7.17.** Sei  $U \in \mathbb{R}^{n \times n}$ . Dann sind die folgenden Aussagen äquivalent:

- i)  $U$  orthogonal,
- ii)  $U^\top$  ist orthogonal,
- iii) Die Spalten von  $U$  (und somit auch die Zeilen) bilden eine Orthonormalbasis des  $\mathbb{R}^n$ .

---

<sup>21</sup>Jørgen Pedersen Gram, 1850–1916

<sup>22</sup>Ehrhart Schmidt, 1876–1959

## 8 Homogene Koordinaten, Quaternionen und Projektionen

**Definition 8.1** (Homogene Koordinaten). Jedem Punkt  $\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$  wird der Punkt im  $\mathbb{R}^{n+1}$  mit den Koordinaten  $(x_1, \dots, x_n, 1)^\top$  zugeordnet. Diese Koordinaten heißen homogene Koordinaten von  $\mathbf{x}$ . Beim Rechnen mit homogenen Koordinaten wird jedem Punkt  $(x_1, \dots, x_n, x_{n+1})^\top \in \mathbb{R}^{n+1}$ ,  $x_{n+1} \neq 0$ , der Punkt  $(x_1/x_{n+1}, \dots, x_n/x_{n+1})^\top \in \mathbb{R}^n$  zugeordnet.

*Beispiel 8.2.* Der Punkt  $(-1, 2)^\top \in \mathbb{R}^2$  entspricht in homogenen Koordinaten dem Punkt  $(-1, 2, 1)^\top \in \mathbb{R}^3$ , bzw. den Punkten  $(-\lambda, 2\lambda, \lambda)^\top \in \mathbb{R}^3$ ,  $\lambda \neq 0$ .

**Bemerkung 8.3.**

- i) Sei  $\mathbf{t} = (t_1, t_2)^\top \in \mathbb{R}^2$ . Die Translation  $T_{\mathbf{t}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  um  $\mathbf{t}$  mit  $T_{\mathbf{t}}(\mathbf{x}) = \mathbf{x} + \mathbf{t}$  lässt sich in homogenen Koordinaten darstellen als

$$\begin{pmatrix} 1 & 0 & t_1 \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 + t_1 \\ x_2 + t_2 \\ 1 \end{pmatrix}.$$

- ii) Die Rotation  $\text{rot}_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  um den Ursprung mit Winkel  $\theta$  lässt sich in homogenen Koordinaten darstellen als

$$\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta x_1 - \sin \theta x_2 \\ \sin \theta x_1 + \cos \theta x_2 \\ 1 \end{pmatrix}.$$

- iii) Die Rotation um den Punkt  $\mathbf{v} = (v_1, v_2)^\top$  mit Winkel  $\theta$  lässt sich in homogenen Koordinaten darstellen als

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & v_1 \\ 0 & 1 & v_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -v_1 \\ 0 & 1 & -v_2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta & -\sin \theta & -\cos \theta v_1 + \sin \theta v_2 + v_1 \\ \sin \theta & \cos \theta & -\sin \theta v_1 - \cos \theta v_2 + v_2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}. \end{aligned}$$

- iv) Sei  $\mathbf{t} = (t_1, t_2, t_3)^\top \in \mathbb{R}^3$ . Die Translation  $T_{\mathbf{t}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  um  $\mathbf{t}$  mit  $T_{\mathbf{t}}(\mathbf{x}) = \mathbf{x} + \mathbf{t}$  lässt sich in homogenen Koordinaten darstellen als

$$\begin{pmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 + t_1 \\ x_2 + t_2 \\ x_3 + t_3 \\ 1 \end{pmatrix}.$$

- v) Die Rotation um die  $x_3$ -Achse mit Winkel  $\theta$  lässt sich mittels homogenen Koordinaten darstellen als

$$\begin{pmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta x_1 - \sin \theta x_2 \\ \sin \theta x_1 + \cos \theta x_2 \\ x_3 \\ 1 \end{pmatrix}.$$

Entsprechend lassen sich die Rotationen um die  $x_1$ - oder  $x_2$ -Achse darstellen. Mittels Verknüpfungen von Rotationen und Translationen lassen sich somit Rotationen um beliebige Achsen durch Matrixmultiplikationen ausdrücken.

**Bemerkung 8.4.** Eine Rotation in  $\mathbb{R}^2 \cong \mathbb{C}$  um den Ursprung mit Winkel  $\theta$  kann auch als Multiplikation mit der komplexen Zahl  $e^{i\theta} = \cos \theta + i \sin \theta$  aufgefasst werden.

**Definition 8.5** (Quaternionen). Die Quaternionen sind der 4-dimensionale  $\mathbb{R}$ -Vektorraum

$$\mathbb{H} = \{a + bi + cj + dk : a, b, c, d \in \mathbb{R}\},$$

zusammen mit einer Multiplikation, gemäss den üblichen Rechenregeln in  $\mathbb{R}$  und den speziellen Rechenregeln

$$i^2 = j^2 = k^2 = ijk = -1$$

für die Zahlen  $i, j, k$ .

**Bemerkung 8.6.** Für die Multiplikation in  $\mathbb{H}$  gilt das Assoziativgesetz und das Distributivgesetz, nicht aber die Kommutativität. Insbesondere gilt

$$ij = k = -ji, \quad jk = i = -kj, \quad \text{und} \quad ki = j = -ik.$$

Die Quaternionen bilden einen sogenannten Schiefkörper, d.h. zusammen mit den Verknüpfungen  $+$  und  $\cdot$  haben sie alle Eigenschaften eines Körpers, bis auf die Kommutativität der Multiplikation.

**Definition 8.7** (Konjugation und Norm in  $\mathbb{H}$ ). Sei  $r = a + bi + cj + dk \in \mathbb{H}$ . Dann ist

- i)  $\bar{r} = a - bi - cj - dk$  konjugiert zu  $r$ ,
- ii)  $\|r\| = \sqrt{a^2 + b^2 + c^2 + d^2}$  die Norm von  $r$ .

**Bemerkung 8.8.** Für  $r, s \in \mathbb{H}$  gilt:

- i)  $\|r\| = \sqrt{r\bar{r}}$ ,
- ii)  $\bar{r}s = \bar{s}r$ .
- iii)  $\|rs\| = \|r\|\|s\|$ ,
- iv) Für  $r \neq 0$  ist  $r^{-1} = \frac{\bar{r}}{\|r\|^2}$  (multiplikatives) Inverses zu  $r$ .

**Satz 8.9.** Sei  $\text{Im}(\mathbb{H}) = \{xi + yj + zk : x, y, z \in \mathbb{R}\}$  und sei  $r = a + bi + cj + dk \in \mathbb{H} \setminus \{0\}$ . Die Abbildung  $\mathbb{R}_r : \text{Im}(\mathbb{H}) \rightarrow \text{Im}(\mathbb{H})$  mit  $R_r(p) = rpr^{-1}$  ist bijektiv und es gilt

- i)  $R_{\lambda r} = R_r$  für alle  $\lambda \in \mathbb{R} \setminus \{0\}$ ,
- ii)  $(R_r)^{-1} = (R_{r^{-1}})$ ,
- iii)  $(R_{\pm 1}) = \text{Identität}$ ,
- iv)  $R_r \circ R_s = R_{r \cdot s}$ .

**Satz 8.10.** Sei  $\text{Im}(\mathbb{H}) = \{xi + yj + zk : x, y, z \in \mathbb{R}\}$  und sei  $r = a + bi + cj + dk \in \mathbb{H} \setminus \{0\}$ .

- i) Identifiziert man  $p = xi + yj + zk \in \text{Im}(\mathbb{H})$  mit dem Vektor  $(x, y, z) \in \mathbb{R}^3$ , dann beschreibt  $R_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  eine Rotation, und zwar für  $\|r\| = 1$  ist  $R_r$  die Rotation um  $\{\lambda \mathbf{v} : \lambda \in \mathbb{R}\}$  mit den Winkel  $\theta$ , wobei

$$\mathbf{v} = (b, c, d)^\top \quad \text{und} \quad \theta = 2 \arccos(a) = 2 \arcsin(\|v\|).$$

- ii) Ist umgekehrt  $\mathbf{v} = (v_1, v_2, v_3)^\top \in \mathbb{R}^3$  ein Punkt mit  $\|\mathbf{v}\| = 1$ , dann lässt sich die Drehung um die Drehachse  $\{\lambda \mathbf{v} : \lambda \in \mathbb{R}\}$  mit dem Winkel  $\theta$  darstellen als die Abbildung  $R_r$ , wobei

$$r = \cos(\phi/2) + \sin(\phi/2) v_1 i + \sin(\phi/2) v_2 j + \sin(\phi/2) v_3 k.$$

- iii) Somit lässt sich die Hintereinanderausführung von "Drehungen"  $(R_r \circ R_s)(p)$  durch Quaternionenmultiplikation  $rsp s^{-1} r^{-1}$  ausdrücken.

**Definition 8.11** (Zentralprojektion im  $\mathbb{R}^n$ ). Sei  $A = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{x} \rangle = b\}$  mit  $\mathbf{a} \in \mathbb{R}^n \setminus \{0\}$  und  $b \in \mathbb{R}$  die Abbildungsebene (Projektionsebene, Betrachtungsebene), und sei  $\mathbf{v} \in \mathbb{R}^n$  der Betrachtungspunkt (das Projektionszentrum) mit  $\mathbf{v} \notin A$ . Unter der Zentralprojektion  $Z_\pi : \mathbb{R}^n \rightarrow A$  versteht man die Abbildung, die jedem Punkt  $\mathbf{p} \in \mathbb{R}^n$  den Schnittpunkt  $\mathbf{p}'$  (falls vorhanden) der Geraden durch  $\mathbf{p}$  und  $\mathbf{v}$  mit  $A$  zuordnet.

**Satz 8.12.** Ist die Gerade durch  $\mathbf{p}$  und  $\mathbf{v}$  nicht parallel zu  $A$ , d.h.  $\langle \mathbf{a}, \mathbf{p} - \mathbf{v} \rangle \neq 0$ , dann gibt es einen eindeutigen Schnittpunkt  $\mathbf{p}'$ , für den gilt:

$$\mathbf{p}' = \left( \frac{\langle \mathbf{a}, \mathbf{p} \rangle - b}{\langle \mathbf{a}, \mathbf{p} - \mathbf{v} \rangle} \right) \mathbf{v} - \left( \frac{\langle \mathbf{a}, \mathbf{v} \rangle - b}{\langle \mathbf{a}, \mathbf{p} - \mathbf{v} \rangle} \right) \mathbf{p}.$$

**Satz 8.13** (Darstellung der Zentralprojektion in homogenen Koordinaten). Sei  $\bar{\mathbf{a}} = (\mathbf{a}, -b)^\top \in \mathbb{R}^{n+1}$  und  $\bar{A} = \{x \in \mathbb{R}^{n+1} : \langle \bar{\mathbf{a}}, \mathbf{x} \rangle = 0\}$  die Abbildungsebene in "homogenen" Koordinaten. Sei  $\bar{\mathbf{v}} = (\mathbf{v}, 1)^\top \in \mathbb{R}^{n+1}$  der Betrachtungspunkt und  $\bar{\mathbf{p}} = (\mathbf{p}, 1)^\top \in \mathbb{R}^{n+1}$ . Dann ist

$$Z_\pi(\bar{\mathbf{p}}) = \langle \bar{\mathbf{a}}, \bar{\mathbf{p}} \rangle \bar{\mathbf{v}} - \langle \bar{\mathbf{a}}, \bar{\mathbf{v}} \rangle \bar{\mathbf{p}} = (\bar{\mathbf{v}} \bar{\mathbf{a}}^\top - \bar{\mathbf{a}}^\top \bar{\mathbf{v}} I_{n+1}) \bar{\mathbf{p}}.$$

Hierbei bezeichnet  $I_{n+1}$  die  $(n+1) \times (n+1)$ -Einheitsmatrix.

**Bemerkung 8.14.** Für  $n = 3$  erhält man

$$Z_\pi(\bar{\mathbf{p}}) = \begin{pmatrix} -(\bar{a}_2 \bar{v}_2 + \bar{a}_3 \bar{v}_3 + \bar{a}_4 \bar{v}_4) & -(\bar{a}_1 \bar{v}_1 + \bar{a}_3 \bar{v}_3 + \bar{a}_4 \bar{v}_4) & \bar{a}_3 \bar{v}_1 \\ \bar{a}_1 \bar{v}_2 & \bar{a}_2 \bar{v}_3 & \bar{a}_3 \bar{v}_2 \\ \bar{a}_1 \bar{v}_3 & \bar{a}_2 \bar{v}_4 & \bar{a}_3 \bar{v}_4 \end{pmatrix} \bar{\mathbf{p}}.$$

**Definition 8.15** (Parallelprojektion im  $\mathbb{R}^n$ ). Sei  $A = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{x} \rangle = b\}$  mit  $\mathbf{a} \in \mathbb{R}^n \setminus \{0\}$  und  $b \in \mathbb{R}$  die Abbildungsebene (Projektionsebene, Betrachtungsebene), und sei  $\mathbf{v} \in \mathbb{R}^n$  die Projektionsrichtung mit  $\langle \mathbf{a}, \mathbf{v} \rangle \neq 0$ . Unter der Parallelprojektion  $P_\pi : \mathbb{R}^n \rightarrow A$  versteht man die Abbildung, die jedem Punkt  $\mathbf{p} \in \mathbb{R}^n$  den Schnittpunkt  $\mathbf{p}'$  (falls vorhanden) der Geraden durch  $\mathbf{p}$  und Richtung  $\mathbf{v}$  mit  $A$  zuordnet.

**Satz 8.16.** Ist die Gerade mit Richtung  $v$  nicht parallel zu  $A$ , d.h.  $\langle \mathbf{a}, \mathbf{v} \rangle \neq 0$ , dann gibt es einen eindeutigen Schnittpunkt  $\mathbf{p}'$ , für den gilt:

$$\mathbf{p}' = \mathbf{p} - \left( \frac{\langle \mathbf{a}, \mathbf{p} \rangle - b}{\langle \mathbf{a}, \mathbf{v} \rangle} \right) \mathbf{v}.$$

**Satz 8.17** (Darstellung der Parallelprojektion in homogenen Koordinaten). Sei  $\bar{\mathbf{a}} = (\mathbf{a}, -b)^\top \in \mathbb{R}^{n+1}$  und  $\bar{A} = \{\mathbf{x} \in \mathbb{R}^{n+1} : \langle \bar{\mathbf{a}}, \mathbf{x} \rangle = 0\}$  die Abbildungsebene in homogenen Koordinaten. Sei  $\bar{\mathbf{v}} = (\mathbf{v}, 0)^\top \in \mathbb{R}^{n+1}$  die Projektionsrichtung und  $\bar{\mathbf{p}} = (\mathbf{p}, 1)^\top \in \mathbb{R}^{n+1}$ . (Die Punkte auf der Geraden  $\mathbf{p} + \lambda \mathbf{v}$  besitzen die homogenen Koordinaten  $\bar{\mathbf{p}} + \lambda \bar{\mathbf{v}}$ ). Dann ist

$$P_\pi(\bar{\mathbf{p}}) = \langle \bar{\mathbf{a}}, \bar{\mathbf{p}} \rangle \bar{\mathbf{v}} - \langle \bar{\mathbf{a}}, \bar{\mathbf{v}} \rangle \bar{\mathbf{p}} = (\bar{\mathbf{v}} \bar{\mathbf{a}}^\top - \bar{\mathbf{a}}^\top \bar{\mathbf{v}} I_{n+1}) \bar{\mathbf{p}}.$$

**Bemerkung 8.18.** Für  $n = 3$  erhält man

$$P_\pi(\bar{\mathbf{p}}) = \begin{pmatrix} -(\bar{a}_2 \bar{v}_2 + \bar{a}_3 \bar{v}_3) & -(\bar{a}_1 \bar{v}_1 + \bar{a}_3 \bar{v}_3) & \bar{a}_3 \bar{v}_1 \\ \bar{a}_1 \bar{v}_2 & \bar{a}_2 \bar{v}_3 & \bar{a}_3 \bar{v}_2 \\ 0 & 0 & \bar{a}_1 \bar{v}_1 + \bar{a}_2 \bar{v}_2 \end{pmatrix} \bar{\mathbf{p}}.$$

**Bemerkung 8.19.** Parallelprojektion ist nur ein Spezialfall der Zentralprojektion. Der Betrachtungspunkt liegt bei der Parallelprojektion im Unendlichen. Betrachtet man die beiden Matrizen, so erhält man die Matrix für die Parallelprojektion aus der entsprechenden für die Zentralprojektion, indem man  $\bar{\mathbf{v}}_4 = 0$  (allgemein  $\bar{\mathbf{v}}_{n+1} = 0$ ) setzt.

## 9 Determinanten und Eigenwerte

**Notation 9.1.** Für eine  $n \times n$ -Matrix  $A \in \mathbb{K}^{n \times n}$  und  $i, j \in \{1, \dots, n\}$  bezeichnet  $A^{i,j} \in \mathbb{K}^{(n-1) \times (n-1)}$  die  $(n-1) \times (n-1)$  Matrix, die aus  $A$  durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht.

**Definition 9.2** (Determinante, Entwicklungssatz von Laplace<sup>23</sup>). Sei  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ . Die *Determinante* von  $A$ , bezeichnet mit  $\det(A)$ , ist eine Zahl aus  $\mathbb{K}$ , die wie folgt rekursiv definiert ist:

- i) Ist  $n = 1$ , so ist  $\det(A) = a_{11}$ .
- ii) Für  $n > 1$  und einen Zeilenindex  $i \in \{1, \dots, n\}$  ist

$$\begin{aligned}\det(A) &= \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(A^{i,j}) \\ &= (-1)^{i+1} a_{11} \det(A^{1,1}) + (-1)^{i+2} a_{12} \det(A^{1,2}) + \cdots + (-1)^{i+n} a_{1n} \det(A^{1,n}).\end{aligned}$$

(Entwicklung nach der  $i$ -ten Zeile)

*Beispiel 9.3.*

- i) Für  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$  ist
- $$\det(A) = a_{11} a_{22} - a_{12} a_{21}.$$

Dies entspricht dem (orientierten) Flächeninhalt des von den Spaltenvektoren aufgespannten Parallelogramms.

- ii) Sei  $A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 0 & 5 \\ 7 & 6 & 8 \end{pmatrix}$ . Entwicklung nach der 2-ten Zeile (also  $i = 2$ ) ergibt

$$\begin{aligned}\det(A) &= (-1)^{2+1} 4 \det(A^{2,1}) + (-1)^{2+2} 0 \det(A^{2,2}) + (-1)^{2+3} 5 \det(A^{2,3}) \\ &= -4 \det\left(\begin{pmatrix} 1 & 3 \\ 6 & 8 \end{pmatrix}\right) - 5 \det\left(\begin{pmatrix} 2 & 1 \\ 7 & 6 \end{pmatrix}\right) \\ &= -4(8 - 18) - 5(12 - 7) = 40 - 25 = 15.\end{aligned}$$

**Satz 9.4.** Für  $A \in \mathbb{K}^{n \times n}$  ist  $\det(A) = \det(A^\top)$ . Insbesondere kann  $\det(A)$  auch durch Entwicklung nach der  $j$ -ten Spalte berechnet werden, d.h. für  $j \in \{1, \dots, n\}$  ist

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

*Beispiel 9.5.* Sei  $A = \begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 5 \\ 2 & 5 & 13 \end{pmatrix}$ . Entwicklung nach der 1-ten Spalte (also  $j = 1$ ) ergibt

$$\begin{aligned}\det(A) &= (-1)^{1+1} 0 \det(A^{1,1}) + (-1)^{2+1} 1 \det(A^{2,1}) + (-1)^{3+1} 2 \det(A^{3,1}) \\ &= -1 \det\left(\begin{pmatrix} 1 & 3 \\ 5 & 13 \end{pmatrix}\right) + 2 \det\left(\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}\right) \\ &= -(13 - 15) + 2(5 - 6) = 0.\end{aligned}$$

---

<sup>23</sup>Pierre-Simon Laplace, 1749-1827

**Satz 9.6** (Eigenschaften der Determinante).

- i) Die Determinante einer (oberen oder unteren) Dreiecksmatrix ist gleich dem Produkt ihrer Diagonalelemente. Insbesondere ist  $\det(I_n) = 1$ .
- ii) Die Determinante ist alternierend, d.h. bei Vertauschen zweier Zeilen ändert sich das Vorzeichen der Determinante.
- iii) Die Determinante ist linear in jeder Zeile, d.h. seien  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{K}^n$  Zeilenvektoren und sei  $\mathbf{v}$  ein weiterer Zeilenvektor,  $\lambda \in \mathbb{K}$ . Dann ist

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i + \mathbf{v} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{v} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

und

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \lambda \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \lambda \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

**Bemerkung 9.7.** Wegen  $\det(A) = \det(A^\top)$  gelten die Eigenschaften ii) und iii) aus Satz 9.6 auch, wenn man „Zeile“ durch „Spalte“ ersetzt.

**Korollar 9.8.** Sei  $A \in \mathbb{K}^{n \times n}$  und  $\lambda \in \mathbb{K}$ .

- i) Enthält  $A$  eine Zeile (Spalte) mit lauter Nullen, dann ist  $\det(A) = 0$ .
- ii) Enthält  $A$  zwei gleiche Zeilen (Spalten), dann ist  $\det(A) = 0$ .
- iii) Addition des  $\lambda$ -fachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte) ändert die Determinante nicht.
- iv) Sind die Zeilen (Spalten) von  $A$  linear abhängig, dann ist  $\det(A) = 0$ .
- v)  $\det(\lambda A) = \lambda^n \det(A)$ .

**Bemerkung 9.9.** Aufgrund von Satz 9.6 i), ii) und Korollar 9.8 iii) lässt sich die Determinante auch berechnen, indem man die Matrix mit dem Gauß-Algorithmus auf Dreiecksform (Zeilenstufenform) bringt.

**Beispiel 9.10.** Sei  $A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 0 & 5 \\ 7 & 6 & 8 \end{pmatrix}$ .

$$\begin{pmatrix} 2 & 1 & 3 \\ 4 & 0 & 5 \\ 7 & 6 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 3 \\ 0 & -2 & -1 \\ 0 & \frac{5}{2} & -\frac{5}{2} \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 3 \\ 0 & -2 & -1 \\ 0 & 0 & -\frac{15}{4} \end{pmatrix}$$

Also  $\det(A) = 2 \cdot (-2) \cdot (-15/4) = 15$ .

**Satz 9.11.** Für  $A \in \mathbb{K}^{n \times n}$  sind die folgenden Aussagen äquivalent:

- i)  $A$  ist regulär (invertierbar).
- ii) Die Zeilen (Spalten) von  $A$  sind linear unabhängig.
- iii)  $\text{rg}(A) = n$ .
- iv)  $\det(A) \neq 0$ .

**Satz 9.12.** Seien  $A, B \in \mathbb{K}^{n \times n}$ . Dann ist  $\det(A \cdot B) = \det(A) \cdot \det(B)$ .

**Korollar 9.13.**

- i) Sei  $A \in \mathbb{K}^{n \times n}$  regulär. Dann ist  $\det(A^{-1}) = \frac{1}{\det(A)}$ .
- ii) Sei  $U \in \mathbb{K}^{n \times n}$  orthogonal. Dann ist  $|\det(U)| = 1$ .

**Bemerkung 9.14.** Seien  $A, B \in \mathbb{K}^{n \times n}$ . Im Allgemeinen ist  $\det(A + B) \neq \det(A) + \det(B)$ .

**Definition 9.15** (Eigenwert, Eigenvektor). Sei  $A \in \mathbb{K}^{n \times n}$ ,  $\lambda \in \mathbb{K}$  heißt *Eigenwert* der Matrix  $A$  (bzw. der linearen Abbildung  $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ), wenn es einen Vektor  $\mathbf{u} \in \mathbb{K}^n$ ,  $\mathbf{u} \neq \mathbf{0}$ , gibt, so dass

$$A \mathbf{u} = \lambda \mathbf{u}.$$

$\mathbf{u}$  heißt *Eigenvektor* zum Eigenwert  $\lambda$ .

*Beispiel 9.16.* Sei  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  (Spiegelung an der 45-Grad Achse). 1 und -1 sind die Eigenwerte von  $A$ . Die Eigenvektoren zum Eigenwert 1 sind alle Vektoren der Form  $(\alpha, \alpha)^T$ ,  $\alpha \neq 0$ . Die Eigenvektoren zum Eigenwert -1 sind alle Vektoren der Form  $(-\alpha, \alpha)^T$ ,  $\alpha \neq 0$ .

**Bemerkung 9.17.** Sei  $A \in \mathbb{K}^{n \times n}$ ,  $\lambda \in \mathbb{K}$  ist Eigenwert von  $A$  genau dann, wenn es  $\mathbf{u} \neq \mathbf{0} \in \mathbb{K}^n$  gibt mit

$$(A - \lambda I_n) \mathbf{u} = \mathbf{0},$$

d.h. falls die Spaltenvektoren von  $A - \lambda I_n$  linear abhängig sind, was wiederum äquivalent dazu ist, dass  $\det(A - \lambda I_n) = 0$ .

**Definition 9.18** (Charakteristisches Polynom). Sei  $A \in \mathbb{K}^{n \times n}$ . Dann ist

$$\chi_A(\lambda) = \det(A - \lambda I_n)$$

ein Polynom vom Grad  $n$  in  $\lambda$ . Es heißt das *charakteristische Polynom* von  $A$ . Die Nullstellen  $\lambda \in \mathbb{K}$  von  $\chi_A(\lambda)$  sind die Eigenwerte von  $A$ .

*Beispiel 9.19.*

- i) Sei  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Dann ist

$$\begin{aligned} \chi_A(\lambda) &= \det(A - \lambda I_2) = \det \left( \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} \right) \\ &= \lambda^2 - 1 = (\lambda + 1)(\lambda - 1). \end{aligned}$$

Die Eigenwerte von  $A$  sind 1 und -1.

- ii) Sei  $A = I_n$ . Dann ist

$$\chi_A(\lambda) = \det(I_n - \lambda I_n) = (1 - \lambda)^n.$$

Eigenwert von  $I_n$  ist nur 1.

**Satz 9.20.** Sei  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$  und sei

$$\chi_A(\lambda) = \det(A - \lambda I_n) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0,$$

mit  $a_i \in \mathbb{K}$ . Dann ist  $a_n = (-1)^n$ ,  $a_0 = \det(A)$  und  $a_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$ .

**Definition 9.21** (Spur). Sei  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ . Die Summe der Diagonalelemente von  $A$  heißt *Spur* der Matrix  $A$  und wird mit  $\text{tr}(A)$  bezeichnet, d.h.

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

**Bemerkung 9.22.**  $A$  und  $A^\top$  haben das gleiche charakteristische Polynom.

**Bemerkung 9.23.** Sei  $A \in \mathbb{K}^{n \times n}$ , und sei  $\lambda \in \mathbb{K}$  Eigenwert von  $A$ .  $\mathbf{u} \neq \mathbf{0}$  ist Eigenvektor zum Eigenwert  $\lambda$  genau dann, wenn

$$(A - \lambda I_n) \mathbf{u} = \mathbf{0}.$$

Somit ist die Menge aller Eigenvektoren zum Eigenwert  $\lambda$  gegeben durch die von Null verschiedenen Lösungen des homogenen Gleichungssystems  $(A - \lambda I_n) \mathbf{x} = \mathbf{0}$ , was wiederum gleich der Menge  $\text{Kern}(A - \lambda I_n) \setminus \{\mathbf{0}\}$  ist.

**Definition 9.24** (Eigenraum). Sei  $A \in \mathbb{K}^{n \times n}$ , und sei  $\lambda \in \mathbb{K}$  Eigenwert von  $A$ .  $\text{Kern}(A - \lambda I_n)$ , d.h. die Menge aller Eigenvektoren zum Eigenwert  $\lambda$  (plus den Nullvektor), heißt *Eigenraum* zum Eigenwert  $\lambda$ .

**Bemerkung 9.25** (Bestimmen der Eigenwerte und Eigenräume/-vektoren). Sei  $A \in \mathbb{K}^{n \times n}$ .

- i) Berechne das charakteristische Polynom  $\chi_A(\lambda)$ .
- ii) Bestimme alle Nullstellen  $\lambda_1, \dots, \lambda_k \in \mathbb{K}$  von  $\chi_A(\lambda)$ .
- iii) Für  $\lambda_1, \dots, \lambda_k$  bestimme man die Eigenräume, d.h.  $\text{Kern}(A - \lambda_i I_n)$ ,  $1 \leq i \leq k$ .

*Beispiel 9.26.* Sei  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .  $A$  hat die beiden Eigenwerte  $\lambda_1 = 1$  und  $\lambda_2 = -1$ .

- $\lambda_1 = 1$ . In diesem Fall ist  $A - \lambda_1 I_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ , und

$$\begin{aligned} \text{Kern}(A - \lambda_1 I_2) &= \left\{ \mathbf{x} \in \mathbb{R}^2 : \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{x} = \mathbf{0} \right\} \\ &= \{(\alpha, \alpha)^\top : \alpha \in \mathbb{R}\}. \end{aligned}$$

- $\lambda_2 = -1$ . In diesem Fall ist  $A - \lambda_2 I_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ , und

$$\begin{aligned} \text{Kern}(A - \lambda_2 I_2) &= \left\{ \mathbf{x} \in \mathbb{R}^2 : \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \mathbf{x} = \mathbf{0} \right\} \\ &= \{(-\alpha, \alpha)^\top : \alpha \in \mathbb{R}\}. \end{aligned}$$

**Bemerkung 9.27.** Sei  $U \in \mathbb{K}^{n \times n}$ ,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , eine orthogonale Matrix. Dann gilt für jeden Eigenwert  $\lambda$  von  $U$

$$|\lambda| = 1.$$

**Definition 9.28** (Ähnliche Matrizen). Zwei Matrizen  $A, B \in \mathbb{K}^{n \times n}$  heißen *ähnlich*, falls es eine reguläre Matrix  $C \in \mathbb{K}^{n \times n}$  gibt mit

$$B = C^{-1} A C.$$

**Bemerkung 9.29.** Seien  $A, B \in \mathbb{K}^{n \times n}$  ähnliche Matrizen und  $B = C^{-1} A C$  mit  $C \in \mathbb{K}^{n \times n}$  reguläre. Dann ist für  $m \in \mathbb{N}$

$$B^m = C^{-1} A^m C.$$

**Satz 9.30.** Ähnliche Matrizen haben das gleiche charakteristische Polynom.

**Definition 9.31** (Diagonalisierbarkeit). Eine Matrix  $A \in \mathbb{K}^{n \times n}$  (bzw. eine lineare Abbildung  $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ) heißt *diagonalisierbar*, falls es eine reguläre Matrix  $C \in \mathbb{K}^{n \times n}$  gibt, so dass  $C^{-1} A C$  eine Diagonalmatrix ist, d.h.  $A$  ist ähnlich zu einer Diagonalmatrix.

**Satz 9.32.** Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

**Satz 9.33.** Sei  $A \in \mathbb{K}^{n \times n}$ . Dann sind die folgenden Aussagen äquivalent:

- i)  $A$  ist diagonalisierbar.
- ii) Es gibt eine Basis des  $\mathbb{K}^n$  bestehend aus Eigenvektoren von  $A$ .
- iii) Das charakteristische Polynom  $\chi_A(\lambda)$  zerfällt über  $\mathbb{K}$  in Linearfaktoren, d.h. es gibt  $\lambda_1, \dots, \lambda_k$ ,  $\lambda_i \neq \lambda_j$  für  $i \neq j$ , und  $m_i \in \mathbb{N}$  mit

$$\chi_A(\lambda) = (-1)^n (\lambda - \lambda_1)^{m_1} \cdot (\lambda - \lambda_2)^{m_2} \cdots \cdot (\lambda - \lambda_k)^{m_k},$$

und  $m_i = \dim_{\mathbb{K}} \text{Kern}(A - \lambda_i I_n)$ ,  $1 \leq i \leq k$ , d.h. die algebraische Vielfachheit ( $m_i$ ) ist gleich der geometrischen Vielfachheit ( $\dim_{\mathbb{K}} \text{Kern}(A - \lambda_i I_n)$ ) für alle Eigenwerte.

**Satz 9.34.** Jede symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  ist diagonalisierbar, und es gibt sogar eine Orthonormalbasis des  $\mathbb{R}^n$  bestehend aus Eigenvektoren von  $A$ . Insbesondere gibt es eine orthogonale Matrix  $U$ , so dass  $U^\top A U$  eine Diagonalmatrix ist.

**Definition 9.35** (Positiv (negativ) definit). Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch.  $A$  heißt *positiv definit*, falls

$$\mathbf{x}^\top A \mathbf{x} > 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

Ist hingegen

$$\mathbf{x}^\top A \mathbf{x} < 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\},$$

dann heißt  $A$  *negativ definit*.

Gilt statt  $>$  nur  $\geq 0$ , bzw. statt  $<$  nur  $\leq$ , so heißen die Matrizen *positiv semi-definit*, bzw. *negativ semi-definit*. Trifft keine der Bedingungen zu, heisst die Matrix *indefinit*.

**Satz 9.36.** Eine symmetrische Matrix  $A$  ist genau dann positiv (negativ) definit, wenn alle Eigenwerte von  $A$  positiv (negativ) sind. Sie ist genau dann positiv (negativ) semi-definit, wenn alle Eigenwerte von  $A \geq 0$  ( $\leq 0$ ) sind.

**Satz 9.37** (Hurwitz<sup>24</sup>-Kriterium). Eine symmetrische  $n \times n$ -Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  ist genau dann positiv definit, wenn für  $1 \leq m \leq n$  gilt:

$$\det((a_{ij})_{i,j=1}^m) > 0.$$

---

<sup>24</sup>Adolf Hurwitz, 1859-1919

**Definition 9.38** (Stochastische Matrix, Markov<sup>25</sup>-Matrix). Eine Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  heißt *stochastische Matrix* oder *Markov-Matrix* falls

- i) alle Einträge nichtnegativ sind, d.h.  $a_{ij} \geq 0$ ,  $1 \leq i, j \leq n$ , und
- ii) die Summen der Einträge in jeder Spalte gleich eins ist, d.h.  $\sum_{i=1}^n a_{ij} = 1$  für  $1 \leq j \leq n$ .

**Satz 9.39.** Sei  $A$  eine stochastische Matrix. Alle Eigenwerte von  $A$  sind von Betrag kleiner gleich 1, und sie hat immer den Eigenwert 1. Zum Eigenwert 1 gibt es immer einen Eigenvektor mit nichtnegativen Komponenten.

---

<sup>25</sup>Andrej Andrejewitsch Markov, 1856–1922

**Teil II**  
**Mathematik II**

## 10 Gruppen, Ringe, Körper

**Definition 10.1** (Gruppe (vergl. Def. 3.36)). Sei  $G$  eine nichtleere Menge, und sei  $\otimes : G \times G \rightarrow G$  mit  $(x, y) \mapsto x \otimes y$  eine Abbildung (Verknüpfung).

$(G, \otimes)$  heißt *Gruppe*, wenn die folgenden Bedingungen 1.–3. erfüllt sind:

1. Für alle  $x, y, z \in G$  gilt:  $(x \otimes y) \otimes z = x \otimes (y \otimes z)$  (Assoziativgesetz).
  2. Es gibt ein *neutrales Element*  $e \in G$ , so dass für alle  $x \in G$  gilt:  $e \otimes x = x \otimes e = x$ .
  3. Zu jedem  $x \in G$  gibt es ein *inverses Element*  $x' \in G$ , so dass  $x \otimes x' = x' \otimes x = e$ .  $x'$  wird auch mit  $x^{-1}$  bezeichnet.
- Gilt zusätzlich  $x \otimes y = y \otimes x$  für alle  $x, y \in G$ , dann heißt  $(G, \otimes)$  *kommutative (abelsche<sup>26</sup>) Gruppe*.

*Beispiel 10.2.* (siehe auch Kapitel 3)

- $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +)$  sind kommutative Gruppen mit neutralem Element 0, und für  $a$  ist  $-a$  das inverse Element.
- $(\mathbb{Q} \setminus \{0\}, \cdot)$  ist kommutative Gruppe mit neutralem Element 1, und für  $a \in \mathbb{Q} \setminus \{0\}$  ist  $1/a$  das inverse Element. Ebenso ist  $(\mathbb{R} \setminus \{0\}, \cdot)$  kommutative Gruppe.
- $(S_n, \circ)$  ist Gruppe mit neutralem Element  $\text{id}_{\{1, \dots, n\}}$ , und das inverse Element von  $\sigma \in S_n$  ist durch die Umkehrabbildung gegeben.  $S_n$  ist nicht kommutativ für  $n \geq 3$ .
- $(\mathbb{Z}_m, \oplus)$  ist kommutative Gruppe mit neutralem Element  $[0]_m$ , und für  $[a]_m \in \mathbb{Z}_m$  ist  $[-a]_m = [m-a]_m$  das inverse Element.
- $(\mathbb{Z}_4 \setminus \{[0]_4\}, \odot)$  ist keine Gruppe, aber  $(\mathbb{Z}_5 \setminus \{[0]_5\}, \odot)$  ist kommutative Gruppe.
- $(\mathbb{K}^n, +)$  (also insbesondere  $(\mathbb{R}^n, +)$ ) ist eine abelsche Gruppe für einen Körper  $\mathbb{K}$ .
- $\text{GL}(n, \mathbb{R}) = \{A \in \mathbb{R}^{n \times n} : \det A \neq 0\}$  ist eine Gruppe mit der üblichen Matrizenmultiplikation.

**Satz 10.3.** Sei  $(G, \otimes)$  eine Gruppe, und seien  $x, y \in G$ .

- i) Es gibt genau ein  $g \in G$  mit  $x \otimes g = y$  und genau ein  $h \in G$  mit  $h \otimes x = y$ .
- ii) Es gilt  $(x \otimes y)^{-1} = y^{-1} \otimes x^{-1}$ .

**Definition 10.4** (Untergruppe). Sei  $(G, \otimes)$  Gruppe. Eine nichtleere Teilmenge  $U \subseteq G$  heißt *Untergruppe* von  $G$ , falls  $U$  mit der Verknüpfung  $\otimes$  die Gruppeneigenschaften erfüllt.  $(U, \otimes)$  ist also selbst wieder eine Gruppe. In diesem Falle schreibt man  $(U, \otimes) \leq (G, \otimes)$ , bzw. nur  $U \leq G$ .

**Satz 10.5** (Untergruppenkriterium (vergl. Satz 5.6)). Sei  $(G, \otimes)$  Gruppe und  $U \subseteq G$ .  $U$  ist genau dann Untergruppe von  $G$ , falls  $U \neq \emptyset$  und  $x \otimes y^{-1} \in U$  für alle  $x, y \in U$ .

*Beispiel 10.6.*

- Sei  $m \in \mathbb{N}$  und  $m\mathbb{Z} = \{m \cdot z : z \in \mathbb{Z}\}$ . Dann ist  $(m\mathbb{Z}, +) \leq (\mathbb{Z}, +)$ .
- Jeder lineare Teilraum (Untervektorraum) von  $\mathbb{R}^n$  ist eine Untergruppe von  $(\mathbb{R}^n, +)$ .
- $\text{SL}(n, \mathbb{R}) = \{A \in \mathbb{R}^{n \times n} : \det A = 1\}$  ist eine Untergruppe von  $\text{GL}(n, \mathbb{R})$ .

---

<sup>26</sup>Niels Abel, 1802–1829

### Notation 10.7.

i) Sei  $(G, \otimes)$  Gruppe mit neutralem Element  $e$ , und sei  $x \in G$ . Für  $k \in \mathbb{Z}$  versteht man unter

$$x^k = \begin{cases} \underbrace{x \otimes x \otimes \cdots \otimes x}_{k\text{-mal}} & : k \geq 1, \\ e & : k = 0, \\ \underbrace{x^{-1} \otimes x^{-1} \otimes \cdots \otimes x^{-1}}_{k\text{-mal}} & : k \leq -1. \end{cases}$$

ii) Für  $x \in G$  sei

$$\langle x \rangle = \{x^k : k \in \mathbb{Z}\}.$$

**Bemerkung 10.8.** Sei  $(G, \otimes)$  Gruppe und  $x \in G$ . Dann ist  $\langle x \rangle \leq G$ .

**Definition 10.9** (Zyklische Gruppe). Eine Gruppe  $G$  heißt *zyklisch*, falls ein  $x \in G$  existiert mit  $G = \langle x \rangle$ , d.h. es gibt ein Element  $x \in G$ , welches die Gruppe erzeugt.

*Beispiel 10.10.*

- $\mathbb{Z} = \langle 1 \rangle$ ,  $\mathbb{Z}_m = \langle [1]_m \rangle$ .
- $(\mathbb{Z}_5 \setminus \{[0]_5\}, \odot) = \langle [2]_5 \rangle$  (vergl. Satz 3.40)
- $\langle 2 \rangle = \{2^k : k \in \mathbb{Z}\} = \{\dots, \frac{1}{4}, \frac{1}{2}, 1, 2, 4, \dots\}$  ist eine zyklische Untergruppe von  $(\mathbb{Q}, \cdot)$ .

**Definition 10.11** (Ordnung einer Gruppe / eines Elements). Sei  $(G, \otimes)$  eine Gruppe. Die Anzahl der Elemente von  $G$ , bezeichnet mit  $|G|$ , heißt die *Ordnung der Gruppe* (unendlich möglich). Für  $x \in G$  heißt  $|\langle x \rangle|$  die *Ordnung von  $x$* .

*Beispiel 10.12.*

- $|\mathbb{Z}| = \infty$
- $|\mathbb{Z}_m| = m$ , und in  $(\mathbb{Z}_6, \oplus)$  ist  $|\langle [2]_6 \rangle| = 3$ .

**Bemerkung 10.13.** Sei  $(G, \otimes)$  Gruppe mit neutralem Element  $e_G$ . Sei  $|G|$  endlich und sei  $x \in G$ . Dann ist

$$|\langle x \rangle| = \min\{n \in \mathbb{N} : x^n = e_G\}.$$

**Definition 10.14** (Nebenklassen). Sei  $(G, \otimes)$  Gruppe, und sei  $U \leq G$  Untergruppe von  $G$ . Die Relation  $R$  auf  $G$  mit

$$xRy \Leftrightarrow x^{-1} \otimes y \in U \quad (\Leftrightarrow y \in x \otimes U)$$

ist eine Äquivalenzrelation (vergl. Def. 2.5). Hierbei ist  $x \otimes U = \{x \otimes u : u \in U\}$ . Die durch diese Äquivalenzrelation definierten Äquivalenzklassen  $[x]_R = x \otimes U$  heißen (*Links-*)Nebenklassen von  $U$  bzgl.  $G$ .

**Definition 10.15** (Index einer Untergruppe). Sei  $(G, \otimes)$  Gruppe, und sei  $U \leq G$  Untergruppe von  $G$ . Die Anzahl der Nebenklassen von  $U$  bzgl.  $G$  heißt *Index von  $U$  in  $G$*  und wird mit  $|G : U|$  bezeichnet.

**Satz 10.16** (Lagrange<sup>27</sup>). Sei  $G$  endliche Gruppe, und sei  $U \leq G$  Untergruppe von  $G$ . Dann ist

$$|G| = |U| \cdot |G : U|.$$

Insbesondere ist die Ordnung einer Untergruppe Teiler der Gruppenordnung.

---

<sup>27</sup>Joseph-Louis Lagrange, 1736–1813

**Satz 10.17** (Euler<sup>28</sup>). Sei  $G$  endliche Gruppe, und sei  $x \in G$ . Dann ist  $|x|$  ein Teiler von  $|G|$ .

**Korollar 10.18.** Sei  $G$  endliche Gruppe mit neutralem Element  $e$ , und sei  $x \in G$ . Dann ist  $x^{|G|} = e$ .

**Satz 10.19** (Kleiner Satz von Fermat<sup>29</sup>, siehe Satz 3.30). Sei  $a \in \mathbb{Z}$ , und sei  $p$  Primzahl mit  $\text{ggT}(p, a) = 1$ . Dann ist

$$a^{p-1} \equiv 1 \pmod{p}.$$

**Satz 10.20** (Chinesischer Restsatz). Seien  $m_1, \dots, m_n \in \mathbb{N}$  mit  $\text{ggT}(m_i, m_j) = 1$  für  $1 \leq i \neq j \leq n$ , und seien  $a_1, \dots, a_n \in \mathbb{Z}$ . Dann gibt es ein  $x \in \mathbb{Z}$  mit

$$x \equiv a_i \pmod{m_i}, \quad 1 \leq i \leq n.$$

$x$  ist modulo  $m = m_1 \cdot m_2 \cdot \dots \cdot m_n$  eindeutig bestimmt.

**Bemerkung 10.21** (Bestimmen einer Lösung).

- Für  $k = 1, \dots, n$  sei  $M_k = m/m_k$ .
- Dann ist  $\text{ggT}(M_k, m_k) = 1$  und es gibt ein  $N_k \in \mathbb{Z}$  mit  $N_k \cdot M_k \equiv 1 \pmod{m_k}$  (siehe Lemma 3.25).
- Sei  $x = \sum_{i=1}^n a_i M_i N_i$ .

*Beispiel 10.22.* Gesucht ist eine Zahl  $x \in \mathbb{Z}$ , die bei Division durch 3 den Rest 2 lässt, bei Division durch 5 den Rest 3, und bei Division durch 7 den Rest 2, d.h. gesucht ist  $x \in \mathbb{Z}$  mit  $x \equiv 2 \pmod{3}$ ,  $x \equiv 3 \pmod{5}$  und  $x \equiv 2 \pmod{7}$ .

- $m_1 = 3, a_1 = 2, m_2 = 5, a_2 = 3, m_3 = 7, a_3 = 2, m = 105$ .
- $M_1 = 35, M_2 = 21, M_3 = 15$ .
- $N_1 = 2, N_2 = 1, N_3 = 1$ .
- $x = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 233$ .
- Jede Zahl der Form  $233 + z \cdot 105, z \in \mathbb{Z}$ , ist Lösung; also ist die kleinste positive Lösung 23.

**Definition 10.23** (Ring). Sei  $R$  eine nichtleere Menge, und seien  $+, \cdot : R \times R \rightarrow R$  mit  $(x, y) \mapsto x + y$  (Addition), bzw.  $(x, y) \mapsto x \cdot y$  (Multiplikation) Abbildungen.  $(R, +, \cdot)$  heißt *Ring*, wenn die folgenden Bedingungen 1.–4. erfüllt sind:

1.  $(R, +)$  ist kommutative Gruppe. Das neutrale Element von  $(R, +)$  wird mit 0 bezeichnet.

2. Assoziativität bzgl. Multiplikation, d.h. für alle  $x, y, z \in R$  gilt

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

3. Bzgl. der Multiplikation existiert ein Einselement, bezeichnet mit 1, so dass gilt:  $1 \cdot x = x \cdot 1 = x$  für alle  $x \in R$ .

4. Distributivgesetze: Für alle  $x, y, z \in R$  gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z \text{ und } (x + y) \cdot z = x \cdot z + y \cdot z.$$

5.  $R$  heißt *kommutativer Ring*, wenn zusätzlich  $x \cdot y = y \cdot x$  für alle  $x, y \in R$  gilt.

---

<sup>28</sup>Leonhard Euler, 1707–1783

<sup>29</sup>Pierre de Fermat, 1601–1665

*Beispiel 10.24.*

- Jeder Körper (vergl. Def. 3.38) ist ein Ring.
- $(\mathbb{Z}, +, \cdot)$  ist ein Ring (aber kein Körper).
- $(\mathbb{Z}_m, \oplus, \odot)$  ist ein Ring (aber i.A. kein Körper, vergl. Satz 3.40).
- $\mathbb{Z}[i] = \{a+bi : a, b \in \mathbb{Z}\}$  ist ein Ring mit der gewöhnlichen Addition und Multiplikation von komplexen Zahlen. Er heißt Ring der *Gauß'schen ganzen Zahlen*.

**Definition 10.25** (Polynomring). Sei  $(R, +, \cdot)$  ein Ring, und seien  $a_0, a_1, \dots, a_n \in R$ . Die Abbildung  $f : R \rightarrow R$  mit  $f(x) = a_n \cdot x^n + a_{n-1} x^{n-1} + \dots + a_1 \cdot x + a_0$  heißt *Polynomfunktion* oder kurz *Polynom*.  $a_i$  heißen die Koeffizienten des Polynoms.

Der größte Index  $i$  mit  $a_i \neq 0$  heißt der *Grad* von  $f$  und wird mit  $\text{grad}(f)$  bezeichnet. Ist  $a_{\text{grad}(f)} = 1$ , dann heißt  $f$  *normiert*.

Die Menge aller Polynome mit Koeffizienten aus  $R$  wird mit  $R[x]$  bezeichnet. Mit den Verknüpfungen

$$(p \oplus q)(x) = p(x) + q(x) \text{ und } (p \odot q)(x) = p(x) \cdot q(x),$$

$p, q \in R[x]$  ist  $(R[x], \oplus, \odot)$  (oder auch nur  $R(x)$ ) ein Ring und heißt *Polynomring*. Statt  $\oplus, \odot$  schreibt man oft auch nur  $+, \cdot$ .

*Beispiel 10.26.*  $\mathbb{R}[x]$  ist der Ring aller Polynome mit Koeffizienten aus  $\mathbb{R}$ . Sei  $p(x) = 2x^2 - x + 3$  und  $q(x) = x^5 + 3x^3 - 4x^2 + 6x + 2$ . Dann ist

$$\begin{aligned} (p + q)(x) &= (2x^2 - x + 3) + (x^5 + 3x^3 - 4x^2 + 6x + 2) \\ &= x^5 + 3x^3 - 2x^2 + 5x + 5, \\ (p \cdot q)(x) &= (2x^2 - x + 3) \cdot (x^5 + 3x^3 - 4x^2 + 6x + 2) \\ &= 2x^7 - x^6 + 9x^5 - 11x^4 + 25x^3 - 14x^2 + 16x + 6. \end{aligned}$$

**Definition 10.27** (Ideale). Sei  $(R, +, \cdot)$  ein kommutativer Ring. Eine Menge  $I \subseteq R$  heißt *Ideal* von  $R$ , falls

1.  $(I, +)$  ist Untergruppe von  $(R, +)$ .
2. Für alle  $x \in R$  und für alle  $y \in I$  gilt:  $x \cdot y \in I$ .

*Beispiel 10.28.*

- $m\mathbb{Z}$  ist ein Ideal in  $(\mathbb{Z}, +, \cdot)$ .
- In jedem Ring  $R$  sind  $\{0\}$  und  $R$  Ideale.
- Die Menge aller Polynome in  $\mathbb{R}[x]$  mit Nullstelle 1 bildet ein Ideal  $I$ , welches sich beschreiben lässt als  $I = (x - 1)\mathbb{R}[x]$ .

**Bemerkung 10.29.** Sei  $(R, +, \cdot)$  Ring. Ist  $(R \setminus \{0\}, \cdot)$  eine kommutative Gruppe, dann heißt  $(R, +, \cdot)$  Körper (siehe Def. 3.38). Jeder Körper ist also ein Ring, aber nicht umgekehrt.

**Satz 10.30** (Polynomdivision). Sei  $\mathbb{K}$  Körper und  $\mathbb{K}[x]$  der zugehörige Polynomring. Für  $f, g \in \mathbb{K}[x]$ ,  $\text{grad}(g) \geq 1$ , gibt es  $q, r \in \mathbb{K}[x]$  mit

$$f = q \cdot g + r \text{ und } \text{grad}(r) < \text{grad}(g).$$

Man schreibt auch  $r = f \bmod g$  (vgl. Def. 3.22)

**Definition 10.31** (ggT von Polynomen). Sei  $\mathbb{K}$  Körper und  $\mathbb{K}[x]$  der zugehörige Polynomring.  $g \in \mathbb{K}[x]$  heisst Teiler von  $f \in \mathbb{K}[x]$ , falls es ein  $q \in \mathbb{K}[x]$  gibt mit  $f = q \cdot g$ , also  $f \bmod g = 0$ .

Seien  $f_1, f_2 \in \mathbb{K}[x]$ . Das *normierte* Polynom mit maximalem Grad, das Teiler von  $f_1$  und  $f_2$  ist, heisst *größter gemeinsamer Teiler* von  $f_1$  und  $f_2$ , und wird mit  $\text{ggT}(f_1, f_2)$  bezeichnet.

**Bemerkung 10.32.** Analog zum Euklidischen Algorithmus für Zahlen (vgl. Satz 3.23) lässt sich auch der ggT von Polynomen  $f_1, f_2$  bestimmen, bzw. eine Darstellung des ggT's mittels der Polynome  $f_1, f_2$  (s. Bem. 3.26).

**Satz 10.33.** Sei  $\mathbb{K}$  Körper und  $\mathbb{K}[x]$  der zugehörige Polynomring. Sei  $f(x) \in \mathbb{K}[x]$ , und sei  $a \in \mathbb{K}$ .  $f(x)$  lässt sich genau dann durch den Linearfaktor  $x - a \in \mathbb{K}[x]$  teilen, d.h.,  $f(x) = q(x) \cdot (x - a)$  für ein  $q(x) \in \mathbb{K}[x]$ , falls  $a$  Nullstelle von  $f(x)$  ist, d.h.,  $f(a) = 0$ .

**Definition 10.34** (Kongruenz von Polynomen, Restklassenring). Sei  $K$  Körper und  $\mathbb{K}[x]$  der zugehörige Polynomring.

i) Seien  $f, g, m \in \mathbb{K}[x]$ . Dann heissen  $f$  und  $g$  *kongruent modulo m*, falls  $f$  und  $g$  den gleichen Rest bei Division durch  $m$  lassen, d.h.  $m$  ist ein Teiler von  $f - g$ . Man schreibt  $f \equiv g \pmod{m}$  (vgl. Def. 3.28).

ii) Sei  $m \in \mathbb{K}[x]$  normiert. Die Menge

$$\mathbb{K}[x]_m = \{f \bmod m : f \in \mathbb{K}[x]\}$$

heisst *Restklassenring modulo m*.

**Bemerkung 10.35.**  $\mathbb{K}[x]_m$  ist in der Tat ein Ring mittels der Operationen  $\tilde{\oplus}, \tilde{\odot}$  definiert als

$$(f \bmod m) \tilde{\oplus} (g \bmod m) := (f + g) \bmod m \text{ und } (f \bmod m) \tilde{\odot} (g \bmod m) := (f \cdot g) \bmod m.$$

Vergleiche Notation 3.31.

**Definition 10.36** (Irreduzible Polynome). Ein Polynom  $p \in \mathbb{K}[x]$  heisst *irreduzibel*, falls  $p$  von keinem Polynom  $q$  mit  $0 < \text{grad}(q) < \text{grad}(p)$  geteilt wird.

**Satz 10.37.** Sei  $K$  Körper und  $\mathbb{K}[x]$  der zugehörige Polynomring. Sei  $f \in \mathbb{K}[x]$  ein normiertes Polynom. Dann gibt es eindeutig bestimmte normierte irreduzible Polynome  $q_i \in \mathbb{K}[x]$ ,  $1 \leq i \leq m$ , so dass  $f = \prod_{i=1}^m q_i$  (vgl. Satz 3.16).

**Satz 10.38.** Sei  $K$  Körper und  $\mathbb{K}[x]$  der zugehörige Polynomring. Für  $m \in \mathbb{K}[x]$  gilt:  $\mathbb{K}[x]_m$  ist genau dann ein Körper, falls  $m$  irreduzibel ist (vgl. Satz 3.40).

## 11 Folgen und Reihen

**Definition 11.1** (Folgen). Eine (reelle) *Folge* ist eine Abbildung  $a : \mathbb{N}_0 \rightarrow \mathbb{R}$ ,  $n \mapsto a_n$ , auch geschrieben als  $a_0, a_1, a_2, \dots$ . Die reellen Zahlen  $a_n$  heißen die *Glieder* der Folge. Die Folge wird auch mit  $(a_n)_{n \in \mathbb{N}_0}$  bezeichnet.

*Beispiel 11.2.*

1.  $a_n = \frac{1}{n+1}$  ist die Folge  $1, \frac{1}{2}, \frac{1}{3}, \dots$
2.  $a_n = n^2$  ist die Folge  $0, 1, 4, 9, 16, \dots$
3.  $a_n = \frac{n}{n+1}$  ist die Folge  $0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots$

**Bemerkung 11.3.** Eine Folge muss nicht mit dem Index 0 beginnen; sie kann etwa auch mit 1 oder jeder beliebigen ganzen Zahl  $k$  beginnen. Man betrachtet dann Funktionen  $a : \mathbb{N} \rightarrow \mathbb{R}$  oder  $a : \mathbb{Z}_{\geq k} \rightarrow \mathbb{R}$  und schreibt  $(a_n)_{n \in \mathbb{N}}$  oder  $(a_n)_{n \in \mathbb{Z}_{\geq k}}$ .

Auch kann man als Folgenglieder komplexe Zahlen betrachten. In diesem Fall betrachtet man Funktionen  $a : \mathbb{N}_0 \rightarrow \mathbb{C}$ .

*Beispiel 11.4.*

1. *Rekursive Folgen:*

- (a) Sei  $h_0 = 2$  und für  $n \geq 1$  sei  $h_n = \frac{1}{2}(h_{n-1} + \frac{2}{h_{n-1}})$ . Dies ergibt die Folge  $2, 1.5, 1.416666\dots, 1.414215\dots, 1.414213\dots, \dots$
- (b) Sei  $f_0 = 0, f_1 = 1$  und für  $n \geq 2$  sei  $f_n = f_{n-1} + f_{n-2}$ . Dies ergibt die *Fibonacci-Folge*  $0, 1, 2, 3, 5, 8, \dots$  (vergl. Beispiel 3.7).
- (c) Sei  $a_0 = 1$  und für  $n \geq 1$  sei  $a_n = n \cdot a_{n-1}$ . Dann ist  $a_n = n!$  (vergl. Bemerkung 3.5).

2. *Alternierende Folgen*, d. h. die Folgenglieder haben abwechselndes Vorzeichen:

- (a)  $a_n = (-1)^{n+1}$  ist die Folge  $-1, 1, -1, 1, -1, 1, -1, 1, \dots$
- (b)  $a_n = (-1/2)^n$  ist die Folge  $1, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \dots$

**Definition 11.5** (Monotone/beschränkte Folgen).

- i) Eine (reelle) Folge  $(a_n)_{n \in \mathbb{N}_0}$  heißt *monoton steigend* (bzw. *monoton fallend*) falls  $a_n \leq a_{n+1}$  (bzw.  $a_n \geq a_{n+1}$ ) für alle  $n \in \mathbb{N}_0$ .
- ii) Eine (reelle) Folge  $(a_n)_{n \in \mathbb{N}_0}$  heißt nach *oben beschränkt*, wenn es ein  $C \in \mathbb{R}$  gibt, so dass  $a_n \leq C$  für alle  $n \in \mathbb{N}_0$ . Entsprechend heißt eine Folge nach *unten beschränkt*, wenn es ein  $c \in \mathbb{R}$  gibt, so dass  $a_n \geq c$  für alle  $n \in \mathbb{N}_0$ .
- iii) Eine (reelle)  $(a_n)_{n \in \mathbb{N}_0}$  heißt *beschränkt*, wenn sie nach oben und unten beschränkt ist, d.h. wenn es ein  $M \in \mathbb{R}$  gibt, mit  $|a_n| \leq M$  für alle  $n \in \mathbb{N}_0$ .

**Definition 11.6** (Konvergenz von Folgen). Eine reelle Folge  $(a_n)_{n \in \mathbb{N}_0}$  heißt *konvergent* gegen eine Zahl  $a \in \mathbb{R}$ , wenn es für jede (noch so kleine) positive reelle Zahl  $\epsilon > 0$  einen Folgenindex  $n_0$  gibt, so dass alle Folgenglieder mit  $n \geq n_0$  in dem Intervall  $(a - \epsilon, a + \epsilon)$  liegen, d.h.

$$|a_n - a| < \epsilon \text{ für alle } n \geq n_0.$$

In diesem Fall schreibt man

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \rightarrow a \text{ für } n \rightarrow \infty.$$

$a$  heißt der *Grenzwert* der Folge, und man sagt „Die Folge  $(a_n)_{n \in \mathbb{N}_0}$  konvergiert gegen  $a$ “.

*Beispiel 11.7.*

Sei  $a_n = \frac{n}{n+1}$ . Offensichtlich nähern sich die Folgenglieder immer mehr der Zahl 1. Um zu zeigen, dass 1 (tatsächlich) der Grenzwert dieser Folge ist, muß für jedes  $\epsilon > 0$  ein  $n_0$  existieren, so dass

$$|a_n - 1| = \left| \frac{n}{n+1} - 1 \right| < \epsilon \text{ für alle } n \geq n_0.$$

Wegen

$$\left| \frac{n}{n+1} - 1 \right| = \left| \frac{-1}{n+1} \right| = \frac{1}{n+1}$$

kann man in diesem Fall für  $n_0$  irgendeine positive ganze Zahl  $> 1/\epsilon - 1$  wählen.

**Bemerkung 11.8.**

- i) Eine Folge  $(a_n)_{n \in \mathbb{N}_0}$  konvergiert gegen eine Zahl  $a$  genau dann, wenn für jedes  $\epsilon > 0$  nur endlich viele Folgenglieder außerhalb von  $(a - \epsilon, a + \epsilon)$  liegen.
- ii) Der Index  $n_0$  hängt i. A. immer von  $\epsilon$  ab; je kleiner  $\epsilon$  desto größer  $n_0$ .

**Definition 11.9** (Teilfolge). Sei  $(a_n)_{n \in \mathbb{N}_0}$  eine Folge und sei  $M \subset \mathbb{N}_0$  mit  $|M| = \infty$ . Dann heißt die Folge  $(a_n)_{n \in M}$  Teilfolge von  $(a_n)_{n \in \mathbb{N}_0}$ . Sie besteht also aus allen Folgengliedern  $a_n$ , deren Index in  $M$  liegt.

*Beispiel 11.10.* Sei  $a_n = (-1)^n$  und sei  $M \subset \mathbb{N}_0$  die Menge aller geraden Zahlen. Dann ist  $(a_n)_{n \in M}$  die Folge 1, 1, 1, 1, 1, 1, 1, 1, ... Man kann sie auch beschreiben als  $(a_{2n})_{n \in \mathbb{N}_0}$ .

**Bemerkung 11.11.** Ein Folge komplexer Zahlen  $(a_n)_{n \in \mathbb{N}_0}$  mit  $a_n = x_n + iy_n$  heißt konvergent, falls die Folgen bestehend aus den Realteilen  $(x_n)_{n \in \mathbb{N}_0}$  und den Imaginärteilen  $(y_n)_{n \in \mathbb{N}_0}$  konvergieren, d.h. es gibt  $x, y \in \mathbb{R}$  mit  $x_n \rightarrow x$  und  $y_n \rightarrow y$  für  $n \rightarrow \infty$ . Der Grenzwert der Folge  $(a_n)_{n \in \mathbb{N}_0}$  ist dann  $x + iy$ .

**Satz 11.12.**

- i) Der Grenzwert einer Folge ist eindeutig bestimmt. Konvergiert eine Folge, so konvergiert auch jede Teilfolge gegen den Grenzwert.
- ii) Jede konvergente Folge ist beschränkt, bzw. jede unbeschränkte Folge ist nicht konvergent.

**Definition 11.13** (Divergente Folgen).

- i) Eine Folge, die nicht konvergiert, heißt divergent.
- ii) Eine Folge  $(a_n)_{n \in \mathbb{N}_0}$  heißt bestimmt divergent gegen  $\infty$ , falls  $\lim_{n \rightarrow \infty} 1/a_n = 0$  und ab einem Index  $n_0$  gilt  $a_n > 0$  für  $n \geq n_0$ . In diesem Fall schreibt man  $\lim_{n \rightarrow \infty} a_n = \infty$ . Man sagt auch „ $(a_n)_{n \in \mathbb{N}_0}$  konvergiert gegen  $\infty$ “.
- iii) Analog heißt eine Folge  $(a_n)_{n \in \mathbb{N}_0}$  bestimmt divergent gegen  $-\infty$ , falls  $\lim_{n \rightarrow \infty} 1/a_n = 0$  und ab einem Index  $n_0$  gilt  $a_n < 0$  für  $n \geq n_0$ . In diesem Fall schreibt man  $\lim_{n \rightarrow \infty} a_n = -\infty$ . Man sagt dann auch „ $(a_n)_{n \in \mathbb{N}_0}$  konvergiert gegen  $-\infty$ “.

*Beispiel 11.14.*  $a_n = (-1)^n$  oder  $b_n = 2^n$  sind divergente Folgen, und  $(b_n)_{n \in \mathbb{N}_0}$  ist bestimmt divergent gegen  $\infty$ .

**Satz 11.15** (Rechenregeln für konvergente Folgen). Seien  $(a_n)_{n \in \mathbb{N}_0}$ ,  $(b_n)_{n \in \mathbb{N}_0}$  konvergente Folgen mit den Grenzwerten  $a$ ,  $b$ , und sei  $\alpha \in \mathbb{C}$ . Dann sind auch die Folgen  $(\alpha a_n)_{n \in \mathbb{N}_0}$ ,  $(a_n \pm b_n)_{n \in \mathbb{N}_0}$ ,  $(a_n \cdot b_n)_{n \in \mathbb{N}_0}$  und  $(a_n/b_n)_{n \in \mathbb{N}_0, b_n \neq 0}$  (falls  $b \neq 0$ ) konvergent mit den Grenzwerten:

$$\begin{aligned} \lim_{n \rightarrow \infty} \alpha a_n &= \alpha a, & \lim_{n \rightarrow \infty} (a_n \pm b_n) &= a \pm b, \\ \lim_{n \rightarrow \infty} (a_n \cdot b_n) &= a \cdot b, & \lim_{n \rightarrow \infty} \frac{a_n}{b_n} &= \frac{a}{b}. \end{aligned}$$

**Satz 11.16.** Sei  $(a_n)_{n \in \mathbb{N}_0}$  die rationale Folge

$$a_n = \frac{\alpha_k n^k + \alpha_{k-1} n^{k-1} + \cdots + \alpha_1 n + \alpha_0}{\beta_m n^m + \beta_{m-1} n^{m-1} + \cdots + \beta_1 n + \beta_0},$$

wobei  $k, m \in \mathbb{N}_0$  und  $\alpha_i, \beta_i \in \mathbb{R}$  mit  $\alpha_k \neq 0, \beta_m \neq 0$ . Dann gilt

$$\lim_{n \rightarrow \infty} a_n = \begin{cases} 0, & \text{für } k < m, \\ a_k/b_m, & \text{für } k = m, \\ \infty, & \text{für } k > m \text{ und } a_k/b_m > 0, \\ -\infty, & \text{für } k > m \text{ und } a_k/b_m < 0. \end{cases}$$

**Satz 11.17.** Jede beschränkte und monoton wachsende (oder fallende) Folge konvergiert.

**Bemerkung 11.18.** Sei  $x > 0$ . Die Heron'sche Folge

$$h_n = \frac{1}{2} \left( h_{n-1} + \frac{x}{h_{n-1}} \right)$$

konvergiert für jeden „Startwert“  $h_0 > 0$  gegen  $\sqrt{x}$ .

**Bemerkung 11.19.** Eine Folge  $(a_n)_{n \in \mathbb{N}_0}$  mit  $\lim_{n \rightarrow \infty} a_n = 0$  heißt Nullfolge. Für  $q \in \mathbb{R}$  mit  $|q| < 1$  ist  $a_n = q^n$  eine Nullfolge.

**Definition 11.20** (Reihe, Partialsummen). Sei  $(a_n)_{n \in \mathbb{N}_0}$  eine Folge reeller oder komplexer Zahlen. Man nennt den (formalen) Ausdruck

$$\sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + \cdots$$

eine (*unendliche*) Reihe. Die Folge  $(s_n)_{n \in \mathbb{N}_0}$  mit den Folgegliedern

$$s_n = \sum_{k=0}^n a_k = a_0 + a_1 + a_2 + \cdots + a_n$$

heißt *Folge der Partialsummen* (oder auch *Teilsummen*) der Reihe.

**Definition 11.21** (Konvergenz von Reihen). Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe und  $(s_n)_{n \in \mathbb{N}_0}$  die Folge ihrer Partialsummen. Ist  $(s_n)_{n \in \mathbb{N}_0}$  konvergent mit Grenzwert  $s$ , dann heißt die Reihe konvergent mit Grenzwert  $s$ , und man schreibt

$$s = \sum_{k=0}^{\infty} a_k.$$

Andernfalls heißt die Reihe *divergent*.

*Beispiel 11.22.*

1. Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

heißt *harmonische Reihe* und ist *divergent*!

2. Die Reihe

$$\sum_{k=0}^{\infty} \left( \frac{1}{2} \right)^k$$

ist konvergent mit Grenzwert 2.

3. Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$

ist konvergent mit Grenzwert  $\pi^2/6$ .

**Satz 11.23.** Ist  $\sum_{k=0}^{\infty} a_k$  eine konvergente Reihe, dann ist  $(a_n)_{n \in \mathbb{N}_0}$  eine Nullfolge, d.h.  $\lim_{n \rightarrow \infty} a_n = 0$ . Die Umkehrung gilt im Allgemeinen nicht, siehe z.B. harmonische Reihe.

**Satz 11.24** (Geometrische Reihe). Sei  $q \in \mathbb{R}$  oder  $\mathbb{C}$ . Eine Reihe der Form

$$\sum_{k=0}^{\infty} q^k$$

heißt geometrische Reihe, wobei der Index  $k$  nicht unbedingt bei 0 beginnen muss. Sei  $(s_n)_{n \in \mathbb{N}_0}$  die Folge ihrer Partialsummen. Dann gilt für  $q \neq 1$

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

Die geometrische Reihe ist genau dann konvergent, wenn  $|q| < 1$ , und in diesem Fall ist

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

**Satz 11.25** (Rechenregeln für konvergente Reihen). Seien  $\sum_{k=0}^{\infty} a_k$ ,  $\sum_{k=0}^{\infty} b_k$  konvergente Reihen, und sei  $\alpha \in \mathbb{C}$ . Dann sind auch die Reihen  $\sum_{k=0}^{\infty} (\alpha a_k)$  und  $\sum_{k=0}^{\infty} (a_k \pm b_k)$  konvergent, und es gilt

$$\begin{aligned} \sum_{k=0}^{\infty} (\alpha a_k) &= \alpha \sum_{k=0}^{\infty} a_k, \\ \sum_{k=0}^{\infty} (a_k \pm b_k) &= \sum_{k=0}^{\infty} a_k \pm \sum_{k=0}^{\infty} b_k. \end{aligned}$$

**Definition 11.26.** Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe. Ist die Reihe  $\sum_{k=0}^{\infty} |a_k|$  konvergent, dann heißt  $\sum_{k=0}^{\infty} a_k$  absolut konvergent.

**Bemerkung 11.27.** Jede absolut konvergente Reihe ist konvergent.

**Bemerkung 11.28.** Seien  $\sum_{k=0}^{\infty} a_k$ ,  $\sum_{k=0}^{\infty} b_k$  absolut konvergente Reihen. Dann ist auch ihr Cauchy-Produkt

$$\sum_{k=0}^{\infty} c_k \text{ mit } c_k = \sum_{i=0}^k a_i b_{k-i}$$

(absolut) konvergent, und es gilt

$$\sum_{k=0}^{\infty} c_k = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{k=0}^{\infty} b_k \right).$$

**Definition 11.29** (Alternierende Reihe). Ist  $(a_k)_{k \in \mathbb{N}_0}$  eine alternierende Folge, d.h. die Folgenglieder sind abwechselnd nicht-negativ und nicht-positiv, dann heißt  $\sum_{k=0}^{\infty} a_k$  einen alternierende Reihe.

**Satz 11.30** (Konvergenzkriterium von Leibniz<sup>30</sup>). Sei  $(a_k)_{k \in \mathbb{N}_0}$  eine monoton fallende Nullfolge nichtnegativer Zahlen. Dann konvergiert die alternierende Reihe

$$\sum_{k=0}^{\infty} (-1)^k a_k.$$

Beispiel 11.31.

$$\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$$

ist konvergent, aber nicht absolut konvergent.

**Satz 11.32** (Majorantenkriterium). Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe.

- i) Wenn es eine konvergente Reihe  $\sum_{k=0}^{\infty} b_k$  und eine  $k_0 \in \mathbb{N}_0$  gibt mit

$$|a_k| \leq b_k \text{ für alle } k \geq k_0,$$

dann ist die Reihe  $\sum_{k=0}^{\infty} a_k$  (absolut) konvergent.  $\sum_{k=0}^{\infty} b_k$  heißt Majorante für  $\sum_{k=0}^{\infty} a_k$ .

- ii) Wenn es eine divergente Reihe  $\sum_{k=0}^{\infty} b_k$  und eine  $k_0 \in \mathbb{N}_0$  gibt mit

$$|a_k| \geq b_k \geq 0 \text{ für alle } k \geq k_0,$$

dann ist die Reihe  $\sum_{k=0}^{\infty} a_k$  divergent. Man nennt dann die Reihe  $\sum_{k=0}^{\infty} b_k$  eine (divergente) Minorante für  $\sum_{k=0}^{\infty} a_k$ .

**Satz 11.33** (Quotientenkriterium). Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe.

- i) Wenn es eine Zahl  $q < 1$  und ein  $k_0 \in \mathbb{N}_0$  gibt, so dass

$$\left| \frac{a_{k+1}}{a_k} \right| \leq q \text{ für alle } k \geq k_0,$$

dann ist die Reihe  $\sum_{k=0}^{\infty} a_k$  (absolut) konvergent.

- ii) Wenn es ein  $k_0 \in \mathbb{N}_0$  gibt, so dass

$$\left| \frac{a_{k+1}}{a_k} \right| \geq 1 \text{ für alle } k \geq k_0,$$

dann ist die Reihe  $\sum_{k=0}^{\infty} a_k$  divergent.

**Bemerkung 11.34.** Sei  $\sum_{k=0}^{\infty} a_k$  eine Reihe, und sei

$$b_k = \left| \frac{a_{k+1}}{a_k} \right|$$

eine konvergente Folge mit Grenzwert  $b$ . Ist  $b < 1$ , dann ist  $\sum_{k=0}^{\infty} a_k$  (absolut) konvergent. Ist  $b > 1$ , dann ist  $\sum_{k=0}^{\infty} a_k$  divergent. Im Falle  $b = 1$  ist keine Aussage möglich, die Reihe kann sowohl konvergent als auch divergent sein.

---

<sup>30</sup>Gottfried Wilhelm Freiherr von Leibniz, 1646-1716

## 12 Stetigkeit von Funktionen

**Definition 12.1** ((Strenge) Montonie von Funktionen). Sei  $D \subseteq \mathbb{R}$  und sei  $f : D \rightarrow \mathbb{R}$  eine Funktion.

1.  $f$  heißt *streng monoton wachsend*, falls  $f(x) > f(y)$  für alle  $x, y \in D$  mit  $x > y$ .
2.  $f$  heißt *streng monoton fallend*, falls  $f(x) < f(y)$  für alle  $x, y \in D$  mit  $x > y$ .
3. Gilt „lediglich“  $f(x) \geq f(y)$ , bzw.  $f(x) \leq f(y)$ , dann heißt die Funktion „nur“ *monoton wachsend*, bzw. *monoton fallend*.

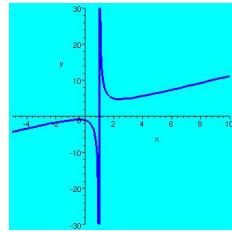
**Definition 12.2** (Beschränktheit von Funktionen). Eine Funktion  $f : D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}$  heißt *beschränkt*, falls es ein  $M > 0$  gibt mit  $|f(x)| \leq M$  für alle  $x \in D$ .

**Definition 12.3** (Rationale Funktionen). Seien  $p, q \in \mathbb{R}[x]$  Polynome. Dann heißt die Funktion  $f(x) = p(x)/q(x)$  *rationale Funktion*. Sie ist überall dort definiert, wo das Nennerpolynom  $q(x) \neq 0$  ist.

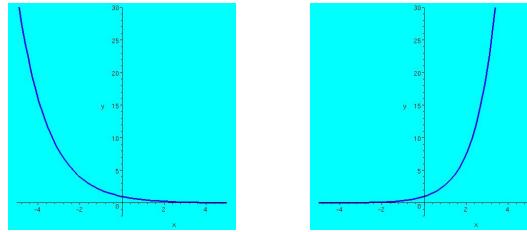
*Beispiel 12.4.* Sei

$$f(x) = \frac{x^3 + x^2 + x + 1}{x^2 - 1}.$$

Sie ist definiert auf der Menge  $\mathbb{R} \setminus \{-1, 1\}$ .



**Definition 12.5** (Exponentialfunktion). Sei  $a > 0$ . Eine Funktion der Form  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = a^x$  heißt *Exponentialfunktion zur Basis a*. Ist die Basis die *Eulersche Zahl e* = 2,7182818284..., dann heißt die Funktion einfach *Exponentialfunktion* (oder auch *e-Funktion*).



Basis  $a = 1/2$

Basis e

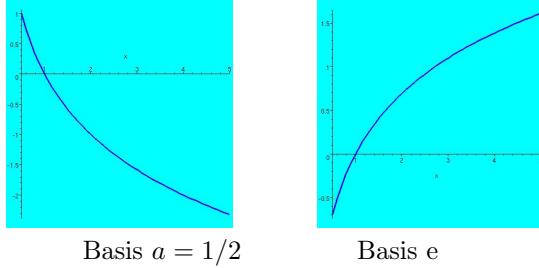
**Satz 12.6** (Eigenschaften der Exponentialfunktion). Sei  $f(x) = a^x$  mit  $a > 0$ . Dann gilt:

- i)  $f(0) = 1$ .
- ii)  $f(x) > 0$  für alle  $x \in \mathbb{R}$ .
- iii) Ist  $a < 1$ , dann ist  $f(x)$  eine streng monoton fallende Funktion. Ist  $a > 1$ , dann ist  $f(x)$  eine streng monoton wachsende Funktion. Die Exponentialfunktion ist unbeschränkt.
- iv)  $f(x+y) = f(x) \cdot f(y)$  für alle  $x, y \in \mathbb{R}$ .

**Definition 12.7** (Logarithmusfunktion). Sei  $a > 0, a \neq 1$ . Die Umkehrfunktion der Exponentialfunktion  $f : \mathbb{R} \rightarrow (0, \infty), f(x) = a^x$ , heißt *Logarithmusfunktion zur Basis a*, und wird mit

$$\log_a : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \log_a(x),$$

bezeichnet. Ist die Basis e, dann spricht man von dem *natürlichen Logarithmus*, der mit  $\ln(x)$  bezeichnet wird.



**Satz 12.8** (Eigenschaften der Logarithmusfunktion). Sei  $a > 0, a \neq 1$ .

- i)  $\log_a(1) = 0$  und  $\log_a(a) = 1$ .
- ii) Ist  $a < 1$ , dann ist  $\log_a(x)$  eine streng monoton fallende Funktion. Ist  $a > 1$ , dann ist  $\log_a(x)$  eine streng monoton wachsende Funktion. Die Logarithmusfunktion ist unbeschränkt.
- iii)  $\log_a(x \cdot y) = \log_a(x) + \log_a(y)$  für  $x, y \in (0, \infty)$ .
- iv)  $\log_a(x^b) = b \cdot \log_a(x)$ ,  $b \in \mathbb{R}$ .

**Bemerkung 12.9.** Sei  $a > 0, a \neq 1$ . Dann gilt:

- i)  $\log_a(a^x) = x$ .
- ii)  $a^x = b^{x \cdot \log_b(a)}$ .
- iii)  $\log_a(\frac{x}{y}) = \log_a(x) - \log_a(y)$ .
- iv)  $\log_a(x) = \log_b(x) / \log_b(a)$ .

Weiterhin ist  $\ln(x) \leq x - 1$  für alle  $x \in (0, \infty)$ .

**Notation 12.10** (Intervalle). Seien  $a, b \in \mathbb{R}, a \leq b$ .

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

heißt *abgeschlossenes Intervall*, und

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}$$

heißt *offenes Intervall*. Analog definiert man *halboffene* (bzw. *halbabgeschlossene*) Intervalle, z.B.

$$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}.$$

Unter  $[a, \infty)$  (bzw.  $(a, \infty)$ ) versteht man alle Zahlen  $\geq a$  (bzw.  $> a$ ).

**Definition 12.11** ((Un)gerade Funktionen). Sei  $D \subseteq \mathbb{R}$  und  $f : D \rightarrow \mathbb{R}$  eine Funktion.

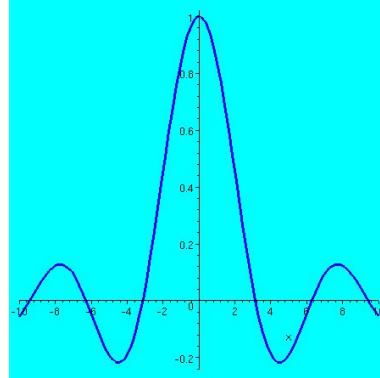
1.  $f$  heißt *gerade*, falls  $f(-x) = f(x)$  für alle  $x, -x \in D$ .
2.  $f$  heißt *ungerade*, falls  $f(-x) = -f(x)$  für alle  $x, -x \in D$ .

**Definition 12.12** (Periodische Funktionen). Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine Funktion.  $f$  heißt *periodisch* mit Periode  $p > 0$ , falls für alle  $x \in \mathbb{R}$  gilt

$$f(p + x) = f(x).$$

*Beispiel 12.13.*  $\sin, \cos : \mathbb{R} \rightarrow \mathbb{R}$  sind periodische Funktionen mit Periode  $2\pi$ . Desweiteren ist  $\sin$  ungerade, und  $\cos$  ist eine gerade Funktion.

*Beispiel 12.14.* Sei  $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  mit  $f(x) = \frac{\sin(x)}{x}$ .



**Definition 12.15** (Grenzwert einer Funktion). Sei  $D \subseteq \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion, und sei  $x^* \in \mathbb{R}$ . Wenn es ein  $y^* \in \mathbb{R}$  gibt, so dass für jede Folge  $(x_n)_{n \in \mathbb{N}_0}$  mit  $x_n \in D \setminus \{x^*\}$  und

$$\lim_{n \rightarrow \infty} x_n = x^*$$

gilt, dass die Folge der Funktionswerte  $(f(x_n))_{n \in \mathbb{N}_0}$  gegen  $y^*$  konvergiert, also

$$\lim_{n \rightarrow \infty} f(x_n) = y^*,$$

dann heißt  $y^*$  der Grenzwert von  $f$  für  $x$  gegen  $x^*$ . Man schreibt dafür

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

Hierbei ist auch  $y^* = \pm\infty$  erlaubt (siehe (11.13)) und man spricht dann von bestimmter Divergenz.

*Beispiel 12.16.*

i)

$$\lim_{x \rightarrow 0} \frac{1}{|x|} = \infty.$$

ii)

$$\lim_{x \rightarrow 5} 2x + 1 = 11.$$

iii) Sei  $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  mit  $f(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$ . Dann existiert  $\lim_{x \rightarrow 0} f(x)$  nicht, d.h.  $f$  hat keinen Grenzwert für  $x$  gegen 0.

**Definition 12.17** (Rechts-Linksseitiger Grenzwert). Sei  $D \subseteq \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion, und sei  $x^* \in \mathbb{R}$ . Wenn es ein  $y^* \in \mathbb{R}$  gibt, so dass für jede Folge  $(x_n)_{n \in \mathbb{N}_0}$  mit  $x_n \in D \setminus \{x^*\}$ ,  $x_n > x^*$  (!), und

$$\lim_{n \rightarrow \infty} x_n = x^*$$

gilt, dass

$$\lim_{n \rightarrow \infty} f(x_n) = y^*,$$

dann heißt  $y^*$  der rechtsseitige Grenzwert von  $f$  für  $x$  gegen  $x^*$ . Man schreibt dafür

$$\lim_{x \rightarrow x^*+} f(x) = y^*.$$

Analog definiert man den linkssseitigen Grenzwert von  $f$  für  $x$  gegen  $x^*$ , indem man nur Folgen  $(x_n)_{n \in \mathbb{N}_0}$  mit  $x_n \in D \setminus \{x^*\}$ ,  $x_n < x^*$  (!), und  $\lim_{n \rightarrow \infty} x_n = x^*$  betrachtet. In diesem Falle schreibt man

$$\lim_{x \rightarrow x^*-} f(x) = y^*.$$

*Beispiel 12.18.* Sei  $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  mit  $f(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$ . Dann gilt

$$\lim_{x \rightarrow 0^+} f(x) = 1 \text{ und } \lim_{x \rightarrow 0^-} f(x) = -1.$$

**Bemerkung 12.19.** Es gilt

$$\lim_{x \rightarrow x^*} f(x) = y^* \Leftrightarrow \lim_{x \rightarrow x^*+} f(x) = \lim_{x \rightarrow x^*-} f(x) = y^*.$$

**Satz 12.20** (Rechenregeln für Grenzwerte). Seien  $f, g$  Funktionen mit  $\lim_{x \rightarrow x^*} f(x) = \hat{y}$  und  $\lim_{x \rightarrow x^*} g(x) = \tilde{y}$ . Dann gilt

i)  $\lim_{x \rightarrow x^*} \alpha f(x) = \alpha \hat{y}$  für  $\alpha \in \mathbb{R}$ .

ii)  $\lim_{x \rightarrow x^*} (f(x) \pm g(x)) = \hat{y} \pm \tilde{y}$ .

iii)  $\lim_{x \rightarrow x^*} (f(x) \cdot g(x)) = \hat{y} \cdot \tilde{y}$ .

iv)  $\lim_{x \rightarrow x^*} \frac{f(x)}{g(x)} = \frac{\hat{y}}{\tilde{y}}$  falls  $\tilde{y} \neq 0$ .

**Definition 12.21** (Asymptotisches Verhalten). Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine Funktion. Unter dem asymptotischen Verhalten von  $f$  für  $x \rightarrow \infty$  versteht man den Grenzwert (falls existent und  $\pm\infty$  sind zugelassen)

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

Analog definiert man das asymptotischen Verhalten von  $f$  für  $x \rightarrow -\infty$  als

$$\lim_{x \rightarrow -\infty} f(x).$$

*Beispiel 12.22.*

i)

$$\lim_{x \rightarrow \infty} \frac{1}{x} = 0 \text{ und } \lim_{x \rightarrow -\infty} \frac{1}{x} = 0$$

ii)

$$\lim_{x \rightarrow \infty} x^3 = \infty \text{ und } \lim_{x \rightarrow -\infty} x^3 = -\infty.$$

*Beispiel 12.23.*

i)  $\lim_{x \rightarrow \infty} x^a = \infty$  für  $a > 0$ .

- ii)  $\lim_{x \rightarrow \infty} x^a = 0$  für  $a < 0$ .
- iii)  $\lim_{x \rightarrow -\infty} x^n = (-1)^n \infty$  für  $n \in \mathbb{N}$ .
- iv) Für  $a > 0$  ist  $\lim_{x \rightarrow \infty} e^{ax} = \infty$  und  $\lim_{x \rightarrow -\infty} e^{ax} = 0$ .
- v)  $\lim_{x \rightarrow \infty} \ln(x) = \infty$  und  $\lim_{x \rightarrow 0} \ln(x) = -\infty$ .

**Definition 12.24** (Stetigkeit). Sei  $D \subseteq \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion, und sei  $x^* \in D$ .  $f$  heißt *stetig im Punkt  $x^*$* , falls

$$\lim_{x \rightarrow x^*} f(x) = f(x^*),$$

d.h. der Grenzwert  $\lim_{x \rightarrow x^*} f(x)$  existiert und ist gleich  $f(x^*)$ . Dies bedeutet, für jede Folge  $(x_n)_{n \in \mathbb{N}_0}$ ,  $x_n \in D \setminus \{x^*\}$ , mit  $\lim_{n \rightarrow \infty} x_n = x^*$  gilt

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

Ist die Funktion  $f$  in jedem Punkt ihres Definitionsbereiches  $D$  stetig, dann heißt  $f$  *stetig*.

**Bemerkung 12.25.** *Stetigkeit von  $f : D \rightarrow \mathbb{R}$  in  $x^*$  bedeutet also, dass für  $x$  „in der Nähe von“  $x^*$  auch  $f(x)$  „nahe“ an  $f(x^*)$  ist. Genauer: Zu jedem  $\delta > 0$  existiert ein  $\epsilon > 0$  mit*

$$|f(x) - f(x^*)| < \delta \text{ für alle } x \in D \text{ mit } |x - x^*| < \epsilon.$$

*Beispiel 12.26.*

- i) Jede konstante Funktion ist stetig.
- ii)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = x$  ist stetig.
- iii)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = \lfloor x \rfloor$  ist nicht stetig in  $z \in \mathbb{Z}$ . Hierbei ist  $\lfloor x \rfloor$  die größte ganze Zahl kleiner oder gleich  $x$ .
- iv)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = |x|$  ist stetig.
- v)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0 \end{cases}$  ist nicht stetig im Punkt 0.

**Bemerkung 12.27.** Die Exponentialfunktionen, die Logarithmusfunktionen,  $\sin(x)$ ,  $\cos(x)$ , oder auch  $\sqrt{x}$  sind stetig.

**Satz 12.28.** Seien  $f(x), g(x)$  Funktionen, die in  $x^*$  stetig sind. Dann sind auch die Funktionen

$$\begin{aligned} \alpha f(x) \text{ mit } \alpha \in \mathbb{R}, \quad & f(x) \pm g(x), \quad f(x) \cdot g(x), \\ \frac{f(x)}{g(x)}, \text{ falls } & g(x^*) \neq 0, \end{aligned}$$

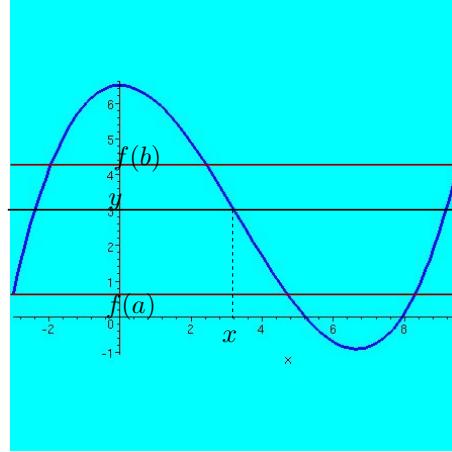
stetig in  $x^*$ . Ist die Verknüpfung  $f(g(x))$  der beiden Funktionen definiert, dann ist auch  $f(g(x))$  stetig in  $x^*$ .

*Beispiel 12.29.*

- i) Jedes Polynom ist stetig.
- ii) Alle rationalen Funktionen sind in ihrem Definitionsbereich stetig.

- iii)  $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  mit  $f(x) = 1/x$  ist stetig.
- iv)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = |2x^2 - x|$  ist stetig.
- v)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = e^{\sin(5 \cos(x^2)) + x^3}$  ist stetig.

**Satz 12.30** (Zwischenwertsatz). Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Dann nimmt  $f$  jeden Wert zwischen  $f(a)$  und  $f(b)$  an, d.h.  
für jedes  $y \in [f(a), f(b)]$  (bzw.  $y \in [f(b), f(a)]$  falls  $f(b) < f(a)$ ) gibt es ein  $x \in [a, b]$  mit  $f(x) = y$ .



**Satz 12.31** (Weierstraß<sup>31</sup>). Jede auf einem abgeschlossenen Intervall  $[a, b]$  stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  nimmt auf diesem Intervall ihr Maximum und Minimum an, d.h.  
es gibt  $\tilde{x}, \hat{x} \in [a, b]$  mit

$$\begin{aligned} f(\tilde{x}) &= \max\{f(x) : x \in [a, b]\}, \text{ und} \\ f(\hat{x}) &= \min\{f(x) : x \in [a, b]\}. \end{aligned}$$

---

<sup>31</sup>Karl Theodor Wilhelm Weierstraß 1815 - 1897

## 13 Differenzierbarkeit I

**Definition 13.1** (Sekante). Sei  $D \subseteq \mathbb{R}$  und sei  $f : D \rightarrow \mathbb{R}$  eine Funktion. Für  $x_0, x_1 \in D$  heißt die Gerade

$$s(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0)$$

Sekante von  $f$  durch die Punkte  $(x_0, f(x_0))$  und  $(x_1, f(x_1))$ .

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

ist die *Steigung der Sekante*.

**Definition 13.2** (Differenzierbarkeit). Sei  $D \subseteq \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion, und sei  $x^* \in D$ .  $f$  heißt differenzierbar in  $x^*$ , wenn der Grenzwert (der Sekantensteigungen)

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

existiert. In diesem Fall wird der Grenzwert mit  $f'(x^*)$  oder auch  $\frac{df}{dx}(x^*)$  bezeichnet, und heißt *Ableitung von  $f$  in  $x^*$* . Ist  $f$  in jedem Punkt  $x^* \in D$  differenzierbar, dann heißt  $f$  differenzierbar, und die Funktion

$$f : D \rightarrow \mathbb{R} \text{ mit } x \mapsto f'(x)$$

heißt *Ableitung von  $f$* . Für  $f'(x)$  schreibt man auch  $\frac{df}{dx}(x)$  bzw. einfach  $\frac{df}{dx}$ .

Ist  $f'$  eine stetige Funktion, dann heißt  $f$  stetig differenzierbar.

**Bemerkung 13.3.** Ist  $f$  in  $x^*$  differenzierbar, dann ist

$$t(x) = f(x^*) + f'(x^*)(x - x^*), \quad x \in \mathbb{R},$$

die Gleichung der Tangente durch  $(x^*, f(x^*))$ .

*Beispiel 13.4.*

- i)  $f(x) = x$  ist (stetig) differenzierbar mit  $f'(x) = 1$ .
- ii)  $f(x) = x^2$  ist (stetig) differenzierbar mit  $f'(x) = 2x$ .
- iii)  $f(x) = |x|$  ist nicht differenzierbar in  $x^* = 0$ .

**Satz 13.5.** Jede differenzierbare Funktion ist stetig (aber nicht umgekehrt).

**Satz 13.6.** Sei  $D \subseteq \mathbb{R}$ , und seien  $f, g : D \rightarrow \mathbb{R}$  Funktionen, die in  $x^* \in D$  differenzierbar sind. Dann sind auch  $\alpha f$ ,  $f \pm g$ ,  $f \cdot g$  in  $x^*$  differenzierbar, und es gilt:

- i)  $(\alpha f)'(x^*) = \alpha f'(x^*)$ ,
- ii)  $(f \pm g)'(x^*) = f'(x^*) \pm g'(x^*)$ ,
- iii)  $(f \cdot g)'(x^*) = f'(x^*) \cdot g(x^*) + f(x^*) \cdot g'(x^*)$ . (Produktregel)

Ist  $g(x^*) \neq 0$ , dann ist auch  $\frac{f}{g}$  in  $x^*$  differenzierbar mit

$$\text{iv)} \quad \left(\frac{f}{g}\right)'(x^*) = \frac{f'(x^*) \cdot g(x^*) - f(x^*) \cdot g'(x^*)}{g(x^*)^2}. \quad (\text{Quotientenregel})$$

Ist die Verknüpfung  $f \circ g$  definiert, und ist  $f$  in  $g(x^*)$  diffbar, dann ist  $f \circ g$  in  $x^*$  differenzierbar mit

$$\text{v)} \quad (f \circ g)'(x^*) = f'(g(x^*)) \cdot g'(x^*). \quad (\text{Kettenregel})$$

**Satz 13.7** (Ableitung elementarer Funktionen I).

Funktion	Ableitung
$c$ (Konstante)	0
$x^n, n \in \mathbb{Z} \setminus \{0\}$ ,	$n x^{n-1}$
$e^x$	$e^x$
$a^x, a > 0$ ,	$a^x \ln(a)$
$x^a, x > 0, a > 0$ ,	$a x^{a-1}$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x) = \frac{\sin(x)}{\cos(x)}$	$\frac{1}{\cos^2(x)}$

**Bemerkung 13.8.** Sei  $f : D \rightarrow \mathbb{R}$  differenzierbar.  $f$  ist genau dann monoton steigend (bzw. fallend), falls  $f'(x) \geq 0$  (bzw.  $f'(x) \leq 0$ ) für alle  $x \in D$ . Ist  $f'(x) > 0$  (bzw.  $f'(x) < 0$ ) für alle  $x \in D$ , dann ist  $f$  streng monoton steigend (bzw. fallend).

**Satz 13.9** (Ableitung der Umkehrfunktion). Sei  $D \subseteq \mathbb{R}$ ,  $f : D \rightarrow \mathbb{R}$  eine stetige, streng monotone Funktion, und sei  $\tilde{D} = f(D)$ . Dann existiert die Umkehrfunktion  $f^{-1} : \tilde{D} \rightarrow \mathbb{R}$ . Ist  $f$  in  $x^*$  differenzierbar mit  $f'(x^*) \neq 0$ , so ist  $f^{-1}$  in  $y^* = f(x^*)$  differenzierbar, und es gilt

$$(f^{-1})'(y^*) = \frac{1}{f'(x^*)} = \frac{1}{f'(f^{-1}(y^*))}$$

**Satz 13.10** (Ableitung elementarer Funktionen II).

Funktion	Ableitung
$\ln(x)$	$\frac{1}{x}$
$\log_a(x), a > 0, a \neq 1$	$\frac{1}{x \ln(a)}$
$\arcsin(x), x \in (-1, 1)$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos(x), x \in (-1, 1)$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\frac{1}{1+x^2}$

Beispiel 13.11.

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

**Satz 13.12** (Regel von de l'Hospital<sup>32</sup>). Sei  $D \subseteq \mathbb{R}$ ,  $f, g : D \rightarrow \mathbb{R}$  differenzierbar, und sei  $x^* \in D$  ( $\pm\infty$  zugelassen). Gilt

$$\lim_{x \rightarrow x^*} f(x) = \lim_{x \rightarrow x^*} g(x) = 0,$$

oder

$$\lim_{x \rightarrow x^*} |f(x)| = \lim_{x \rightarrow x^*} |g(x)| = \infty,$$

und existiert  $\lim_{x \rightarrow x^*} \frac{f'(x)}{g'(x)}$  ( $\pm\infty$  zugelassen), dann gilt

$$\lim_{x \rightarrow x^*} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x^*} \frac{f'(x)}{g'(x)}.$$

Beispiel 13.13.

$$1. \lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1.$$

$$2. \lim_{x \rightarrow \infty} x^2 e^{-x} = \lim_{x \rightarrow \infty} \frac{x^2}{e^x} = \lim_{x \rightarrow \infty} \frac{2x}{e^x} = \lim_{x \rightarrow \infty} \frac{2}{e^x} = 0.$$

---

<sup>32</sup>Guillaume de l'Hospital, 1661-1704

$$3. \lim_{x \rightarrow 0+} x \ln(x) = \lim_{x \rightarrow 0+} \frac{\ln(x)}{1/x} = \lim_{x \rightarrow 0+} \frac{1/x}{-1/x^2} = \lim_{x \rightarrow 0+} (-x) = 0.$$

**Definition 13.14** (Höhere Ableitungen). Sei  $D \subseteq \mathbb{R}$  und  $f : D \rightarrow \mathbb{R}$  eine differenzierbare Funktion. Ist die Ableitung  $f' : D \rightarrow \mathbb{R}$  in  $x^* \in D$  differenzierbar, dann heißt

$$(f')'(x^*)$$

die *zweite* Ableitung von  $f$  in  $x^*$  und wird mit  $f''(x^*)$  bezeichnet.

Analog definiert man rekursiv die 3-te, 4-te usw. Ableitung in  $x^*$ . Allgemein wird  $n$ -te Ableitung von  $f$  in  $x^*$  mit  $f^{(n)}(x^*)$  bezeichnet. Die Funktion selbst wird auch als 0-te Ableitung  $f^{(0)}$  bezeichnet.

**Definition 13.15** (Konvexe/konkave Funktionen). Sei  $D \subseteq \mathbb{R}$ . Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt

- i) *konvex* in einem Intervall  $[a, b] \subseteq D$ , falls für alle  $x_0, x_1 \in [a, b]$  und alle  $\lambda \in [0, 1]$  gilt

$$f((1 - \lambda)x_0 + \lambda x_1) \leq (1 - \lambda)f(x_0) + \lambda f(x_1),$$

d.h. die Sekante, die  $(x_0, f(x_0))$  und  $(x_1, f(x_1))$  verbindet, liegt in diesem Bereich oberhalb des Funktionsgraphen.

- ii) *konkav* in einem Intervall  $[a, b] \subseteq D$ , falls für alle  $x_0, x_1 \in [a, b]$  und alle  $\lambda \in [0, 1]$  gilt

$$f((1 - \lambda)x_0 + \lambda x_1) \geq (1 - \lambda)f(x_0) + \lambda f(x_1),$$

d.h. die Sekante, die  $(x_0, f(x_0))$  und  $(x_1, f(x_1))$  verbindet, liegt in diesem Bereich unterhalb des Funktionsgraphen.

**Satz 13.16.** Sei  $D \subseteq \mathbb{R}$  und  $f : D \rightarrow \mathbb{R}$  in  $[a, b] \subseteq D$  zweimal differenzierbar. Dann gilt

- i)  $f$  ist konvex in  $[a, b]$  genau dann, wenn  $f''(x) \geq 0$  für alle  $x \in [a, b]$ , d.h. die erste Ableitung ist monoton steigend.
- ii)  $f$  ist konkav in  $[a, b]$  genau dann, wenn  $f''(x) \leq 0$  für alle  $x \in [a, b]$ , d.h. die erste Ableitung ist monoton fallend.

**Definition 13.17** (lokale/globale Extrema). Sei  $D \subseteq \mathbb{R}$  und  $f : D \rightarrow \mathbb{R}$  eine Funktion.  $x^* \in D$  heißt

- i) *lokales (relatives) Maximum*, falls ein  $\epsilon > 0$  existiert, so dass

$$f(x) \leq f(x^*) \text{ für alle } x \in (x^* - \epsilon, x^* + \epsilon) \cap D.$$

- ii) *lokales (relatives) Minimum*, falls ein  $\epsilon > 0$  existiert, so dass

$$f(x) \geq f(x^*) \text{ für alle } x \in (x^* - \epsilon, x^* + \epsilon) \cap D.$$

- iii) *globales Maximum*, falls

$$f(x) \leq f(x^*) \text{ für alle } x \in D.$$

- iv) *globales Minimum*, falls

$$f(x) \geq f(x^*) \text{ für alle } x \in D.$$

- v) Lokale/globale Minima oder Maxima werden als lokale/globale Extrema bzw. Extremwerte der Funktion bezeichnet.

**Satz 13.18** (Kriterium für lokale Extrema).

- i) Sei  $f : [a, b] \rightarrow \mathbb{R}$  differenzierbar, und sei  $x^* \in (a, b)$ . Ist  $x^*$  ein lokales Minimum oder Maximum, dann gilt  $f'(x^*) = 0$ .
- ii) Sei  $f : [a, b] \rightarrow \mathbb{R}$  zweimal differenzierbar, und sei  $x^* \in (a, b)$ .
  - Ist  $f'(x^*) = 0$  und  $f''(x^*) < 0$ , dann ist  $x^*$  ein lokales Maximum.
  - Ist  $f'(x^*) = 0$  und  $f''(x^*) > 0$ , dann ist  $x^*$  ein lokales Minimum.

**Satz 13.19.** Sei  $f : [a, b] \rightarrow \mathbb{R}$   $n$ -mal differenzierbar, und sei  $x^* \in (a, b)$ . Sei  $f'(x^*) = 0$  und  $f^{(n)}(x^*)$  die erste Ableitung, die an der Stelle  $x^*$  nicht gleich 0 ist, d.h.,  $f^{(i)}(x^*) = 0$ ,  $1 \leq i \leq n - 1$  und  $f^{(n)}(x^*) \neq 0$ . Ist  $n$  gerade und

- i)  $f^{(n)}(x^*) < 0$ , dann ist  $x^*$  ein lokales Maximum.
- ii)  $f^{(n)}(x^*) > 0$ , dann ist  $x^*$  ein lokales Minimum.

**Definition 13.20** (Wendepunkte, Sattelpunkte). Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion.  $x^* \in (a, b)$  heißt *Wendepunkt*, wenn es ein  $\epsilon > 0$  gibt, so dass  $f$  in  $[x^* - \epsilon, x^*]$  konkav und in  $[x^*, x^* + \epsilon]$  konvex ist (oder umgekehrt).

Ist  $f$  in  $x^*$  differenzierbar, und gilt zudem  $f'(x^*) = 0$ , dann wird der Wendepunkt auch *Sattelpunkt* genannt.

**Satz 13.21** (Kriterium für Wendepunkte).

- i) Sei  $f : [a, b] \rightarrow \mathbb{R}$  zweimal differenzierbar, und sei  $x^* \in (a, b)$ . Ist  $x^*$  ein Wendepunkt, dann gilt  $f''(x^*) = 0$ .
- ii) Sei  $f : [a, b] \rightarrow \mathbb{R}$  dreimal differenzierbar, und sei  $x^* \in (a, b)$ . Ist  $f''(x^*) = 0$  und  $f'''(x^*) \neq 0$ , dann ist  $x^*$  ein Wendepunkt.

**Satz 13.22.** Sei  $f : [a, b] \rightarrow \mathbb{R}$   $n$ -mal differenzierbar, und sei  $x^* \in (a, b)$ . Sei  $f'(x^*) = 0$  und  $f^{(n)}(x^*)$  die erste Ableitung, die an der Stelle  $x^*$  nicht gleich 0 ist, d.h.,  $f^{(i)}(x^*) = 0$ ,  $1 \leq i \leq n - 1$  und  $f^{(n)}(x^*) \neq 0$ . Ist  $n$  ungerade, dann ist  $x^*$  ein Sattelpunkt.

**Bemerkung 13.23.** Maxima und Minima am Rand eines Intervalls können i.A. nicht mit Hilfe der Differentialrechnung bestimmt werden.

*Beispiel 13.24.* Sei  $f : [0, 1] \rightarrow \mathbb{R}$  mit  $f(x) = x$ . Dann ist das globale Maximum bei 1 und das globale Minimum bei 0, aber in beiden Punkten ist die Ableitung gleich 1.

**Bemerkung 13.25.** Um alle Maxima und Minima einer Funktion zu bestimmen, müssen die Randpunkte und alle Punkte, in denen die Funktion nicht differenzierbar ist, gesondert untersucht werden.

*Beispiel 13.26.* Sei  $f : [-1, 1] \rightarrow \mathbb{R}$  mit

$$f(x) = \begin{cases} 1 + x, & -1 \leq x \leq 0, \\ 1 - x + x^2, & 0 \leq x \leq 1. \end{cases}$$

**Satz 13.27** (Satz von Rolle<sup>33</sup>). Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine differenzierbare Funktion, mit  $f(a) = f(b)$ . Dann gibt es ein  $x^* \in (a, b)$  mit  $f'(x^*) = 0$ .

**Satz 13.28** (Mittelwertsatz der Differentialrechnung). Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine differenzierbare Funktion. Dann gibt es ein  $x^* \in (a, b)$  mit

$$f'(x^*) = \frac{f(b) - f(a)}{b - a}.$$

---

<sup>33</sup>Michel Rolle, 1652–1719

**Satz 13.29.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine differenzierbare Funktion mit  $f'(x) = 0$  für alle  $x \in (a, b)$ . Dann ist  $f$  konstant.

**Satz 13.30.** Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine differenzierbare Funktion, und für eine Konstante  $c$  gelte  $f'(x) = c f(x)$  für alle  $x \in \mathbb{R}$ . Dann ist  $f(x) = f(0)e^{cx}$ .

## 14 Taylor- und Potenzreihen

**Definition 14.1** (Taylorpolynom<sup>34</sup>). Sei  $D \subseteq \mathbb{R}$ , und sei  $f : D \rightarrow \mathbb{R}$  eine Funktion, die an der Stelle  $x^* \in D$  mindestens  $n$ -mal differenzierbar ist. Dann heißt

$$\begin{aligned} T_n(x) &= \sum_{k=0}^n \frac{f^{(k)}(x^*)}{k!} (x - x^*)^k \\ &= f(x^*) + f'(x^*) (x - x^*) + \frac{f''(x^*)}{2} (x - x^*)^2 + \cdots + \frac{f^{(n)}(x^*)}{n!} (x - x^*)^n \end{aligned}$$

das zu  $f$  gehörige *Taylorpolynom vom Grad  $n$*  an der Stelle  $x^*$ , bzw. mit *Entwickelpunkt  $x^*$* . Man schreibt daher auch manchmal  $T_n(x^*; x)$ .

**Bemerkung 14.2.**  $T_1(x)$  ist gerade die Tangente an der Stelle  $x^*$  (siehe 13.3).

*Beispiel 14.3.* Sei  $f(x) = \sin(x)$ . Dann ist

$$f^{(k)}(x) = \begin{cases} (-1)^{k/2} \sin(x) & : k \text{ gerade} \\ (-1)^{(k-1)/2} \cos(x) & : k \text{ ungerade} \end{cases}.$$

Für den Entwickelpunkt  $x^* = 0$  ergibt sich somit  $f^{(k)}(0) = 0$ , falls  $k$  gerade und  $f^{(k)}(0) = (-1)^{(k-1)/2}$  für  $k$  ungerade. Damit ist das Taylorpolynom von  $\sin(x)$  an der Stelle  $x^* = 0$  gegeben durch

$$\begin{aligned} T_n(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^{\lfloor (n-1)/2 \rfloor} \frac{x^n}{n!} \\ &= \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} (-1)^k \cdot \frac{x^{2k+1}}{(2k+1)!}. \end{aligned}$$

**Definition 14.4** (Restglied). Sei  $D \subseteq \mathbb{R}$ , und sei  $f : D \rightarrow \mathbb{R}$  eine Funktion, die an der Stelle  $x^* \in D$  mindestens  $n$ -mal differenzierbar ist. Der „Fehler“

$$R_n(x) = f(x) - T_n(x^*; x)$$

heißt *Restglied*.

**Satz 14.5** (Taylor). Sei  $D \subseteq \mathbb{R}$ , und sei  $f : D \rightarrow \mathbb{R}$  eine  $(n+1)$ -mal stetig differenzierbare Funktion. Dann gilt:

$$|R_n(x)| = |f(x) - T_n(x^*; x)| \leq \frac{M}{(n+1)!} |x - x^*|^{n+1},$$

wobei  $M$  so gewählt ist, dass  $|f^{(n+1)}(x)| \leq M$  für alle  $x \in D$ .

*Beispiel 14.6.* Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = \sin(x)$ ,  $x^* = 0$  und  $n = 11$ . Dann ist  $f^{(12)}(x) = \sin(x)$ , und somit gilt die Fehlerabschätzung

$$|R_{11}(1/2)| = |\sin(1/2) - T_{11}(0; 1/2)| \leq \frac{1}{(12)!} |1/2|^{12} < 10^{-12}.$$

---

<sup>34</sup>Brook Taylor, 1685-1731

**Definition 14.7** (Landausymbole<sup>35</sup>). Sei  $D \subseteq \mathbb{R}$ , seien  $f, g : D \rightarrow \mathbb{R}$ , und sei  $x^* \in D$ .

- i)  $f$  heißt von *Ordnung groß O von  $g$  für  $x$  gegen  $x^*$*  falls es Konstanten  $M > 0$ ,  $\delta > 0$  gibt, so dass für  $x \in (x^* - \delta, x^* + \delta)$  gilt

$$\left| \frac{f(x)}{g(x)} \right| \leq M.$$

Man schreibt dann  $f = \mathcal{O}(g)$  für  $x \rightarrow x^*$ . In diesem Fall wächst  $f$  höchstens so schnell wie  $g$ . Die Bedingung ist sicherlich erfüllt, falls  $\lim_{x \rightarrow x^*} \left| \frac{f(x)}{g(x)} \right| \leq M$  gilt. Ist  $x^* = \pm\infty$ , dann schreibt man  $f = \mathcal{O}(g)$  für  $x \rightarrow x^*$  falls

$$\lim_{x \rightarrow x^*} \left| \frac{f(x)}{g(x)} \right| \leq M.$$

- ii)  $f$  heißt von *Ordnung klein o von  $g$  für  $x$  gegen  $x^*$*  falls

$$\lim_{x \rightarrow x^*} \left| \frac{f(x)}{g(x)} \right| = 0.$$

Man schreibt dann  $f = o(g)$  für  $x \rightarrow x^*$ . In diesem Fall wächst  $f$  (echt) langsamer als  $g$ .

**Bemerkung 14.8.** Mit dieser Notation ist also für  $x$  gegen  $x^*$

$$f(x) = T_n(x^*; x) + \mathcal{O}((x - x^*)^{n+1}).$$

**Definition 14.9** (Taylorreihe). Sei  $D \subseteq \mathbb{R}$ , und sei  $f : D \rightarrow \mathbb{R}$  eine beliebig oft differenzierbare Funktion. Sei  $x^* \in D$ . Die Reihe

$$T_\infty(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x^*)}{k!} (x - x^*)^k$$

heißt *Taylorreihe* mit Entwicklungspunkt  $x^*$ . Man schreibt auch manchmal  $T_\infty(x^*; x)$ .

**Beispiel 14.10.** Sei  $f(x) = \frac{1}{1-x}$  und  $x^* = 0$ . Dann ist

$$T_\infty(x) = \sum_{k=0}^{\infty} x^k,$$

und es ist  $f(x) = T_\infty(x)$  für  $|x| < 1$  (siehe 11.24).

**Definition 14.11** (Potenzreihen). Sei  $(a_n)_{n \in \mathbb{N}_0}$  eine Folge reeller (bzw. komplexer) Zahlen und  $x^* \in \mathbb{R}$  (bzw.  $\in \mathbb{C}$ ). Für  $x \in \mathbb{R}$  (bzw.  $\in \mathbb{C}$ ) heißt die Reihe

$$\sum_{k=0}^{\infty} a_k (x - x^*)^k$$

*Potenzreihe* mit Entwicklungspunkt  $x^*$ .

**Bemerkung 14.12.** Jede Taylorreihe ist eine Potenzreihe.

**Lemma 14.13.** Konvergiert die Potenzreihe  $\sum_{k=0}^{\infty} a_k (x - x^*)^k$  für ein  $x_1 \in \mathbb{R}$ , dann konvergiert sie (absolut) für alle  $x \in \mathbb{R}$  mit

$$|x - x^*| < |x_1 - x^*|.$$

---

<sup>35</sup>Edmund Georg Hermann Landau, 1877-1938

**Definition 14.14** (Supremum/Infimum). Sei  $A \subseteq \mathbb{R}$ . Ein Element  $s \in \mathbb{R}$  heißt *obere Schranke* von  $A$ , falls  $a \leq s$  für alle  $a \in A$ . Ein Element  $r \in \mathbb{R}$  heißt *untere Schranke* von  $A$ , falls  $r \leq a$  für alle  $a \in A$ .

- i)  $s \in \mathbb{R}$  heißt *Supremum von  $A$*  im Zeichen  $s = \sup A$ , falls  $s$  die kleinste obere Schranke von  $A$  ist.
- ii)  $r \in \mathbb{R}$  heißt *Infimum von  $A$*  im Zeichen  $r = \inf A$ , falls  $r$  die größte untere Schranke von  $A$  ist.
- iii) Falls keine oberen (unteren) Schranken existieren, dann setzt man  $\sup A = \infty$  ( $\inf A = -\infty$ ).
- iv) Ist  $s = \sup A < \infty$  und  $s \in A$ , dann heißt  $s$  *maximales Element (Maximum)* von  $A$ , und man schreibt  $s = \max A$ .
- v) Ist  $r = \inf A > -\infty$  und  $r \in A$ , dann heißt  $r$  *minimales Element (Minimum)* von  $A$ , und man schreibt  $r = \min A$ .

*Beispiel 14.15.*

1. Sei  $A = (0, 1)$ . Dann ist  $\sup A = 1$  und  $\inf A = 0$ .
2. Sei  $A = [0, 1]$ . Dann ist  $\max A = 1$  und  $\min A = 0$ .
3.  $\inf \mathbb{N} = 1$  und  $\sup \mathbb{N} = \infty$ .

**Definition 14.16** (Konvergenzbereich, -radius). Sei  $(a_n)_{n \in \mathbb{N}_0}$  eine Folge reeller Zahlen,  $x^* \in \mathbb{R}$ , und  $\sum_{k=0}^{\infty} a_k (x - x^*)^k$  die zugehörige Potenzreihe.

$$C = \{x \in \mathbb{R} : \sum_{k=0}^{\infty} a_k (x - x^*)^k \text{ konvergiert}\}$$

heißt *Konvergenzbereich* der Potenzreihe, und

$$\rho = \sup\{|x - x^*| : x \in C\}$$

heißt der *Konvergenzradius* der Potenzreihe.

*Beispiel 14.17.* Für die Potenzreihe  $\sum_{k=0}^{\infty} x^k$  (Entwicklungsplatz  $x^* = 0$ ) ist

$$C = \{x \in \mathbb{R} : |x| < 1\} \text{ und } \rho = 1.$$

**Satz 14.18.** Sei  $\sum_{k=0}^{\infty} a_k (x - x^*)^k$  eine Potenzreihe mit Konvergenzradius  $\rho$ . Dann ist die Reihe

- i) absolut konvergent für alle  $x \in \mathbb{R}$  mit  $|x - x^*| < \rho$ , und
- ii) divergent für alle  $x \in \mathbb{R}$  mit  $|x - x^*| > \rho$ .

**Satz 14.19.** Sei  $\sum_{k=0}^{\infty} a_k (x - x^*)^k$  eine Potenzreihe mit Konvergenzradius  $\rho$ , und es sei  $a_k \neq 0$  für fast alle  $k \in \mathbb{N}_0$ . Existiert der Grenzwert  $\lim_{k \rightarrow \infty} |a_k/a_{k+1}|$  ( $\infty$  zugelassen), dann ist

$$\rho = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right|.$$

*Beispiel 14.20.* Für die Potenzreihe  $\sum_{k=1}^{\infty} \frac{1}{k} x^k$ , also  $a_k = 1/k$ , ist

$$\lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| = \lim_{k \rightarrow \infty} \frac{k+1}{k} = 1.$$

Somit ist der Konvergenzradius  $\rho = 1$ , und die Reihe konvergiert für  $\{x \in \mathbb{R} : |x| < 1\}$ . Am Rand dieses Bereiches erhalten wir für  $x = 1$  die divergente harmonische Reihe, und für  $x = -1$  die konvergente alternierende harmonische Reihe (siehe 1, 11.31). Also ist der Konvergenzbereich dieser Reihe gegeben durch

$$C = \{x \in \mathbb{R} : -1 \leq x < 1\}.$$

**Satz 14.21.** Sei  $\sum_{k=0}^{\infty} a_k (x - x^*)^k$  eine Potenzreihe mit Konvergenzradius  $\rho$ . Existiert der Grenzwert  $\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}$  ( $\infty$  zugelassen), dann gilt

$$\rho = \frac{1}{\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}}.$$

Ist der Grenzwert 0, dann ist  $\rho = \infty$ , und ist der Grenzwert  $\infty$  (also bestimmte Divergenz), dann ist  $\rho = 0$ .

Beispiel 14.22. Für die Potenzreihe  $\sum_{k=0}^{\infty} k^k x^k$ , also  $a_k = k^k$ , ist

$$\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \lim_{k \rightarrow \infty} k = \infty.$$

Somit ist der Konvergenzradius  $\rho = 0$ , d.h. die Reihe konvergiert nur für  $x = 0$ .

**Bemerkung 14.23.**

- i) Vorsicht! Auch wenn die Taylorreihe  $T_{\infty}(x)$  einer Funktion  $f(x)$  an einer Stelle  $\hat{x}$  konvergiert, muss nicht unbedingt gelten:  $T_{\infty}(\hat{x}) = f(\hat{x})$ .
- ii) Die Taylorreihe konvergiert genau für diejenigen  $x$  aus dem Konvergenzbereich gegen  $f(x)$ , für die das Restglied aus (14.4) gegen 0 konvergiert mit  $n \rightarrow \infty$ .

**Satz 14.24.** Einige Taylorreihen und ihr Konvergenzradien  $\rho$ :

- i)  $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k, \quad x^* = 0, \rho = 1,$
- ii)  $\ln(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x-1)^k, \quad x^* = 1, \rho = 1,$
- iii)  $e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \quad x^* = 0, \rho = \infty,$
- iv)  $\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}, \quad x^* = 0, \rho = \infty,$
- v)  $\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}, \quad x^* = 0, \rho = \infty,$

**Satz 14.25.** Eine Potenzreihe  $g(x) = \sum_{k=0}^{\infty} a_k (x - x^*)^k$  ist innerhalb ihres Konvergenzbereiches beliebig oft stetig differenzierbar. Die Ableitungen können durch gliedweises Differenzieren erhalten werden, d.h.

$$g'(x) = \sum_{k=1}^{\infty} k a_k (x - x^*)^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} (x - x^*)^k,$$

und der Konvergenzbereich von  $g'(x)$  ist gleich dem Konvergenzbereich von  $g(x)$ .

Beispiel 14.26.

$$\cos(x) = (\sin(x))' = \left( \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} \right)' = \sum_{k=0}^{\infty} (2k+1) \frac{(-1)^k}{(2k+1)!} x^{2k} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}.$$

## 15 Integralrechnung I

**Definition 15.1** (Zerlegung). Sei  $[a, b] \subset \mathbb{R}$  ein Intervall. Eine Menge  $Z = \{x_0, x_1, \dots, x_n\}$  mit  $a = x_0 < x_1 < x_2 \dots < x_n = b$  heißt *Zerlegung* von  $[a, b]$  bzgl. den Stützstellen  $x_0, x_1, \dots, x_n$ .

**Definition 15.2** (Ober- Undersumme). Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion, und sei  $Z = \{x_0, \dots, x_n\}$  eine Zerlegung von  $[a, b]$ . Dann heißt

1.  $U_f(Z) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \cdot \inf\{f(x) : x \in [x_i, x_{i+1}]\}$  *Undersumme* von  $f$  bzgl.  $Z$ .
2.  $O_f(Z) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \cdot \sup\{f(x) : x \in [x_i, x_{i+1}]\}$  *Obersumme* von  $f$  bzgl.  $Z$ .

**Bemerkung 15.3.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion.

1. Für jede Zerlegung  $Z$  von  $[a, b]$  gilt  $U_f(Z) \leq O_f(Z)$ .
2. Sind  $Z_1, Z_2$  Zerlegungen von  $[a, b]$  mit  $Z_1 \subseteq Z_2$ , dann gilt

$$U_f(Z_2) \geq U_f(Z_1) \quad \text{und} \quad O_f(Z_2) \leq O_f(Z_1).$$

3. Für zwei beliebige Zerlegung  $Z_1, Z_2$  von  $[a, b]$  gilt  $U_f(Z_1) \leq O_f(Z_2)$ .

**Bemerkung 15.4.** Insbesondere sind Undersummen nach oben beschränkt, und Obersummen nach unten.

**Definition 15.5** ((Riemann<sup>36</sup>)-Integral). Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte Funktion.

1.  $U_f = \sup\{U_f(Z) : Z$  Zerlegung von  $[a, b]\}$  heißt (*Riemann'sches*) *Unterintegral* von  $f$  über  $[a, b]$ .
2.  $O_f = \inf\{O_f(Z) : Z$  Zerlegung von  $[a, b]\}$  heißt (*Riemann'sches*) *Oberintegral* von  $f$  über  $[a, b]$ .
3.  $f$  heißt (*Riemann-*) *integrierbar*, falls  $U_f, O_f$  existieren und  $U_f = O_f$  gilt.
4. Ist  $f$  (*Riemann-*) integrierbar, dann heißt das Unterintegral  $U_f$  (bzw. Oberintegral  $O_f$ ) das (*Riemann-*) *Integral* von  $f(x)$  über  $[a, b]$  und wird mit

$$\int_a^b f(x) dx$$

bezeichnet. Abkürzend schreibt man auch nur  $\int_a^b f$ .

**Notation 15.6.** Für ein Integral  $\int_a^b f(x) dx$  heißt die Funktion  $f$  der *Integrand*,  $x$  heißt die *Integrationsvariable* und  $a, b$  heißen die *Integrationsgrenzen*.

**Satz 15.7.** Seien  $f, g : [a, b] \rightarrow \mathbb{R}$  integrierbar über  $[a, b]$ , und sei  $\alpha \in \mathbb{R}$ . Dann gilt

- i)  $\alpha f$  und  $f + g$  sind integrierbar, und es ist:

$$\int_a^b \alpha f = \alpha \int_a^b f \quad \text{und} \quad \int_a^b (f + g) = \int_a^b f + \int_a^b g.$$

- ii)  $f \cdot g$  ist integrierbar.

---

<sup>36</sup>Georg Friedrich Bernhard Riemann, 1826-1866

**Bemerkung 15.8.** Seien  $f, g : [a, b] \rightarrow \mathbb{R}$  integrierbar über  $[a, b]$ . Ist  $f(x) \leq g(x)$  für alle  $x \in [a, b]$ , dann ist

$$\int_a^b f \leq \int_a^b g.$$

Insbesondere ist im Falle  $f(x) \geq 0$  für alle  $x \in [a, b]$ , auch  $\int_a^b f \geq 0$ .

**Lemma 15.9.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine integrierbare Funktion, und sei  $c \in [a, b]$ . Dann ist  $f$  auch über  $[a, c]$  und  $[c, b]$  integrierbar, und es gilt:

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

**Satz 15.10.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion.

- i) Ist  $f$  monoton, dann ist  $f$  integrierbar.
- ii) Ist  $f$  stetig, dann ist  $f$  integrierbar.

**Definition 15.11.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  integrierbar. Dann setzen wir

$$\int_c^c f(x) dx = 0 \text{ und } \int_d^c f(x) dx = - \int_c^d f(x) dx$$

für für  $a \leq c \leq d \leq b$ .

**Definition 15.12** (Stammfunktion). Seien  $f, F : [a, b] \rightarrow \mathbb{R}$  zwei Funktion. Ist  $F$  differenzierbar mit  $F' = f$ , dann heißt  $F$  Stammfunktion von  $f$ .

**Satz 15.13** (Hauptsatz der Differential- und Integralrechnung). Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion. Die Funktion  $F : [a, b] \rightarrow \mathbb{R}$  mit

$$F(x) = \int_a^x f(t) dt$$

ist Stammfunktion von  $f$

**Lemma 15.14.** Sei  $F : [a, b] \rightarrow \mathbb{R}$  eine Stammfunktion der Funktion  $f : [a, b] \rightarrow \mathbb{R}$ . Dann ist  $G : [a, b] \rightarrow \mathbb{R}$  genau dann eine (weiter) Stammfunktion von  $f$ , wenn es ein  $c \in \mathbb{R}$  gibt, mit  $F(x) = G(x) + c$  für alle  $x \in [a, b]$ .

Die Stammfunktion ist also nur bis auf eine Konstante eindeutig bestimmt.

**Korollar 15.15.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig, und sei  $F$  eine Stammfunktion von  $f$ . Dann gilt:

$$\int_a^b f(x) dx = F(b) - F(a).$$

**Notation 15.16.** Sei  $F$  Stammfunktion von  $f : [a, b] \rightarrow \mathbb{R}$ . Dann setzt man

$$F(x) \Big|_a^b = F(b) - F(a).$$

**Definition 15.17** (Bestimmtes und Unbestimmtes Integral). Sei  $f : [a, b] \rightarrow \mathbb{R}$  integrierbar. Eine Stammfunktion  $F$  von  $f$  heißt auch *unbestimmtes Integral* von  $f$  und wird mit

$$\int f(x) dx$$

(bzw.  $\int f$ ) bezeichnet. Manchmal versteht man unter dem unbestimmten Integral auch die Menge aller Stammfunktionen, und schreibt (relativ unpräzise)

$$\int f(x) dx = F(x) + C,$$

wobei  $F$  eine Stammfunktion ist, und  $C$  deutet die Menge aller Konstanten an.

Hingegen nennt man das Integral  $\int_a^b f(x) dx$  mit gegebenen Integrationsgrenzen  $a, b$  auch *bestimmtes Integral*.

*Beispiel 15.18.* Für  $n \in \mathbb{N}_0$  ist

$$\int x^n dx = \frac{x^{n+1}}{n+1} (+C).$$

**Satz 15.19.** Sei  $f(x) = \sum_{k=0}^{\infty} a_k (x - x^*)^k$  eine Potenzreihe mit Konvergenzradius  $\rho$ . Dann ist auch die Funktion  $F(x)$  die durch gliedweises Integrieren der Potenzreihe entsteht, d.h.

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x - x^*)^{k+1},$$

konvergent in  $\{x \in \mathbb{R} : |x - x^*| < \rho\}$ , und  $F(x)$  ist Stammfunktion von  $f(x)$ .

**Satz 15.20** (Einige Stammfunktionen).

Funktion	Stammfunktion
$c$ , wobei $c \in \mathbb{R}$	$c \cdot x$
$x^a$ , $a \in \mathbb{R}$ , $a \neq -1$	$\frac{x^{a+1}}{a+1}$
$x^{-1}$ , $x \neq 0$	$\ln( x )$
$e^x$	$e^x$
$a^x$ , $a \in \mathbb{R}$ , $a > 0$	$\frac{a^x}{\ln(a)}$
$\sin(x)$	$-\cos(x)$
$\cos(x)$	$\sin(x)$

**Satz 15.21** (Substitutionsregel). Sei  $g : [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar, und  $f : g([a, b]) \rightarrow \mathbb{R}$  eine stetige Funktion mit Stammfunktion  $F$ . Dann ist  $F \circ g$  Stammfunktion von  $(f \circ g) \cdot g'(x)$  und es gilt

$$\int_a^b f(g(x)) \cdot g'(x) dx = \int_{g(a)}^{g(b)} f(y) dy = F(g(b)) - F(g(a)).$$

Hier wurde  $y = g(x)$  substituiert, und mit der „informellen“ Schreibweise

$$\frac{dy}{dx} = \frac{dg}{dx} = g'(x) \text{ bzw. } dy = g'(x) dx$$

kann man sich auch (für das unbestimmte Integral) merken

$$\int f(g(x)) \cdot g'(x) dx = \int f(y) dy.$$

*Beispiele 15.22.*  $\int_{1/\pi}^{2/\pi} -\frac{\sin(1/x)}{x^2} dx$ . Sei  $g : [1/\pi, 2/\pi] \rightarrow \mathbb{R}$  mit  $g(x) = 1/x$ . Dann ist  $g$  stetig differenzierbar mit Ableitung  $g'(x) = -1/x^2$ . Die Funktion  $f(x) = \sin x$  ist auf  $[\pi/2, \pi] = g([1/\pi, 2/\pi])$  stetig mit Stammfunktion

$-\cos x$ . Somit ist

$$\begin{aligned} \int_{1/\pi}^{2/\pi} -\frac{\sin(1/x)}{x^2} dx &= \int_{1/\pi}^{2/\pi} f(g(x)) \cdot g'(x) dx \\ &= \int_{g(1/\pi)}^{g(2/\pi)} f(y) dy = \int_{\pi}^{\pi/2} \sin y dy \\ &= - \int_{\pi/2}^{\pi} \sin y dy = \cos(y) \Big|_{\pi/2}^{\pi} \\ &= \cos(\pi) - \cos(\pi/2) = -1. \end{aligned}$$

**Bemerkung 15.23.**

$$\int \frac{f'(x)}{f(x)} = \ln(|f(x)|) + C.$$

*Beispiel 15.24.*

$$\int_a^b \tan x dx = \int_a^b \frac{\sin x}{\cos x} dx = -\ln(|\cos x|) \Big|_a^b.$$

**Satz 15.25** (Partielle Integration). *Es sei  $g : [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar und  $f : [a, b] \rightarrow \mathbb{R}$  stetig mit Stammfunktion  $F$ . Dann ist  $F \cdot g$  Stammfunktion von  $f \cdot g + F \cdot g'$  und es gilt*

$$\int_a^b f(x) \cdot g(x) dx = (F(x) \cdot g(x)) \Big|_a^b - \int_a^b F(x) \cdot g'(x) dx.$$

*Bzw. man kann sich diese Regel auch merken als*

$$\int u' \cdot v = (u \cdot v) \Big| - \int u \cdot v'.$$

*Beispiel 15.26.*

1. Wir betrachten  $\int_a^b \cos(x) \cdot x dx$ . Mit  $g(x) = x$ ,  $f(x) = \cos(x)$  und so  $g'(x) = 1$ ,  $F(x) = \sin(x)$  erhalten wir

$$\begin{aligned} \int_a^b \cos(x) \cdot x dx &= \int_a^b f(x) \cdot g(x) dx \\ &= (F(x) \cdot g(x)) \Big|_a^b - \int_a^b F(x) \cdot g'(x) dx \\ &= (\sin(x) \cdot x) \Big|_a^b - \int_a^b \sin(x) dx \\ &= (\sin(b) b - \sin(a) a) - (-\cos(b) + \cos(a)) \\ &= \sin(b) b - \sin(a) a + \cos(b) - \cos(a). \end{aligned}$$

Insbesondere ist  $\sin(x) x + \cos(x)$  (eine) Stammfunktion von  $\cos(x) x$ .

2. Wir betrachten  $\int_a^b \ln(x) dx$ . Mit  $g(x) = \ln(x)$ ,  $f(x) = 1$  und so  $g'(x) = 1/x$ ,  $F(x) = x$  erhalten wir

$$\begin{aligned} \int_a^b \ln(x) dx &= \int_a^b f(x) \cdot g(x) dx \\ &= (F(x) \cdot g(x)) \Big|_a^b - \int_a^b F(x) \cdot g'(x) dx \\ &= (x \ln(x)) \Big|_a^b - \int_a^b x \cdot \frac{1}{x} dx \\ &= (b \ln(b) - a \ln(a)) - (b - a). \end{aligned}$$

Insbesondere ist  $x \ln(x) - x$  (eine) Stammfunktion von  $\ln(x)$ .

**Definition 15.27** (Uneigentliche Integrale). Sei  $a \in \mathbb{R} \cup \{-\infty\}$ ,  $b \in \mathbb{R} \cup \{\infty\}$  und  $c \in (a, b)$ . Weiterhin sei  $f : (a, b) \rightarrow \mathbb{R}$  eine Funktion.

1. Ist  $f$  auf jedem Teilintervall  $[c, d] \subset [c, b)$ , also  $c \leq d < b$ , integrierbar und existiert der Grenzwert

$$\lim_{d \rightarrow b^-} \int_c^d f(x) dx$$

so bezeichnet man ihn mit  $\int_c^b f(x) dx$ .

2. Ist  $f$  auf jedem Teilintervall  $[d, c] \subset (a, c]$ , also  $a < d \leq c b$ , integrierbar und existiert der Grenzwert

$$\lim_{d \rightarrow a^+} \int_d^c f(x) dx$$

so bezeichnet man ihn mit  $\int_a^c f(x) dx$ .

3. Existieren die Grenzwerte

$$\lim_{d \rightarrow b^-} \int_c^d f(x) dx \text{ und } \lim_{d \rightarrow a^+} \int_d^c f(x) dx$$

so setzt man

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

*Beispiel 15.28.* 1. Für  $s > 1$  betrachten wir  $\int_1^\infty \frac{1}{x^s} dx$ . Für  $d \in [1, \infty)$  ist

$$\int_1^d \frac{1}{x^s} dx = \frac{1}{1-s} \frac{1}{x^{s-1}} \Big|_1^d = \frac{1}{1-s} \left( \frac{1}{d^{s-1}} - 1 \right).$$

Wegen

$$\begin{aligned} \lim_{d \rightarrow \infty} \frac{1}{1-s} \left( \frac{1}{d^{s-1}} - 1 \right) &= \frac{1}{1-s} \cdot (-1) = \frac{1}{s-1} \\ \int_1^\infty \frac{1}{x^s} dx &= \lim_{d \rightarrow \infty} \int_1^d \frac{1}{x^s} dx = \frac{1}{s-1}. \end{aligned}$$

Für  $0 \leq s \leq 1$  existiert der obige Grenzwert nicht.

2. Für  $0 \leq s \leq 1$  betrachten wir  $\int_0^1 \frac{1}{x^s} dx$ . Für  $d \in (0, 1]$  ist

$$\int_d^1 \frac{1}{x^s} dx = \frac{1}{1-s} \frac{1}{x^{s-1}} \Big|_d^1 = \frac{1}{1-s} \left( 1 - \frac{1}{d^{s-1}} \right).$$

Wegen

$$\lim_{d \rightarrow 0} \frac{1}{1-s} \left( 1 - \frac{1}{d^{s-1}} \right) = \frac{1}{1-s} \cdot (1) = \frac{1}{1-s}$$

ist

$$\int_0^1 \frac{1}{x^s} dx = \lim_{d \rightarrow 0} \int_d^1 \frac{1}{x^s} dx = \frac{1}{1-s}.$$

Für  $1 < s$  existiert der obige Grenzwert nicht.

**Satz 15.29** (Konvergenzkriterium für Integrale).

- i) Sei  $g : (a, b) \rightarrow \mathbb{R}$  mit  $g(x) \geq 0$ . Existiert das Integral  $\int_a^b g(x) dx$  und ist  $|f(x)| \leq g(x)$  für alle  $x \in (a, b)$ , dann existiert auch  $\int_a^b f(x) dx$ . Man sagt, das  $g$  eine (konvergente) Majorante von  $f$  ist.
- ii) Sei  $g : (a, b) \rightarrow \mathbb{R}$  mit  $g(x) \geq 0$ . Existiert das Integral  $\int_a^b g(x) dx$  nicht und ist  $f(x) \geq g(x)$  für alle  $x \in (a, b)$ , dann existiert auch  $\int_a^b f(x) dx$  nicht. Man sagt, das  $g$  eine (divergente) Minorante von  $f$  ist.

**Satz 15.30** (Integralvergleichskriterium für Reihen). Sei  $f : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$  eine nicht-negative monoton fallende Funktion. Die Reihe  $\sum_{n=1}^{\infty} f(n)$  konvergiert genau dann, wenn  $\int_1^{\infty} f(x) dx$  existiert.

Beispiel 15.31. Die Reihe  $\sum_{n=1}^{\infty} \frac{1}{n^s}$  konvergiert genau dann, wenn  $s > 1$ .

## 16 Fourierreihen

**Definition 16.1** (Trigonometrisches Polynom). Sei  $T > 0$  und  $\omega = 2\pi/T$ . Eine Funktion  $T_n : \mathbb{R} \rightarrow \mathbb{R}$  der Form

$$T_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(k\omega x) + b_k \sin(k\omega x))$$

mit  $a_0, a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$  heißt *trigonometrisches Polynom* vom Grad  $n$ .

**Bemerkung 16.2.**  $T_n$  ist periodisch mit Periode  $T$ , d.h.

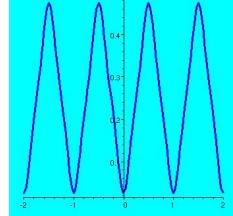
$$T_n(x+T) = T_n(x) \text{ für alle } x \in \mathbb{R}.$$

Es genügt daher solch ein Polynom auf dem Intervall  $[-\frac{T}{2}, \frac{T}{2}]$  (oder auch  $[0, T]$ ) zu untersuchen.

*Beispiel 16.3.*

$$T_3(x) = \frac{1}{4} - \frac{2}{\pi^2} \cos(2\pi x) - \frac{2}{9\pi^2} \cos(6\pi x)$$

ist ein trigonometrisches Polynom vom Grad 3 mit Periode  $T = 1$  und  $a_0 = \frac{1}{2}$ ,  $a_1 = -\frac{2}{\pi^2}$ ,  $b_1 = 0$ ,  $a_2 = b_2 = 0$ ,  $a_3 = -\frac{2}{9\pi^2}$  und  $b_3 = 0$ .



$T_3(x)$

**Definition 16.4** (Fourierpolynom<sup>37</sup>, Fourierkoeffizienten). Sei  $T > 0$  und  $f : [-T/2, T/2] \rightarrow \mathbb{R}$  eine integrierbare Funktion, und sei  $n \in \mathbb{N}$ .

$$\begin{aligned} a_k &= \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos(k\omega t) dt, \quad k = 0, \dots, n \\ b_k &= \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin(k\omega t) dt, \quad k = 1, \dots, n \end{aligned}$$

heißen die *Fourierkoeffizienten* von  $f$ , und

$$F_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(k\omega x) + b_k \sin(k\omega x))$$

heißt das *Fourierpolynom*  $n$ -ten Grades von  $f$ . Hierbei ist wieder  $\omega = 2\pi/T$ .

**Satz 16.5.** Seien  $k, l \in \mathbb{N}$ ,  $T > 0$  und  $\omega = 2\pi/T$ . Dann gilt:

$$\begin{aligned} \int_{-T/2}^{T/2} \cos(k\omega x) \cos(l\omega x) dx &= \int_{-T/2}^{T/2} \sin(k\omega x) \sin(l\omega x) dx = \begin{cases} T/2 & k = l \\ 0 & k \neq l \end{cases}, \\ \int_{-T/2}^{T/2} \cos(k\omega x) \sin(l\omega x) dx &= 0, \\ \int_{-T/2}^{T/2} \cos(k\omega x) dx &= \int_{-T/2}^{T/2} \sin(k\omega x) dx = 0. \end{aligned}$$

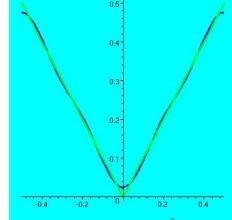
---

<sup>37</sup>Jean Baptiste Joseph Fourier, 1768-1830

**Satz 16.6.** Ist  $f(x)$  eine gerade Funktion, so ist  $b_k = 0$ ,  $k = 1, \dots, n$ , und ist  $f(x)$  eine ungerade Funktion so ist  $a_k = 0$ ,  $k = 0, \dots, n$ .

**Beispiel 16.7.** Das Fourierpolynom 3-ten Grades der Funktion  $f : [-1/2, 1/2] \rightarrow \mathbb{R}$  mit  $f(x) = |x|$  ist das Polynom (siehe (16.3))

$$F_3(x) = \frac{1}{4} - \frac{2}{\pi^2} \cos(2\pi x) - \frac{2}{9\pi^2} \cos(6\pi x).$$



$F_3(x)$

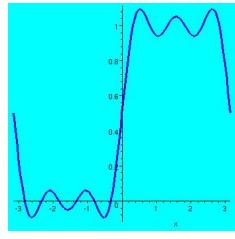
**Definition 16.8** (Fourierreihe). Sei  $T > 0$ ,  $\omega = 2\pi/T$ , und sei  $f : [-T/2, T/2] \rightarrow \mathbb{R}$  eine integrierbare Funktion mit Fourierkoeffizienten  $a_k$ ,  $b_k$ . Dann heißt

$$F_\infty(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega x) + b_k \sin(k\omega x))$$

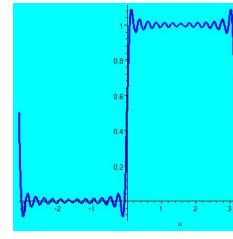
die *Fourierreihe* von  $f$ .

**Beispiel 16.9.** Sei  $f : [-\pi, \pi] \rightarrow \mathbb{R}$  mit  $f(t) = 1$  für  $t > 0$  und  $f(t) = 0$  für  $t \leq 0$ . Dann ist

$$F_\infty(x) = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{1 - (-1)^k}{k\pi} \sin(kx).$$



$F_5(x)$



$F_{25}(x)$

**Definition 16.10** (Skalarprodukt und Norm von Funktionen). Sei  $C[a, b]$  der Vektorraum der stetigen Funktionen auf dem Intervall  $[a, b] \subset \mathbb{R}$ . Für  $f, g \in C[a, b]$  sei

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx, \text{ und } \|f\| = \sqrt{\langle f, f \rangle}.$$

**Bemerkung 16.11.**  $\langle f, g \rangle$  ist ein Skalarprodukt auf  $C[a, b]$  (siehe (7.1)) und  $\|f\|$  eine Norm (siehe (7.4)).

**Satz 16.12.** Die Funktionen

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos(kx)}{\sqrt{\pi}}, \frac{\sin(kx)}{\sqrt{\pi}}, k \in \mathbb{N}$$

bilden ein Orthonormalsystem (siehe (7.13) und Satz 16.5) in  $C[-\pi, \pi]$ . Insbesondere gilt für  $g(x) = \frac{c_0}{2} + \sum_{k=1}^n (c_k \cos(kx) + d_k \sin(kx))$

$$\|g\|^2 = \pi \left( \frac{c_0^2}{2} + \sum_{k=1}^n (c_k^2 + d_k^2) \right).$$

**Satz 16.13** (Approximation durch Fourierpolynome). *Sei  $f \in C[-\pi, \pi]$  mit Fourierpolynom  $F_n$ , und sei  $T_n$  ein beliebiges trigonometrisches Polynom vom Grad  $n$ . Dann gilt*

- i)  $\|f - F_n\| \leq \|f - T_n\|$ .
- ii)  $\lim_{n \rightarrow \infty} \|f - F_n\|^2 = 0$  (*Konvergenz der Fourierreihe im quadratischen Mittel gegen  $f$* ).

## 17 Differenzierbarkeit II

**Definition 17.1** (Reellwertige Funktion mehrerer Variablen). Eine (reellwertige) *Funktion von mehreren Variablen* ist eine Abbildung

$$f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n),$$

wobei  $D$  eine Teilmenge des  $\mathbb{R}^n$  ist. Statt  $(x_1, \dots, x_n)$  wird auch oft  $\mathbf{x}$  geschrieben, und so  $f(\mathbf{x})$  anstelle von  $f(x_1, \dots, x_n)$ .

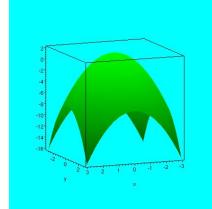
**Bemerkung 17.2.** Für eine Funktion  $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$  ist der Graph

$$\{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in D\}$$

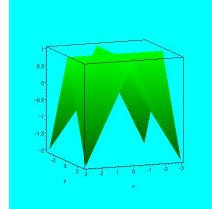
ein 3-dimensionales Gebilde (Gebirge). I.A. für  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  ein  $(n+1)$ -dimensionales Gebilde.

*Beispiel 17.3.*

i)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(\mathbf{x}) = f(x_1, x_2) = 2 - x_1^2 - x_2^2$ .



ii)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(\mathbf{x}) = f(x_1, x_2) = 1 - \min\{|x_1|, |x_2|\}$ .



**Definition 17.4** (Grenzwert einer Folge von Vektoren). Eine Folge  $(\mathbf{x}_k)_{k \in \mathbb{N}_0}$ ,  $\mathbf{x}_k \in \mathbb{R}^n$ , heißt *konvergent gegen  $\mathbf{x}^* \in \mathbb{R}^n$* , bezeichnet als  $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$ , falls der Abstand zwischen  $\mathbf{x}_k$  und  $\mathbf{x}^*$

$$\|\mathbf{x}_k - \mathbf{x}^*\| = \sqrt{(\mathbf{x}_{k,1} - \mathbf{x}_1^*)^2 + (\mathbf{x}_{k,2} - \mathbf{x}_2^*)^2 + \dots + (\mathbf{x}_{k,n} - \mathbf{x}_n^*)^2}$$

gegen 0 konvergiert, d.h.

$$\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^* \Leftrightarrow \lim_{k \rightarrow \infty} \|\mathbf{x}_k - \mathbf{x}^*\| = 0.$$

Man schreibt auch  $\mathbf{x}_k \rightarrow \mathbf{x}^*$ .

**Bemerkung 17.5.**  $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$  ist also gleichbedeutend damit, dass jede Koordinate  $\mathbf{x}_{k,i}$  gegen  $\mathbf{x}_i^*$  konvergiert,  $0 \leq i \leq n$ . Insbesondere übertragen sich die Eigenschaften von Folgen aus Kapitel ?? in „kanonischer Weise“ auf Folgen von Vektoren (s. z.B. Satz 11.12, Satz 11.15).

**Definition 17.6** (Grenzwert einer Funktion). Sei  $D \subseteq \mathbb{R}^n$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion, und sei  $\mathbf{x}^* \in \mathbb{R}^n$ . Wenn es ein  $y^* \in \mathbb{R}$  gibt, so dass für jede Folge  $(\mathbf{x}_k)_{k \in \mathbb{N}_0}$  mit  $\mathbf{x}_k \in D \setminus \{\mathbf{x}^*\}$  und

$$\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$$

gilt, dass die Folge der Funktionswerte  $(f(\mathbf{x}_k))_{k \in \mathbb{N}_0}$  gegen  $y^*$  konvergiert, also

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_k) = y^*,$$

dann heißt  $y^*$  der Grenzwert von  $f$  für  $\mathbf{x}$  gegen  $\mathbf{x}^*$ . Man schreibt dafür

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} f(\mathbf{x}) = y^*.$$

Hierbei ist auch  $y^* = \pm\infty$  erlaubt, und man spricht dann von bestimmter Divergenz (vgl. Def. 12.15).

**Definition 17.7** (Stetigkeit). Sei  $D \subseteq \mathbb{R}^n$ ,  $f : D \rightarrow \mathbb{R}$  eine Funktion, und sei  $\mathbf{x}^* \in D$ .  $f$  heißt stetig im Punkt (an der Stelle)  $\mathbf{x}^*$ , falls

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} f(\mathbf{x}) = f(\mathbf{x}^*),$$

d.h. der Grenzwert  $\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} f(\mathbf{x})$  existiert und ist gleich  $f(\mathbf{x}^*)$ . Dies bedeutet, für jede Folge  $(\mathbf{x}_k)_{k \in \mathbb{N}_0}$ ,  $\mathbf{x}_k \in D \setminus \{\mathbf{x}^*\}$ , mit  $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$  gilt

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_k) = f\left(\lim_{k \rightarrow \infty} \mathbf{x}_k\right) = f(\mathbf{x}^*).$$

Ist die Funktion  $f$  in jedem Punkt ihres Definitionsbereiches  $D$  stetig, dann heißt  $f$  stetig (vgl. Def. 12.24).

**Bemerkung 17.8.** Auch bei den Definitionen 17.6 und 17.7 für eine Funktion mehrerer Variablen lassen sich die bekannten Eigenschaften von Grenzwerten und Stetigkeit einer Funktion einer Veränderlichen (s. z.B. Satz 12.20 oder Bemerkung 12.25) in „kanonischer Weise“ verallgemeinern.

**Satz 17.9.** Seien  $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  Funktionen, die in  $\mathbf{x}^* \in \mathbb{R}^n$  stetig sind. Dann sind auch die Funktionen

$$\begin{aligned} & \alpha f(\mathbf{x}) \text{ mit } \alpha \in \mathbb{R}, \quad f(\mathbf{x}) \pm g(\mathbf{x}), \quad f(\mathbf{x}) \cdot g(\mathbf{x}), \\ & \frac{f(\mathbf{x})}{g(\mathbf{x})}, \text{ falls } g(\mathbf{x}^*) \neq 0, \end{aligned}$$

stetig in  $\mathbf{x}^*$ .

**Definition 17.10** (Vektorwertige Funktion). Sei  $D \subset \mathbb{R}^n$ . Unter einer (vektorwertigen) Funktion  $\mathbf{f} : D \rightarrow \mathbb{R}^m$  versteht man die Abbildung  $\mathbf{x} \mapsto (\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x}), \dots, \mathbf{f}_m(\mathbf{x}))$ , wobei  $\mathbf{f}_i : D \rightarrow \mathbb{R}$ ,  $1 \leq i \leq m$ , reellwertige Funktionen sind.

**Bemerkung 17.11.**

- i) Eine Matrix  $A \in \mathbb{R}^{m \times n}$  beschreibt eine lineare Abbildung von  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  (vgl. Satz 6.5).
- ii) Im Falle  $n = 1$  nennt man  $f$  auch (parameterisierte) Kurven.

**Definition 17.12** (Grenzwert einer Funktion). Sei  $D \subseteq \mathbb{R}^n$ ,  $\mathbf{f} : D \rightarrow \mathbb{R}^m$  eine Funktion, und sei  $\mathbf{x}^* \in \mathbb{R}^n$ . Wenn es ein  $\mathbf{y}^* \in \mathbb{R}^m$  gibt, so dass für jede Folge  $(\mathbf{x}_k)_{k \in \mathbb{N}_0}$  mit  $\mathbf{x}_k \in D \setminus \{\mathbf{x}^*\}$  und

$$\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$$

gilt, dass die Folge der Funktionswerte  $(\mathbf{f}(\mathbf{x}_k))_{k \in \mathbb{N}_0}$  gegen  $\mathbf{y}^*$  konvergiert, also

$$\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{x}_k) = \mathbf{y}^*,$$

dann heißt  $\mathbf{y}^*$  der Grenzwert von  $\mathbf{f}$  für  $\mathbf{x}$  gegen  $\mathbf{x}^*$ . Man schreibt dafür

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \mathbf{f}(\mathbf{x}) = \mathbf{y}^*.$$

Mit anderen Worten:  $\mathbf{y}^*$  ist der Grenzwert von  $\mathbf{f}$  für  $\mathbf{x}$  gegen  $\mathbf{x}^*$  genau dann, wenn  $\mathbf{y}_i^*$  der Grenzwert von  $\mathbf{f}_i$  für  $\mathbf{x}$  gegen  $\mathbf{x}^*$  und für  $1 \leq i \leq m$  ist (vgl. Def. 17.6).

**Definition 17.13** (Stetigkeit). Sei  $D \subseteq \mathbb{R}^n$ ,  $\mathbf{f} : D \rightarrow \mathbb{R}^m$  eine Funktion, und sei  $\mathbf{x}^* \in D$ .  $\mathbf{f}$  heißt *stetig im Punkt (an der Stelle)  $\mathbf{x}^*$* , falls

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^*),$$

d.h. der Grenzwert  $\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \mathbf{f}(\mathbf{x})$  existiert und ist gleich  $\mathbf{f}(\mathbf{x}^*)$ . Dies bedeutet, für jede Folge  $(\mathbf{x}_k)_{k \in \mathbb{N}_0}$ ,  $\mathbf{x}_k \in D \setminus \{\mathbf{x}^*\}$ , mit  $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$  gilt

$$\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{x}_k) = \mathbf{f}\left(\lim_{k \rightarrow \infty} \mathbf{x}_k\right) = \mathbf{f}(\mathbf{x}^*).$$

Ist die Funktion  $\mathbf{f}$  in jedem Punkt ihres Definitionsbereiches  $D$  stetig, dann heißt  $\mathbf{f}$  *stetig*. Mit anderen Worten:  $\mathbf{f}$  ist genau dann in  $\mathbf{x}^*$  stetig, wenn jede Funktion  $f_i : D \rightarrow \mathbb{R}$ ,  $1 \leq i \leq n$ , in  $\mathbf{x}^*$  stetig ist (vgl. Def. 17.7).

**Definition 17.14** ( $\varepsilon$ -Umgebung, offene Menge).

i) Für  $\mathbf{x}^* \in \mathbb{R}^n$  und  $\varepsilon > 0$  heißt

$$B(\mathbf{x}^*, \varepsilon) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}^*\| < \varepsilon\}$$

$\varepsilon$ -Umgebung von  $\mathbf{x}^*$ . Geometrisch ist dies die „offene Kugel mit Radius  $\varepsilon$  und Mittelpunkt  $\mathbf{x}^*$ “.

ii) Eine Menge  $D \subseteq \mathbb{R}^n$  heißt *offen*, falls es für alle  $\mathbf{x}^* \in D$  ein  $\varepsilon > 0$  gibt mit  $B(\mathbf{x}^*, \varepsilon) \subseteq D$ .

**Bemerkung 17.15.** Um nun (zunächst) für Funktionen  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  einen Ableitungsbegriff zu definieren, greifen wir die geometrische Interpretation im Falle  $n = 1$  (s. Bem. 13.3) auf. Anstatt die Funktion (bzw. ihren Graphen) durch eine Tangente zu approximieren, möchten wir nun  $f$  an einer Stelle  $\mathbf{x}^*$  durch eine Hyperebene approximieren (s. Bem. 5.9).

**Definition 17.16** (Partielle Ableitung(en)). Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D$  offen, eine Funktion und  $\mathbf{x}^* \in D$ .

i) Unter der *partiellen Ableitung*  $\frac{\partial f}{\partial x_i}(\mathbf{x}^*)$  von  $f$  nach  $x_i$  in einem Punkt  $\mathbf{x}^*$  versteht man die Ableitung von  $f$  nach der Variablen  $x_i$ , während die übrigen Variablen als Konstante betrachtet werden. Formell bedeutet dies, dass der Grenzwert

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = \lim_{h \rightarrow 0} \frac{f(\mathbf{x}^* + h \mathbf{e}_i) - f(\mathbf{x}^*)}{h}$$

existieren muss, wobei  $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^\top$  den  $i$ -ten Einheitsvektor bezeichnet (s. Def. 5.23).

ii) Ist  $f$  in jedem Punkt  $\mathbf{x}^* \in D$  partiell nach  $x_i$  differenzierbar, dann heißt  $f$  *partiell nach  $x_i$  differenzierbar*, und die Funktion

$$\frac{\partial f}{\partial x_i} : D \rightarrow \mathbb{R} \text{ mit } x \mapsto \frac{\partial f}{\partial x_i}(x)$$

heißt *partielle Ableitung von  $f$  nach  $x_i$* .

iii) Ist  $f$  partiell nach  $x_i$  differenzierbar für  $1 \leq i \leq n$ , dann heißt  $f$  *partiell differenzierbar*.

iv)  $f$  heißt *stetig partiell differenzierbar*, falls zusätzlich alle partiellen Ableitungen  $\frac{\partial f}{\partial x_i}$  stetig sind.

**Bemerkung 17.17.** Im Gegensatz zu Satz 13.5 sind partiell differenzierbare Funktionen nicht notwendigerweise stetig, z.B. ist

$$f(x) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

partiell differenzierbar in 0, aber nicht stetig in 0.

**Definition 17.18** (Gradient). Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D$  offen, eine partiell differenzierbare Funktion. Dann heißt der Vektor

$$\text{grad } f(\mathbf{x}) = \left( \frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)$$

Gradient von  $f$  im Punkt  $\mathbf{x}$ .

**Bemerkung 17.19.** Seien  $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D$  offen, partiell differenzierbare Funktionen. Dann gilt die Produktregel (vgl. Satz 13.6 iii))

$$\text{grad}(f \cdot g)(\mathbf{x}) = g(\mathbf{x}) \cdot \text{grad } f(\mathbf{x}) + f(\mathbf{x}) \cdot \text{grad } g(\mathbf{x}).$$

**Definition 17.20** (Höhere partielle Ableitungen). Sei  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D$  offen, eine partiell differenzierbare Funktion. Sind alle partiellen Ableitungen  $\frac{\partial f}{\partial x_i} : D \rightarrow \mathbb{R}$  selbst wieder partiell differenzierbar, so heißt  $f$  zweimal partiell differenzierbar, und man schreibt statt  $\frac{\partial^2 f}{\partial x_j \partial x_i}$

$$\frac{\partial^2 f}{\partial x_j \partial x_i}$$

für die partielle Ableitung nach  $x_j$  der partielle Ableitung nach  $x_i$ .

Entsprechend sind höhere ( $k$ -te) partielle Ableitungen definiert

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_1}}.$$

**Satz 17.21** (Schwarz<sup>38</sup>). Ist  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D$  offen, zweimal stetig partiell differenzierbar, dann gilt für  $1 \leq i, j \leq n$

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Entsprechende Aussagen gelten für höhere ( $k$ -te) partielle Ableitungen, wenn  $f$  eine  $k$ -mal stetig partiell differenzierbare Funktion ist.

**Definition 17.22** ((totale) Differenzierbarkeit). Sei  $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $D$  offen, eine Abbildung.  $\mathbf{f}$  heißt in  $\mathbf{x}^* \in D$  (total) differenzierbar, falls es eine Matrix  $A(\mathbf{x}^*) \in \mathbb{R}^{m \times n}$  gibt, so dass

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^*) + A(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + r(\mathbf{x})$$

mit

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \frac{r(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}^*\|} = 0.$$

$A(\mathbf{x}^*)$  heißt die Ableitung oder das Differential von  $\mathbf{f}$  in  $\mathbf{x}^*$ .

**Definition 17.23** (Jacobi<sup>39</sup>-Matrix). Sei  $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $D$  offen, eine partiell differenzierbare Abbildung, d.h. alle Koordinaten-Funktionen  $\mathbf{f}_i$  sind partiell differenzierbar, und sei  $\mathbf{x}^* \in D$ . Die  $(m \times n)$ -Matrix

$$J_{\mathbf{f}}(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial \mathbf{f}_1}{\partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial \mathbf{f}_1}{\partial x_n}(\mathbf{x}^*) \\ \frac{\partial \mathbf{f}_2}{\partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial \mathbf{f}_2}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \dots & \vdots \\ \frac{\partial \mathbf{f}_m}{\partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial \mathbf{f}_m}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}$$

heißt Jacobi-Matrix oder auch Funktional-Matrix.

<sup>38</sup>Hermann Amandus Schwarz, 1843-1921

<sup>39</sup>Carl Gustav Jacob Jacobi, 1804-1851

**Bemerkung 17.24.** Ist  $m = 1$ , dann ist  $J_f(\mathbf{x}^*) = \text{grad } f(\mathbf{x}^*)$ .

**Satz 17.25.** Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $D$  offen, und sei  $\mathbf{x}^* \in D$ .

- i) Ist  $f$  in  $\mathbf{x}^*$  differenzierbar mit Ableitung  $A(\mathbf{x}^*)$ . Dann ist  $f$  in  $\mathbf{x}^*$  stetig und es ist  $J_f(\mathbf{x}^*) = A(\mathbf{x}^*)$ .
- ii) Ist  $f_i$  in  $\mathbf{x}^*$  stetig partiell differenzierbar für  $1 \leq i \leq m$ , dann ist  $f$  in  $\mathbf{x}^*$  differenzierbar mit Ableitung  $J_f(\mathbf{x}^*)$ .

**Korollar 17.26.** Ist  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D$  offen, stetig partiell differenzierbar, dann ist  $f$  stetig.

**Satz 17.27** (Kettenregel (vgl. Satz 13.6 v))). Sei  $g : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $C$  offen, in  $x^* \in C$  differenzierbar, und sei  $f : D \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^k$ ,  $D$  offen, mit  $g(C) \subseteq D$ , und es sei  $f$  in  $g(\mathbf{x}^*)$  differenzierbar. Dann ist die Verknüpfung  $f \circ g : C \rightarrow \mathbb{R}^k$  in  $\mathbf{x}^*$  differenzierbar, und es gilt:

$$J_{f \circ g}(\mathbf{x}^*) = J_f(g(\mathbf{x}^*)) \cdot J_g(\mathbf{x}^*).$$

**Definition 17.28** (Lokale Extrema (vgl. Def. 13.17)). Sei  $D \subseteq \mathbb{R}^n$  und  $f : D \rightarrow \mathbb{R}$  eine Funktion.  $\mathbf{x}^* \in D$  heißt

- i) *lokales Maximum*, falls ein  $\epsilon > 0$  existiert, so dass

$$f(\mathbf{x}) \leq f(\mathbf{x}^*) \text{ für alle } \mathbf{x} \in B(\mathbf{x}^*, \epsilon) \cap D.$$

- ii) *lokales Minimum*, falls ein  $\epsilon > 0$  existiert, so dass

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) \text{ für alle } \mathbf{x} \in B(\mathbf{x}^*, \epsilon) \cap D.$$

- iii) *globales Maximum*, falls

$$f(\mathbf{x}) \leq f(\mathbf{x}^*) \text{ für alle } \mathbf{x} \in D.$$

- iv) *globales Minimum*, falls

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) \text{ für alle } \mathbf{x} \in D.$$

- v) Lokale/globale Minima oder Maxima werden als lokale/globale Extrema bzw. Extremwerte der Funktion bezeichnet.

**Satz 17.29** (vgl. Satz 13.18). Sei  $D \subseteq \mathbb{R}^n$  offen, und sei  $f : D \rightarrow \mathbb{R}$  partiell differenzierbar. Besitzt  $f$  in  $\mathbf{x}^* \in D$  ein lokales Extremum, dann ist  $\text{grad } f(\mathbf{x}^*) = \mathbf{0}$ .

**Definition 17.30** (Hesse<sup>40</sup>-Matrix). Sei  $D \subseteq \mathbb{R}^n$  offen, und sei  $f : D \rightarrow \mathbb{R}$  eine 2-mal stetig partiell differenzierbare Funktion. Für  $\mathbf{x}^* \in D$  heißt die symmetrische  $(n \times n)$ -Matrix

$$H_f(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}^*) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}^*) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^*) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}^*) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}^*) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}^*) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}^*) \end{pmatrix}$$

Hesse-Matrix von  $f$  in  $\mathbf{x}^*$ .

**Satz 17.31** ((vgl. Satz 13.18 ii), iii) und Satz 9.37)). Sei  $D \subset \mathbb{R}^n$  offen, und sei  $f : D \rightarrow \mathbb{R}$  eine 2-mal stetig partiell differenzierbare Funktion. Sei  $\mathbf{x}^* \in D$  mit  $\text{grad } f(\mathbf{x}^*) = \mathbf{0}$ .

---

<sup>40</sup>Otto Hesse, 1811-1874

- i) Ist  $H_f(\mathbf{x}^*)$  positiv definit, dann ist  $\mathbf{x}^*$  ein lokales Minimum.
- ii) Ist  $H_f(\mathbf{x}^*)$  negativ definit, dann ist  $\mathbf{x}^*$  ein lokales Maximum.
- iii) Ist  $H_f(\mathbf{x}^*)$  weder positiv noch negativ semi-definit (d.h. indefinit), dann ist  $\mathbf{x}^*$  kein lokales Extremum, sondern ein Sattelpunkt.
- iv) Ist  $H_f(\mathbf{x}^*)$  negativ semi-definit oder positiv-definit, dann ist keine Aussage möglich.

**Bemerkung 17.32.** Eine  $2 \times 2$ -Matrix  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist genau dann positiv definit wenn  $a_{11} > 0$  und  $\det A > 0$ . Sie ist genau dann negativ definit wenn  $a_{11} < 0$  und  $\det A > 0$ .

**Definition 17.33** (Abgeschlossene, kompakte Menge).

- i)  $X \subseteq \mathbb{R}^n$  heißt abgeschlossen, falls  $\mathbb{R}^n \setminus \{X\}$  offen ist.
- ii)  $X \subseteq \mathbb{R}^n$  heißt kompakt, falls  $X$  abgeschlossen und beschränkt (d.h. es gibt ein  $M > 0$  mit  $\|x\| \leq M$  für alle  $x \in X$ ) ist.

**Bemerkung 17.34.** Für  $n = 1$  kann man sich kompakte Mengen als die Vereinigung von endlich vielen abgeschlossenen Intervallen vorstellen.

**Satz 17.35** (vgl. Satz 12.31). Jede stetige Funktion  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  nimmt auf einer kompakten Menge  $D$  ihr Maximum und Minimum an, d.h. es gibt  $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in D$  mit

$$\begin{aligned} f(\tilde{\mathbf{x}}) &= \max\{f(\mathbf{x}) : \mathbf{x} \in D\}, \text{ und} \\ f(\hat{\mathbf{x}}) &= \min\{f(\mathbf{x}) : \mathbf{x} \in D\}. \end{aligned}$$

**Satz 17.36** (Lokale Extrema unter Nebenbedingungen). Sei  $D \subseteq \mathbb{R}^n$  offen, und seien  $f, g : D \rightarrow \mathbb{R}$  stetig partiell differenzierbare Funktionen. Sei  $M = \{x \in D : g(x) = 0\}$  und sei  $\mathbf{x}^* \in M$  mit  $\text{grad } g(\mathbf{x}^*) \neq \mathbf{0}$ . Ferner sei  $\mathbf{x}^*$  ein lokales Maximum (Minimum) von  $f$  unter der Nebenbedingung  $g(\mathbf{x}) = 0$ , d.h. es gibt ein  $\epsilon > 0$ , so dass gilt

$$f(\mathbf{x}) \leq f(\mathbf{x}^*) \quad (\text{bzw. } f(\mathbf{x}) \geq f(\mathbf{x}^*)) \quad \text{für alle } \mathbf{x} \in B(\mathbf{x}^*, \epsilon) \cap M.$$

Dann gibt es ein  $\lambda \in \mathbb{R}$  mit

$$\text{grad } f(\mathbf{x}^*) = \lambda \text{grad } g(\mathbf{x}^*).$$

Man nennt  $\lambda$  einen Lagrange<sup>41</sup>-Multiplikator.

---

<sup>41</sup>Joseph-Louis Lagrange, 1736-1813

## 18 Integralrechnung II

**Definition 18.1** (Quader). Für  $n \in \mathbb{N}$  und  $a_i, b_i \in \mathbb{R}$ ,  $a_i \leq b_i$ ,  $1 \leq i \leq n$ , heißt  $Q \subset \mathbb{R}^n$  mit

$$Q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] = \{\mathbf{x} \in \mathbb{R}^n : a_i \leq x_i \leq b_i, 1 \leq i \leq n\}.$$

Quader (mit Kantenlängen  $b_i - a_i$ ,  $1 \leq i \leq n$ ). Das Volumen  $\text{vol}(Q)$  des Quaders ist gegeben durch

$$\text{vol}(Q) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \cdots \cdot (b_n - a_n).$$

**Bemerkung 18.2.**

- i) Ist  $n = 1$ , so ist ein Quader ein abgeschlossenes Intervall.
- ii) Für  $n = 2$  ist ein Quader  $Q$  ein Rechteck, und  $\text{vol}(Q)$  ist der Flächeninhalt von  $Q$ .

**Definition 18.3** (Zerlegung im  $\mathbb{R}^n$ ). Sei  $Q \subset \mathbb{R}^n$  ein Quader. Eine Menge  $Z = \{Q_1, \dots, Q_m\}$  von Quadern  $Q_i \subseteq Q$  heißt Zerlegung von  $Q$ , falls

$$Q = \bigcup_{i=1}^m Q_i,$$

und je zwei verschiedene Quadern  $Q_i, Q_j$  überlappen sich nicht (vgl. Def. 15.1).

**Definition 18.4** (Ober-, Untersumme im  $\mathbb{R}^n$ ). Sei  $Q \subset \mathbb{R}^n$  ein Quader, und sei  $f : Q \rightarrow \mathbb{R}$  eine beschränkte Funktion. Sei  $Z = \{Q_1, \dots, Q_m\}$  eine Zerlegung von  $Q$ . Dann heißt

- i)  $U_f(Z) = \sum_{i=1}^m \text{vol}(Q_i) \cdot \inf\{f(x) : x \in Q_i\}$  Untersumme von  $f$  bzgl.  $Z$ .
- ii)  $O_f(Z) = \sum_{i=1}^m \text{vol}(Q_i) \cdot \sup\{f(x) : x \in Q_i\}$  Obersumme von  $f$  bzgl.  $Z$ .

(vgl. Def. 15.2)

**Definition 18.5** ((Riemann<sup>42</sup>)-Integral im  $\mathbb{R}^n$ ). Sei  $Q \subset \mathbb{R}^n$  ein Quader, und sei  $f : Q \rightarrow \mathbb{R}$  eine beschränkte Funktion.

- i)  $U_f = \sup\{U_f(Z) : Z$  Zerlegung von  $Q\}$  heißt (Riemann'sches) Unterintegral von  $f$  über  $Q$ .
- ii)  $O_f = \inf\{O_f(Z) : Z$  Zerlegung von  $Q\}$  heißt (Riemann'sches) Oberintegral von  $f$  über  $Q$ .
- iii)  $f$  heißt (Riemann-) integrierbar, falls  $U_f = O_f$ .
- iv) Ist  $f$  (Riemann-) integrierbar, dann heißt das Unterintegral  $U_f$  (bzw. Oberintegral  $O_f$ ) das (Riemann-) Integral von  $f(\mathbf{x})$  über  $Q$  und wird mit

$$\int_Q f(\mathbf{x}) \, d\mathbf{x}$$

bezeichnet. Abkürzend schreibt man auch nur  $\int_Q f$  (vgl. Def. 15.5).

**Bemerkung 18.6.** Im allgemeinen Fall geht man wie folgt vor: Sei  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  eine beschränkte Funktion auf der beschränkten Menge  $D \subset \mathbb{R}^n$ . Wähle einen Quader  $Q$  mit  $D \subseteq Q$  und betrachte

$$\tilde{f} : Q \rightarrow \mathbb{R} \text{ mit } \tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & : \mathbf{x} \in D, \\ 0 & : \mathbf{x} \in Q \setminus D. \end{cases}$$

Dann setzt man

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_Q \tilde{f}(\mathbf{x}) \, d\mathbf{x},$$

falls die rechte Seite existiert.

---

<sup>42</sup>Georg Friedrich Bernhard Riemann, 1826-1866

**Bemerkung 18.7.** Auch für diese Riemann-Integrale im  $\mathbb{R}^n$  gelten die üblichen Eigenschaften in „kanonischer Weise“ (vgl. Satz 15.7, Bem. 15.8, Lemma 15.9).

**Definition 18.8** (Volumen einer Menge im  $\mathbb{R}^n$ ). Sei  $D \subset \mathbb{R}^n$  beschränkt. Dann heißt (falls existent)

$$\text{vol}(D) = \int_D 1 \, d\mathbf{x}$$

das *Volumen der Menge D*. Man integriert also die konstante Funktion 1 über den Bereich.

**Satz 18.9.** Jede stetige Funktion ist integrierbar auf kompakten Mengen, die ein Volumen besitzen (z.B. Quadrate).

**Satz 18.10** (Fubini<sup>43</sup>).

i)  $n = 2$ : Sei  $Q = [a_1, b_1] \times [a_2, b_2]$  und  $f : Q \rightarrow \mathbb{R}$  integrierbar. Dann gilt

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \int_{a_2}^{b_2} \left( \int_{a_1}^{b_1} f(x_1, x_2) \, dx_1 \right) \, dx_2 = \int_{a_1}^{b_1} \left( \int_{a_2}^{b_2} f(x_1, x_2) \, dx_2 \right) \, dx_1.$$

ii)  $n \geq 2$ :  $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$  und  $f : Q \rightarrow \mathbb{R}$  integrierbar. Dann gilt

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \int_{a_n}^{b_n} \left( \int_{a_{n-1}}^{b_{n-1}} \left( \cdots \left( \int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n) \, dx_1 \right) \cdots \right) \, dx_{n-1} \right) \, dx_n,$$

wobei man auf der rechten Seite auch jede andere Reihenfolge beim Integrieren verwenden darf.

*Beispiel 18.11.* Sei  $f : Q = [0, 1] \times [0, 2] \rightarrow \mathbb{R}$  mit  $f(x_1, x_2) = (2 - x_1)x_2 + x_2$

$$\begin{aligned} \int_Q f(\mathbf{x}) \, d\mathbf{x} &= \int_0^1 \left( \int_0^2 f(x_1, x_2) \, dx_2 \right) \, dx_1 = \int_0^1 \left( \int_0^2 (2 - x_1)x_2 + x_2 \, dx_2 \right) \, dx_1 \\ &= \int_0^1 \left( 2x_2 - \frac{1}{2}x_1x_2^2 + \frac{1}{2}x_2^2 \right) \Big|_0^2 \, dx_1 = \int_0^1 (6 - 2x_1) \, dx_1 = (6x_1 - x_1^2) \Big|_0^1 = 5. \end{aligned}$$

**Definition 18.12** (Normalbereich). Ein 2-dimensionaler *Normalbereich* (bzgl. der  $x$ -Achse) ist eine Menge  $D \subset \mathbb{R}^2$  der Form

$$D = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : a \leq x_1 \leq b, \alpha_1(x_1) \leq x_2 \leq \beta_1(x_1) \right\};$$

dabei seien  $a < b$  und  $\alpha_1, \beta_1 : [a, b] \rightarrow \mathbb{R}$  stetige Funktionen. Entsprechend definiert man einen Normalbereich bzgl. der  $y$ -Achse, bzw. Normalbereiche in Dimensionen  $\geq 3$ . Z.B. ist

$$D = \left\{ (x_1, x_2, x_3)^T \in \mathbb{R}^3 : a \leq x_1 \leq b, \alpha_1(x_1) \leq x_2 \leq \beta_1(x_1), \alpha_2(x_1, x_2) \leq x_3 \leq \beta_2(x_1, x_2) \right\}$$

ein Normalbereich in  $\mathbb{R}^3$ . Hierbei sind zusätzlich  $\alpha_2, \beta_2 : U \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$  stetige Funktionen.

**Satz 18.13.** Sei  $D \subset \mathbb{R}^2$  ein Normalbereich, und sei  $f : D \rightarrow \mathbb{R}$  eine integrierbare Funktion auf  $D$ . Dann gilt:

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_a^b \left( \int_{\alpha_1(x_1)}^{\beta_1(x_1)} f(x_1, x_2) \, dx_2 \right) \, dx_1.$$

Eine analoge Aussage gilt für Normalbereiche in Dimension  $\geq 3$ . Mit der Notation aus Definition 18.12 gilt z.B. für  $f : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ :

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_a^b \left( \int_{\alpha_1(x_1)}^{\beta_1(x_1)} \left( \int_{\alpha_2(x_1, x_2)}^{\beta_2(x_1, x_2)} f(x_1, x_2, x_3) \, dx_3 \right) \, dx_2 \right) \, dx_1.$$

---

<sup>43</sup>Guido Fubini, 1879-1943

*Beispiel 18.14.* Sei  $T = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 - x_1\}$ , und sei  $f : T \rightarrow \mathbb{R}$  mit  $f(\mathbf{x}) = 2 - x_1^2 - x_2^2$ .

$$\begin{aligned}\int_T f(\mathbf{x}) d\mathbf{x} &= \int_0^1 \left( \int_0^{1-x_1} f(x_1, x_2) dx_2 \right) dx_1 = \int_0^1 \left( \int_0^{1-x_1} (2 - x_1^2 - x_2^2) dx_2 \right) dx_1 \\ &= \int_0^1 \left( 2x_2 - x_2 x_1^2 - \frac{1}{3} x_2^3 \right) \Big|_0^{1-x_1} dx_1 = \int_0^1 \left( 2(1-x_1) - (1-x_1)x_1^2 - \frac{1}{3}(1-x_1)^3 \right) dx_1 \\ &= \int_0^1 \left( \frac{5}{3} - x_1 - 2x_1^2 + \frac{4}{3}x_1^3 \right) dx_1 = \left( \frac{5}{3}x_1 - \frac{1}{2}x_1^2 - \frac{2}{3}x_1^3 + \frac{1}{3}x_1^4 \right) \Big|_0^1 = \frac{5}{6}.\end{aligned}$$

**Satz 18.15.** Sei  $f : D \rightarrow \mathbb{R}$  integrierbar. Sei  $A \in \mathbb{R}^{n \times n}$  mit  $\det A \neq 0$ , und sei  $t \in \mathbb{R}^n$ . Sei  $\tilde{D} \subset \mathbb{R}^n$  mit  $D = A\tilde{D} + t = \{Ax + t : x \in \tilde{D}\}$ . Dann gilt

$$\int_D f(\mathbf{x}) d\mathbf{x} = |\det A| \cdot \int_{\tilde{D}} f(A\mathbf{x} + t) d\mathbf{x}.$$

Insbesondere ist  $\text{vol}(D) = |\det A| \text{vol}(\tilde{D})$ .

**Korollar 18.16.** Sei  $\tilde{D} \subset \mathbb{R}^n$  beschränkt,  $\lambda \in \mathbb{R}$ ,  $t \in \mathbb{R}^n$  und

$$\lambda\tilde{D} + t = \{\lambda\mathbf{x} + t : \mathbf{x} \in \tilde{D}\}.$$

Dann gilt

$$\text{vol}(\lambda\tilde{D} + t) = |\lambda|^n \text{vol}(\tilde{D}).$$

**Satz 18.17** (Transformationssatz). Sei  $f : D \rightarrow \mathbb{R}$  Riemann-integrierbar. Sei  $\mathbf{T} : \tilde{D} \rightarrow D$  eine bijektive Abbildung von einer Menge  $\tilde{D} \subset \mathbb{R}^n$  nach  $D$ , die stetig partiell differenzierbar ist und deren Umkehrabbildung ebenfalls stetig partiell differenzierbar ist. Dann gilt

$$\int_D f(\mathbf{x}) d\mathbf{x} = \int_{\tilde{D}} f(\mathbf{T}(\mathbf{x})) \cdot |\det J_{\mathbf{T}}(\mathbf{x})| d\mathbf{x}.$$

**Bemerkung 18.18.**

- i) Für  $n = 1$  ist dies (im wesentlichen) die Substitutionsregel aus Satz 15.21.
- ii) Ist  $\mathbf{T}(\mathbf{x}) = A\mathbf{x} + t$  eine affine Abbildung, dann erhält man gerade Satz 18.15.

**Definition 18.19** (Polar(Kugel)- Zylinderkoordinaten).

- i) Jedes  $x \in \mathbb{R}^2$  lässt sich eindeutig darstellen als

$$x = \begin{pmatrix} r \cos \phi \\ r \sin \phi \end{pmatrix}$$

mit  $r \in \mathbb{R}_{\geq 0}$  und  $\phi \in [0, 2\pi]$ .

- ii) Jedes  $x \in \mathbb{R}^3$  lässt sich eindeutig darstellen als

$$\begin{pmatrix} r \cos \phi \sin \vartheta \\ r \sin \phi \sin \vartheta \\ r \cos \vartheta \end{pmatrix}$$

mit  $r \in \mathbb{R}_{\geq 0}$ ,  $\phi \in [0, 2\pi]$  und  $\vartheta \in [0, \pi]$ .

iii) Jedes  $x \in \mathbb{R}^3$  lässt sich eindeutig darstellen als

$$\begin{pmatrix} r \cos \phi \\ r \sin \phi \\ t \end{pmatrix}$$

mit  $r \in \mathbb{R}_{\geq 0}$ ,  $\phi \in [0, 2\pi)$  und  $t \in \mathbb{R}$ .

Die Darstellungen i) und ii) nennt man allgemein *Polarkoordinaten-Darstellungen*, ii) auch *Kugelkoordinaten-Darstellung* und iii) heißt *Zylinderkoordinaten-Darstellung* (siehe auch Definition 7.9).

*Beispiel 18.20.* Volumen der Kugel  $D = B(\mathbf{0}, 1) = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| \leq 1\}$ . Sei  $\tilde{D} = [0, 1] \times [0, 2\pi) \times [0, \pi]$  und  $\mathbf{T} : \tilde{D} \rightarrow B(\mathbf{0}, 1)$  mit

$$\mathbf{T}(r, \phi, \vartheta) = \begin{pmatrix} r \cos \phi \sin \vartheta \\ r \sin \phi \sin \vartheta \\ r \cos \vartheta \end{pmatrix}.$$

Dann ist  $|\det J_{\mathbf{T}}(\mathbf{x})| = r^2 \sin \vartheta$ , und so

$$\begin{aligned} \text{vol}(B(\mathbf{0}, 1)) &= \int_{B(\mathbf{0}, 1)} 1 \, d\mathbf{x} = \int_{\tilde{D}} 1 |\det J_{\mathbf{T}}(\mathbf{x})| \, d\mathbf{x} = \int_{\tilde{D}} |r^2 \sin \vartheta| \, d\mathbf{x} \\ &= \int_0^1 \left( \int_0^{2\pi} \left( \int_0^\pi |r^2 \sin \vartheta| \, d\vartheta \right) \, d\phi \right) \, dr. \\ &= \int_0^1 \left( \int_0^{2\pi} \left( r^2 (-\cos \vartheta) \Big|_0^\pi \right) \, d\phi \right) \, dr = \int_0^1 \left( \int_0^{2\pi} (2r^2) \, d\phi \right) \, dr = \int_0^1 (4\pi r^2) \, dr \\ &= \left( \frac{4}{3} \pi r^3 \right) \Big|_0^1 = \frac{4}{3} \pi. \end{aligned}$$

**Definition 18.21** (Vektorwertige Integrale). Sei  $\mathbf{f} : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ , und es sei jede einzelne Koordinatenfunktion  $f_i : D \rightarrow \mathbb{R}$  Riemann-integrierbar,  $1 \leq i \leq m$ . Dann ist

$$\int_D \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \begin{pmatrix} \int_D f_1(\mathbf{x}) \, d\mathbf{x} \\ \int_D f_2(\mathbf{x}) \, d\mathbf{x} \\ \vdots \\ \int_D f_m(\mathbf{x}) \, d\mathbf{x} \end{pmatrix}.$$

**Definition 18.22** (Schwerpunkt). Sei  $D \subset \mathbb{R}^n$  mit  $\text{vol}(D) > 0$ . Dann heißt der Punkt

$$\frac{1}{\text{vol}(D)} \int_D \mathbf{x} \, d\mathbf{x}$$

der Schwerpunkt von  $D$ .

*Beispiel 18.23.* Sei  $T = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 - x_1\}$ . Dann ist  $\text{vol}(T) = 1/2$ , und für die erste Koordinate des Schwerpunktes  $(x_1^*, x_2^*)$  ergibt sich

$$\begin{aligned} x_1^* &= \frac{1}{\frac{1}{2}} \int_T x_1 \, d\mathbf{x} = 2 \left( \int_0^1 \left( \int_0^{1-x_1} x_1 \, dx_2 \right) \, dx_1 \right) \\ &= 2 \left( \int_0^1 (x_1(1-x_1)) \, dx_1 \right) = 2 \left( \int_0^1 (x_1 - x_1^2) \, dx_1 \right) \\ &= 2 \left( \frac{1}{2}x_1^2 - \frac{1}{3}x_1^3 \right) \Big|_0^1 = \frac{1}{3}. \end{aligned}$$

Ebenso findet man für die 2.te Koordinate  $x_2^* = \frac{1}{3}$ .

# Teil III

# Mathematik III

## 19 Elementare Wahrscheinlichkeitsrechnung

**Definition 19.1** (Zufallsexperiment). Ein *Zufallsexperiment* ist ein Vorgang, der beliebig oft unter den gleichen Bedingungen wiederholt werden kann und dessen Ergebnis nicht mit Sicherheit vorhergesagt werden kann. Die Menge aller möglichen (und sich gegenseitig ausschließenden) Ergebnisse des Zufallsexperiments heißt *Ereignisraum* (*Ergebnisraum*) und wird mit  $\Omega$  bezeichnet.

*Beispiel 19.2.*

- i) Würfeln:  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .
- ii) Lotto:  $\Omega$  = Menge der 6-elementigen Teilmengen aus  $\{1, \dots, 49\}$ .
- iii) Münzwurf:  $\Omega = \{(K)\text{opf}, (Z)\text{ahl}\}$ .

**Definition 19.3** (Ereignis).

- i) Ein *Ereignis*  $A$  eines Zufallsexperiments ist eine Teilmenge des Ereignisraums  $\Omega$ , also  $A \subseteq \Omega$ . Man sagt, dass  $A$  *eingetreten* ist, wenn das Ergebnis des Experiments ein Element von  $A$  ist.
- ii) Ist  $|A| = 1$ , dann heißt  $A$  *Elementarereignis*.
- iii)  $A = \Omega$  heißt *sicheres Ereignis*.
- iv)  $A = \emptyset$  heißt *unmögliches Ereignis*.

*Beispiel 19.4.*

- i) Würfeln einer Primzahl:  $A = \{2, 3, 5\}$ .
- ii) Mindestens einmal Kopf beim zweimaligen Wurf einer Münze:  $\Omega = \{KK, KZ, ZK, ZZ\}$  und  $A = \{KK, KZ, ZK\}$ .

**Definition 19.5.** Seien  $A, B$  Ereignisse eines Zufallsexperiments mit Ereignisraum  $\Omega$ .

- i) Das Ereignis  $A$  und  $B$  entspricht dem Durchschnitt  $A \cap B$ .
- ii) Das Ereignis  $A$  oder  $B$  entspricht der Vereinigung  $A \cup B$ .
- iii) Das *Gegenereignis*  $\bar{A}$  von  $A$  entspricht dem Komplement von  $A$  bzgl.  $\Omega$ , d.h.  $\bar{A} = \Omega \setminus A$ .
- iv) Ist  $A \cap B = \emptyset$ , dann heißen  $A$  und  $B$  *disjunkt*.

**Definition 19.6** (Laplace<sup>44</sup>-Experiment). Ein Zufallsexperiment heißt *Laplace-Experiment*, falls es nur endlich viele mögliche Elementarereignisse gibt (also  $|\Omega| < \infty$ ), und falls jedes Elementarereignis *gleich wahrscheinlich* ist.

**Satz 19.7.** Die Wahrscheinlichkeit  $P(A)$  eines Ereignisses  $A$  bei einem Laplace-Experiment mit Ereignisraum  $\Omega$  ist gleich

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der Elementarereignisse in } A}{\text{Anzahl aller Elementarereignisse}}.$$

---

<sup>44</sup>Pierre-Simon Laplace, 1749-1827

**Bemerkung 19.8** (Urnenmodelle). Sei  $W$  eine Menge mit  $n$  Elementen (Urne mit unterscheidbaren, etwa nummerierten Kugeln). Sei  $k \in \mathbb{N}$ .

i) Die Menge  $W^k = \{(\omega_1, \dots, \omega_k) : \omega_i \in W\}$  hat  $n^k$  Elemente, d.h.,  $\#W^k = n^k$ .

– Mit Zurücklegen und mit Berücksichtigung der Reihenfolge:

Es gibt  $n^k$  Möglichkeiten,  $k$  Kugeln aus der Urne zu ziehen, wenn jede gezogene Kugel sofort wieder zurückgelegt wird und unterschiedliche Reihenfolgen derselben Kugelkombination als unterschiedliche Ereignisse gezählt werden.

ii) Die Menge  $W_s^k = \{(\omega_1, \dots, \omega_k) : \omega_i \in W \text{ und } \omega_1 \leq \omega_2 \leq \dots \leq \omega_k\}$  hat die Mächtigkeit

$$\binom{n+k-1}{k}.$$

– Mit Zurücklegen und ohne Berücksichtigung der Reihenfolge:

Es gibt  $\binom{n+k-1}{k}$  Möglichkeiten,  $k$  Kugeln aus der Urne zu ziehen, wenn jede gezogene Kugel sofort wieder zurückgelegt wird und unterschiedliche Reihenfolgen derselben Kugelkombination als ein Ereignisse gezählt werden.

iii) Die Menge  $\overline{W}^k = \{(\omega_1, \dots, \omega_k) : \omega_i \in W \text{ und } \omega_i \neq \omega_j, i \neq j\}$  hat die Mächtigkeit

$$\prod_{i=0}^{k-1} (n-i) = \frac{n!}{(n-k)!}.$$

– Ohne Zurücklegen und mit Berücksichtigung der Reihenfolge:

Es gibt  $\frac{n!}{(n-k)!}$  Möglichkeiten,  $k$  Kugeln aus der Urne zu ziehen, wenn jede gezogene Kugel nicht zurückgelegt wird und unterschiedliche Reihenfolgen derselben Kugelkombination als unterschiedliche Ereignisse gezählt werden.

iv) Die Menge  $\overline{W}_s^k = \{(\omega_1, \dots, \omega_k) : \omega_i \in W \text{ und } \omega_1 < \omega_2 < \dots < \omega_k\}$  hat die Mächtigkeit

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

– Ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge:

Es gibt  $\binom{n}{k}$  Möglichkeiten,  $k$  Kugeln aus der Urne zu ziehen, wenn jede gezogene Kugel nicht zurückgelegt wird und unterschiedliche Reihenfolgen derselben Kugelkombination als ein Ereignisse gezählt werden.

**Definition 19.9** (Wahrscheinlichkeit(saxiome), Kolmogorov<sup>45</sup>). Die Wahrscheinlichkeit  $P(A)$  eines Ereignisses  $A$  eines Zufallsexperiments mit Ereignisraum  $\Omega$  muss folgende Eigenschaften erfüllen:

i)  $0 \leq P(A) \leq 1$ .

ii)  $P(\Omega) = 1$ .

iii) Ist  $A \cap B = \emptyset$ , dann ist  $P(A \cup B) = P(A) + P(B)$ .

---

<sup>45</sup>Andrey Kolmogorov, 1903-1987

iv)  $\Omega$  zusammen mit einer Wahrscheinlichkeit  $P$  heißt *Wahrscheinlichkeitsraum* und wird mit  $(\Omega, P)$  bezeichnet.

**Satz 19.10.**

- i)  $P(\bar{A}) = 1 - P(A)$ , insbesondere ist  $P(\emptyset) = 0$ .
- ii)  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$  (*Additionsregel*).
- iii) Sind  $A_i$ ,  $1 \leq i \leq m$ , paarweise disjunkte Ereignisse, d.h. es gilt  $A_i \cap A_j = \emptyset$  für  $i \neq j$ , dann ist

$$P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_m) = \sum_{i=1}^m P(A_i).$$

- iv) Ist  $A \subseteq B$ , dann ist  $P(A) \leq P(B)$  (*Monotonität*).

**Bemerkung 19.11.** Sei  $(\Omega, P)$  Wahrscheinlichkeitsraum, und sei  $\Omega = \{\omega_1, \dots, \omega_n\}$  mit Elementarereignissen  $\omega_i$ . Dann gilt für  $A \subseteq \Omega$

$$P(A) = \sum_{\omega_i \in A} P(\omega_i).$$

**Definition 19.12** (Bedingte Wahrscheinlichkeit). Seien  $A, B$  Ereignisse eines Zufallsexperiments mit  $P(A) \neq 0$ . Die Wahrscheinlichkeit von  $B$  unter der Bedingung, dass  $A$  eingetreten ist, ist definiert als

$$P(B|A) = \frac{P(A \cap B)}{P(A)},$$

und heißt *bedingte Wahrscheinlichkeit*.

**Satz 19.13** (Multiplikationssatz). Seien  $A, B$  Ereignisse eines Zufallsexperiments mit  $P(A), P(B) \neq 0$ . Dann gilt

$$P(A \cap B) = P(A) P(B|A) = P(B) P(A|B).$$

**Satz 19.14** (Satz von der totalen Wahrscheinlichkeit). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und seien  $A_i \subseteq \Omega$ ,  $1 \leq i \leq m$ , paarweise disjunkte Ereignisse mit  $\Omega = \bigcup_{i=1}^m A_i$ . Dann gilt für  $B \subseteq \Omega$

$$P(B) = \sum_{i=1}^m P(A_i) \cdot P(B|A_i)$$

**Satz 19.15** (Formel von Bayes<sup>46</sup>). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und seien  $A_i \subseteq \Omega$ ,  $1 \leq i \leq m$ , paarweise disjunkte Ereignisse mit  $\Omega = \bigcup_{i=1}^m A_i$ . Dann gilt für  $B \subseteq \Omega$

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^m P(A_i) \cdot P(B|A_i)}.$$

**Definition 19.16** (Unabhängige Ereignisse). Zwei Ereignisse  $A, B$  eines Zufallsexperiments heißen *unabhängig*, falls

$$P(A \cap B) = P(A) \cdot P(B).$$

**Definition 19.17** (Zufallsvariable). Sei  $\Omega$  ein Ereignisraum. Eine Abbildung, die jedem Elementarereignis eine reelle Zahl zuordnet, d.h.  $X : \Omega \rightarrow \mathbb{R}$  mit  $\omega \mapsto X(\omega)$ , heißt *Zufallsvariable*.  $\{X(\omega) : \omega \in \Omega\}$  werden *Realisierungen* der Zufallsvariable genannt.

---

<sup>46</sup>Thomas Bayes, 1702-1761

**Definition 19.18** (Verteilung). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable. Die Abbildung

$$V : X(\Omega) \rightarrow [0, 1], \quad x \mapsto P(\{\omega \in \Omega : X(\omega) = x\})$$

heißt die *Verteilung der Zufallsvariablen X*.

**Bemerkung 19.19.** Im folgenden werden wir die abkürzenden Schreibweisen verwenden:

$$\begin{aligned} X = x &\text{ beschreibt die Menge } \{\omega \in \Omega : X(\omega) = x\} \\ X \leq x &\text{ beschreibt die Menge } \{\omega \in \Omega : X(\omega) \leq x\} \\ y \leq X \leq x &\text{ beschreibt die Menge } \{\omega \in \Omega : y \leq X(\omega) \leq x\} \end{aligned}$$

**Definition 19.20** (Verteilungsfunktion). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable. Die Abbildung

$$F : \mathbb{R} \rightarrow [0, 1], \quad x \mapsto P(X \leq x)$$

heißt die *Verteilungsfunktion von X*.

**Satz 19.21.** Sei  $F : \mathbb{R} \rightarrow [0, 1]$  die Verteilungsfunktion einer Zufallsvariablen  $X$ . Dann gilt:

i)  $F$  ist monoton steigend, und es gilt

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

ii)  $F$  ist rechtsseitig stetig, und in jedem Punkt existieren rechtsseitiger und linksseitiger Grenzwert:

$$\lim_{x \rightarrow x^*+} F(x) = F(x^*) = P(X \leq x^*), \quad \lim_{x \rightarrow x^*-} F(x) = P(X < x^*).$$

iii)

$$\lim_{x \rightarrow x^*+} F(x) - \lim_{x \rightarrow x^*-} F(x) = P(X \leq x^*) - P(X < x^*) = P(X = x^*).$$

**Definition 19.22** (Diskrete Zufallsvariable). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable. Ist der Wertebereich  $\{X(\omega) : \omega \in \Omega\}$  endlich oder abzählbar unendlich, dann heißt  $X$  *diskrete Zufallsvariable*.

**Bemerkung 19.23.** Sei  $X$  eine diskrete Zufallsvariable mit Wertebereich  $\{x_i : i \in \mathbb{N}\}$ . Dann gilt:

i)

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i),$$

und die Verteilungsfunktion ist eine „Treppenfunktion“.

$$\text{i)} \quad P(X \geq x) = \sum_{x_i \geq x} P(X = x_i) = 1 - P(X < x).$$

$$\text{iii)} \quad P(a \leq X \leq b) = \sum_{a \leq x_i \leq b} P(X = x_i) = P(X \leq b) - P(X < a).$$

**Definition 19.24** (Stetige Zufallsvariable, Dichtefunktion). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable mit Verteilungsfunktion  $F(x)$ . Gibt es eine integrierbare, nicht negative reelle Funktion  $w$ , so dass

$$F(x) = P(X \leq x) = \int_{-\infty}^x w(t) dt,$$

dann heißt  $X$  *stetige Zufallsvariable*, und  $w$  heißt *Dichte(funktion)* der Zufallsvariablen  $X$ .

**Bemerkung 19.25.** Sei  $X$  eine stetige Zufallsvariable mit Verteilungsfunktion  $F$  und Dichte  $w$ . Dann gilt:

- i)  $F$  ist stetig, und es gilt  $F'(x) = w(x)$ .
- ii)  $P(X \geq x) = 1 - P(X \leq x) = \int_x^\infty w(t) dt$ .
- iii)  $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) = \int_a^b w(t) dt$ .

**Satz 19.26.** Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable.  $X$  ist stetig genau dann, wenn  $P(X = x) = 0$  für alle  $x \in \mathbb{R}$ .

**Definition 19.27** (Unabhängige Zufallsvariablen). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und seien  $X_i : \Omega \rightarrow \mathbb{R}$  Zufallsvariablen,  $1 \leq i \leq n$ .  $X_1, \dots, X_n$  heißen unabhängig, falls

$$P((X_1 \leq x_1) \cap (X_2 \leq x_2) \cap \dots \cap (X_n \leq x_n)) = \prod_{i=1}^n P(X_i \leq x_i).$$

**Definition 19.28** (Erwartungswert). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable. Der Erwartungswert  $E(X)$  der Zufallsvariablen  $X$  ist definiert als:

- i) ist  $X$  diskret, dann

$$E(X) = \sum_{x \in X(\Omega)} x \cdot P(X = x).$$

- ii) ist  $X$  stetig, dann

$$E(X) = \int_{-\infty}^{\infty} x w(x) dx.$$

Oft wird der Erwartungswert auch mit  $\mu$  bezeichnet.

**Satz 19.29** (Linearität des Erwartungswerts). Seien  $X, Y$  Zufallsvariablen und sei  $\alpha \in \mathbb{R}$ . Dann gilt

$$E(X + \alpha Y) = E(X) + \alpha E(Y).$$

Allgemein gilt für Zufallsvariablen  $X_1, \dots, X_n$  und reelle Zahlen  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$

$$E\left(\sum_{i=1}^n \alpha_i X_i\right) = \sum_{i=1}^n \alpha_i E(X_i).$$

**Satz 19.30.** Seien  $X, Y$  unabhängige Zufallsvariablen. Dann gilt

$$E(X \cdot Y) = E(X) \cdot E(Y).$$

Allgemein gilt für unabhängige Zufallsvariablen  $X_1, \dots, X_n$

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i).$$

**Definition 19.31** (Varianz, Standardabweichung). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum, und sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable.

$$\text{Var}(X) = E((X - E(X))^2)$$

heißt die Varianz der Zufallsvariablen  $X$ . Die Wurzel aus der Varianz, also  $\sqrt{\text{Var}(X)}$  heißt Standardabweichung von  $X$ . Oft wird die Varianz auch mit  $\sigma^2$  bzw. die Standardabweichung mit  $\sigma$  bezeichnet.

**Bemerkung 19.32.**

i)

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

ii) Ist  $X$  diskret, dann gilt

$$\begin{aligned}\text{Var}(X) &= \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \cdot P(X = x) \\ &= \left( \sum_{x \in X(\Omega)} x^2 \cdot P(X = x) \right) - \mathbb{E}(X)^2.\end{aligned}$$

iii) Ist  $X$  stetig, dann gilt

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 w(x) dx = \left( \int_{-\infty}^{\infty} x^2 w(x) dx \right) - \mathbb{E}(X)^2.$$

iv) Sind  $a, b \in \mathbb{R}$ , dann ist  $\text{Var}(aX + b) = a^2\text{Var}(X)$ .

**Definition 19.33** (Kovarianz). Seien  $X, Y$  Zufallsvariablen.

$$\text{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y)$$

heisst die *Kovarianz* von  $X, Y$ .

**Satz 19.34.** Seien  $X, Y$  Zufallsvariablen. Dann gilt:

i)

$$\text{Var}(X + Y) = \text{Var}(X) + 2\text{Cov}(X, Y) + \text{Var}(Y).$$

ii) Für  $X, Y$  unabhängig gilt:

- a)  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ , d.h.  $\text{Cov}(X, Y) = 0$ .
- b)  $\text{Var}(X \cdot Y) = \mathbb{E}(X^2)\mathbb{E}(Y^2) - \mathbb{E}(X)^2\mathbb{E}(Y)^2$ .

**Definition 19.35** (Standardisierte Zufallsvariable). Ist  $X$  eine Zufallsvariable mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Dann heißt

$$X^* = \frac{X - \mu}{\sigma}$$

die *standardisierte (Zufallsvariable)* von  $X$ . Es gilt  $\mathbb{E}(X^*) = 0$  und  $\text{Var}(X^*) = 1$ .

**Satz 19.36** (Ungleichungen von Markov<sup>47</sup>/Tschebyscheff<sup>48</sup>). Sei  $(\Omega, P)$  ein Wahrscheinlichkeitsraum,  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable, und sei  $c > 0$ .

i)

$$P(|X| \geq c) \leq \frac{\mathbb{E}(|X|)}{c}.$$

ii)

$$P(|X - \mathbb{E}(X)| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

---

<sup>47</sup>Andrey Andreyevich Markov, 1856 – 1922

<sup>48</sup>Pafnuti Lwowitsch Tschebyschow, 1821–1894

**Definition 19.37** (identisch verteilt). Zwei oder mehr Zufallsvariablen heißen *identisch verteilt* (i.d.), wenn ihre Verteilungen gleich sind.

**Satz 19.38** ((Schwaches) Gesetz der großen Zahlen). *Seien  $X_1, \dots, X_n$  unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Dann ist auch das arithmetische Mittel  $\tilde{X} = \frac{1}{n} \sum_{i=1}^n X_i$  eine Zufallsvariable mit Erwartungswert  $\mu$  und Varianz  $\sigma^2/n$ , und für  $\epsilon > 0$  gilt*

$$\lim_{n \rightarrow \infty} P(|\tilde{X} - \mu| < \epsilon) = 1.$$

**Definition 19.39** (Gleichverteilung). Sei  $\Omega$  der endliche Ereignisraum eines Laplace-Experiments, und sei  $P(\omega) = 1/|\Omega|$  für  $\omega \in \Omega$ . Die zugehörige Verteilung einer (diskreten) Zufallsvariable  $X$  heißt *Gleichverteilung*, und es gilt

$$E(X) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} X(\omega), \quad \text{Var}(X) = \left( \sum_{\omega \in \Omega} X(\omega)^2 \frac{1}{|\Omega|} \right) - E(X)^2.$$

**Definition 19.40** (Hypergeometrische Verteilung). Gegeben sei eine Grundmenge  $N$ , von denen  $M$  Elemente eine „bestimmte Eigenschaft“ haben. Man entnimmt nun eine Stichprobe vom Umfang  $n$  ohne Zurücklegen. Sei  $X$  die Zufallsvariable  $X = \text{Anzahl der Elemente in der Stichprobe mit der „bestimmten Eigenschaft“}$ . Dann gilt

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}.$$

Weiterhin gilt

$$E(X) = n \frac{M}{N}, \quad \text{Var}(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}.$$

**Definition 19.41** (Binomialverteilung). Ein *Bernoulli*<sup>49</sup>-Experiment ist ein Zufallsexperiment, bei dem es nur zwei Ausgänge gibt: Ereignis  $A$  tritt ein oder nicht. Sei  $p$  die Wahrscheinlichkeit, mit der  $A$  eintritt. Wiederholt man ein Bernoulli-Experiment  $n$ -mal, so spricht man von einem *Bernoulli-Experiment der Länge  $n$* . Für ein Bernoulli-Experiment der Länge  $n$ , sei  $X$  die (diskrete) Zufallsvariable, die angibt, wie oft das Ereignis  $A$  eingetreten ist. Dann ist  $X$  *binomialverteilt* und es gilt für  $k \in \{0, \dots, n\}$

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Weiterhin gilt

$$E(X) = n p, \quad \text{Var}(X) = n p (1-p).$$

**Definition 19.42** (Poisson<sup>50</sup>-Verteilung). Sei  $\lambda > 0$ . Eine (diskrete) Zufallsvariable  $X$ , die jede natürliche Zahl  $k \in \mathbb{N}_0$  mit der Wahrscheinlichkeit

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

annimmt, heißt *Poissonverteilt*. Es gilt

$$E(X) = \lambda, \quad \text{Var}(X) = \lambda.$$

---

<sup>49</sup>Daniel Bernoulli, 1700 – 1782

<sup>50</sup>Siméon-Denis Poisson, 1781–1840

**Satz 19.43.** Sei  $X$  binomialverteilt bzgl. den Parametern  $n$  und  $p$ . Für  $n \rightarrow \infty$  und  $p \rightarrow 0$ , so dass  $\lambda = np$  konstant ist, gilt

$$\lim_{n \rightarrow \infty, p \rightarrow 0} P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

(Faustregel: für  $n \geq 50$  und  $p \leq 0.1$  kann die Binomialverteilung „gut“ durch die Poisson-Verteilung mit  $\lambda = np$  angenähert werden.

**Definition 19.44** ((Stetige) Gleichverteilung (Rechteckverteilung)). Eine (stetige) Zufallsvariable  $X$  heißt *gleichverteilt* auf dem Intervall  $[a, b]$ , wenn sie die Dichtefunktion

$$w(x) = \begin{cases} \frac{1}{b-a}, & x \in (a, b), \\ 0, & \text{sonst} \end{cases}$$

besitzt. Es gilt

$$E(X) = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}.$$

**Definition 19.45** (Exponentialverteilung). Sei  $\lambda > 0$ . Eine (stetige) Zufallsvariable  $X$  heißt *exponentialverteilt*, wenn sie die Dichtefunktion

$$w(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{sonst} \end{cases}$$

besitzt. Es gilt

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

**Definition 19.46** (Normalverteilung (Gaußverteilung)). Eine (stetige) Zufallsvariable  $X$  heißt *normalverteilt* mit den Parametern  $\mu, \sigma > 0$ , wenn sie die Dichtefunktion

$$w(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}$$

besitzt. Kurzschreibweise  $X \sim N(\mu; \sigma^2)$ . Der Graph der Dichtefunktion heißt *Gauß'sche Glockenkurve*. Es ist

$$E(X) = \mu, \quad \text{Var}(X) = \sigma^2.$$

Speziell für  $\mu = 0, \sigma = 1$  heißt  $N(0; 1)$  *Standardnormalverteilung*. Hier schreibt man für die Dichtefunktion  $\varphi$  und für die Verteilungsfunktion  $\Phi$ .

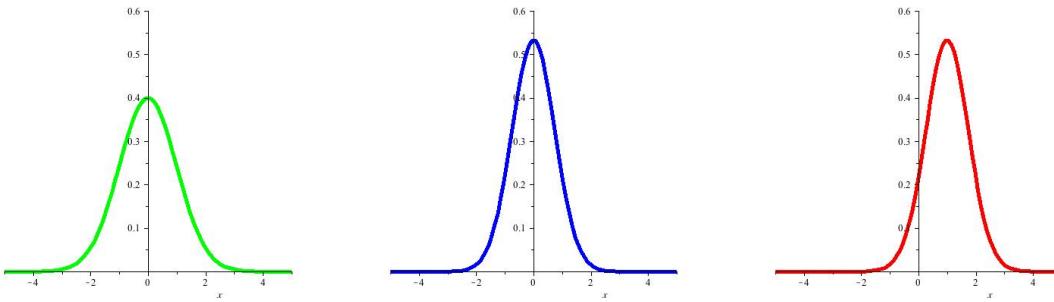


Abbildung 7: Normalverteilung mit  $\mu = 0, \sigma = 1$ ;  $\mu = 0, \sigma = 0.75$ ;  $\mu = 1, \sigma = 0.75$ ;

**Bemerkung 19.47.** Es ist  $\Phi(-x) = 1 - \Phi(x)$ .

**Bemerkung 19.48.** Ist  $X$  nach  $N(\mu; \sigma^2)$  verteilt, dann ist die Zufallsvariable

$$Z = \frac{X - \mu}{\sigma}$$

standardnormalverteilt. Ist  $F(x)$  die Verteilungsfunktion und  $w(x)$  die Dichtefunktion von  $X$ , dann gilt

$$F(x) = \Phi((x - \mu)/\sigma) \text{ und } w(x) = \frac{1}{\sigma} \varphi((x - \mu)/\sigma).$$

**Satz 19.49.** Seien  $X_i, i = 1, 2$ , unabhängige normalverteilte Zufallsvariablen mit Erwartungswerten  $\mu_i$  und Varianzen  $\sigma_i^2$ . Dann ist  $X_1 + X_2$  ebenfalls normalverteilt mit Erwartungswert  $\mu_1 + \mu_2$  und Varianz  $\sigma_1^2 + \sigma_2^2$ .

Beachte: Zum Beispiel ist die Summe von unabhängigen gleichverteilten Zufallsvariablen i.A. nicht gleichverteilt.

**Satz 19.50** (Zentraler Grenzwertsatz). Seien  $X_1, \dots, X_n$  unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Weiterhin sei

$$\tilde{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

das arithmetische Mittel der Zufallsvariablen. Für  $n \rightarrow \infty$  ist  $\tilde{X}$  nach  $N(\mu; \sigma^2/n)$  verteilt. Genauer: Die standardisierte Zufallsvariable  $(\tilde{X} - \mu)/(\sigma/\sqrt{n})$  ist für  $n \rightarrow \infty$  standardnormalverteilt, d.h.

$$\lim_{n \rightarrow \infty} P \left( \frac{\tilde{X} - \mu}{\sigma/\sqrt{n}} \leq z \right) = \Phi(z).$$

**Bemerkung 19.51.**

- i) Sei  $X$  binomialverteilt mit den Parametern  $n$  und  $p$ , und Verteilungsfunktion  $F_B$ . Seien  $np$  und  $n(1-p)$  groß (Faustregel:  $np(1-p) \geq 9$ ). Dann ist

$$F_B(x) \approx \Phi \left( \frac{x + 1/2 - np}{\sqrt{np(1-p)}} \right)$$

- ii) Sei  $X$  poissonverteilt mit den Parametern  $\lambda$  und Verteilungsfunktion  $F_P$ . Sei  $\lambda$  groß (Faustregel:  $\lambda \geq 9$ ). Dann ist

$$F_P(x) \approx \Phi \left( \frac{x + 1/2 - \lambda}{\sqrt{\lambda}} \right)$$

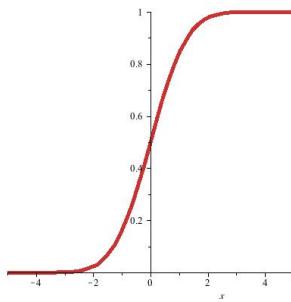


Abbildung 8: Verteilungsfunktion  $\Phi$  der Standardnormalverteilung

## 20 Elementare Statistik

**Definition 20.1** ( $p$ -Quantile, Median). Sei  $X$  eine Zufallsvariable mit Verteilungsfunktion  $F$ . Für  $p \in (0, 1)$  heißt  $x_p \in \mathbb{R}$  mit

$$F(x_p) = p$$

$p$ -Quantil von  $X$ . Ein Quantil zu  $p = 1/2$  heißt *Median*.

**Definition 20.2** (Zufallsstichprobe). Eine *Zufallsstichprobe* vom Umfang  $n$  ist eine Folge  $X_1, \dots, X_n$  von unabhängigen, identisch verteilten Zufallsvariablen. Die  $X_i$  heißen auch *Stichprobenvariablen*.

*Beispiel 20.3.*  $X$ =Abfüllgewicht einer Tafel Schokolade; In einer Stichprobe vom Umfang 10, könnten  $X_1, \dots, X_{10}$  etwa die Werte annehmen: 100, 97, 95, 103, 101, 105, 98, 99, 99, 91.

**Bemerkung 20.4.** Im allgemeinen werden Stichproben „ohne Zurücklegen gezogen“. Wir nehmen aber an, dass das Ziehen eines einzelnen Elementes die Grundmenge nur geringfügig verändert, d.h. wir können annehmen, dass die  $X_i$  unabhängig voneinander sind.

**Bemerkung 20.5** (Schätzungen). Oft möchte man von einer gegebenen Grundmenge (z.B. Menge aller an einem Tag produzierten Schokoladen) einen gewissen Parameter (Größe)  $\Theta$  wissen (z.B. durchschnittliche Abfüllgewicht, also  $E(X)$ ). Da die genaue Bestimmung zu aufwendig ist, versuchen wir  $\Theta$  mit Hilfe einer Stichprobe zu schätzen (z.B. arithmetisches Mittel einer Stichprobe).

**Definition 20.6** (Schätzfunktion). Eine Funktion  $T(X_1, \dots, X_n)$  der Stichprobenvariablen heißt *Stichprobenfunktion*, und ist selbst eine Zufallsvariable. Wird sie zum Schätzen eines Parameters  $\Theta$  verwendet, nennt man sie *Schätzfunktion* (oder *Schätzer*).

*Beispiel 20.7.*

- i)  $\Theta =$  Erwartungswert eines Merkmals  $X$ ; Schätzfunktion:  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  arithmetisches Mittel einer Stichprobe
- ii)  $\Theta =$  Varianz eines Merkmals  $X$ ; Schätzfunktion:  $\bar{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  Varianz einer Stichprobe (empirische Varianz).
- iii)  $\Theta =$  Anteil eines Merkmals  $X$ ; Schätzfunktion:  $\bar{P} =$  Anteil des Merkmals in einer Stichprobe (empirischer Anteil).

**Definition 20.8** ( $\chi^2$ -Verteilung). Seien  $Z_1, \dots, Z_m$  unabhängige und standardnormalverteilte Zufallsvariablen. Dann heißt die Verteilung der Zufallsvariablen

$$X = Z_1^2 + Z_2^2 + \dots + Z_m^2$$

$\chi^2$ -Verteilung mit  $m$  Freiheitsgraden.

Es ist  $E(X) = m$ ,  $\text{Var}(X) = 2m$ , und die Dichtefunktion ist gegeben durch

$$w(x) = \frac{1}{2^{m/2}\Gamma(m/2)} x^{\frac{m}{2}-1} e^{-x/2}.$$

**Bemerkung 20.9.** Die Gammafunktion  $\Gamma : \mathbb{R}_{>0} \rightarrow \mathbb{R}$  ist definiert als

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$

Insbesondere ist  $\Gamma(n) = (n-1)!$  für  $n \in \mathbb{N}$ .

**Satz 20.10.** Sei  $\bar{X}$  das arithmetische Mittel einer zufälligen Stichprobe  $X_1, \dots, X_n$  vom Umfang  $n$  eines normalverteilten Merkmals  $X$  mit Standardabweichung  $\sigma$ . Dann ist

$$\left(\frac{n-1}{\sigma^2}\right) \bar{S}^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i^2 - \bar{X}^2)$$

eine  $\chi^2$ -verteilte Zufallsvariable mit  $n-1$  Freiheitsgraden.

**Definition 20.11** ( $t$ -Verteilung). Seien  $X, Z$  zwei unabhängige Zufallsvariablen, wobei  $X$   $\chi^2$ -verteilt sei mit  $m$  Freiheitsgraden und  $Z$  sei standardnormalverteilt. Dann heißt die Verteilung der Zufallsvariablen

$$Y = \frac{Z}{\sqrt{X/m}}$$

$t$ -Verteilung oder auch Student-Verteilung mit  $m$  Freiheitsgraden.

Es ist  $E(Y) = 0$  ( $m > 1$ ),  $\text{Var}(Y) = m/(m-2)$  ( $m > 2$ ), und die Dichtefunktion ist gegeben durch

$$w(x) = \frac{\Gamma((m+1)/2)}{\sqrt{m\pi}\Gamma(m/2)} \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}.$$

**Satz 20.12.** Sei  $\bar{S}$  die Standardabweichung und  $\bar{X}$  das arithmetische Mittel einer zufälligen Stichprobe  $X_1, \dots, X_n$  vom Umfang  $n$  eines normalverteilten Merkmals  $X$  mit Erwartungswert  $\mu = E(X)$ . Dann besitzt die Zufallsvariable

$$Y = \frac{\bar{X} - \mu}{\bar{S}/\sqrt{n}}$$

eine  $t$ -Verteilung mit  $n-1$  Freiheitsgraden.

**Definition 20.13** (Erwartungstreue). Sei  $T$  eine Schätzfunktion des Parameters  $\Theta$ .

- i)  $T$  heißt erwartungstreu, falls  $E(T) = \Theta$ .
- ii)  $T$  heißt konsistent, falls  $\lim_{n \rightarrow \infty} P(|T - \Theta| < \epsilon) = 1$  für beliebiges  $\epsilon > 0$ .
- iii)  $T$  heißt konsistent im quadratischen Mittel, wenn  $\lim_{n \rightarrow \infty} E((T - \Theta)^2) = 0$ .

**Satz 20.14.** Ist eine Schätzfunktion konsistent im quadratischen Mittel, so ist sie auch konsistent.

**Bemerkung 20.15.**  $\bar{X}$  (arithmetisches Mittel) und  $\bar{P}$  (empirischer Anteil) sind erwartungstreu und konsistent im quadratischen Mittel.  $\bar{S}^2$  (empirische Varianz) ist erwartungstreu und konsistent.

### • Intervallschätzungen

**Definition 20.16** (Konfidenzintervall). Seien  $g_u(X_1, \dots, X_n)$  und  $g_o(X_1, \dots, X_N)$  Stichprobenfunktionen mit  $g_u(X_1, \dots, X_n) \leq g_o(X_1, \dots, X_N)$ , und sei  $\alpha \in [0, 1]$ . Ist

$$P(\Theta \in [g_u(X_1, \dots, X_n), g_o(X_1, \dots, X_N)]) = 1 - \alpha,$$

dann heißt  $[g_u(X_1, \dots, X_n), g_o(X_1, \dots, X_N)]$  Konfidenzintervall (Vertrauensintervall) zum (Konfidenz)niveau (Vertrauensniveau, Sicherheit)  $1 - \alpha$ . Der Wert  $\alpha$  heißt auch Irrtumswahrscheinlichkeit.

**Bemerkung 20.17.** Üblicherweise wählt man für  $\alpha$  Werte wie z.B. 0.1, 0.05, 0.01.

**Satz 20.18** (Konfidenzintervall für  $E(X)$  eines normal verteilten Merkmals  $X$  bei bekanntem  $\sigma$ ).  $E(X)$  wird mittels des arithmetischen Mittels  $\bar{X}$  einer zufälligen Stichprobe geschätzt.

- i) Wähle eine Konfidenzniveau  $1 - \alpha$  (z.B. 0.90, 0.95 oder 0.99).
- ii) Bestimme das  $1 - \alpha/2$ -Quantil  $z_{1-\alpha/2}$  der Standardnormalverteilung.
- iii) Das Konfidenzintervall zum Niveau  $1 - \alpha$  ist gegeben durch

$$\left[ \bar{X} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right],$$

d.h., für gegebenes  $n$  und  $\sigma$  liegt  $E(X)$  mit Wahrscheinlichkeit  $1 - \alpha$  in dem obigen Intervall einer zufälligen Stichprobe, d.h.  $100\% \cdot (1 - \alpha)$  solcher Stichprobenintervallen enthalten  $E(X)$ .

- iv) Ersetze  $\bar{X}$  durch die realisierte Stichprobe  $\bar{x}$ , und  $n, \sigma$  durch die gegebenen Werte.

**Satz 20.19** (Konfidenzintervall für  $\sigma^2$  eines normal verteilten Merkmals  $X$ ).

$\sigma^2$  wird mittels der empirischen Varianz  $\bar{S}^2$  einer zufälligen Stichprobe geschätzt.

- i) Wähle eine Konfidenzniveau  $1 - \alpha$  (z.B. 0.90, 0.95 oder 0.99).
- ii) Bestimme das  $1 - \alpha/2$  und  $\alpha/2$ -Quantil  $z_{1-\alpha/2}, z_{\alpha/2}$  der  $\chi^2$ -Verteilung mit  $n - 1$  Freiheitsgraden.
- iii) Das Konfidenzintervall der Varianz  $\sigma^2$  zum Niveau  $1 - \alpha$  ist gegeben durch

$$\left[ \frac{(n-1)\bar{S}^2}{z_{1-\alpha/2}}, \frac{(n-1)\bar{S}^2}{z_{\alpha/2}} \right].$$

- iv) Das Konfidenzintervall der Standardabweichung  $\sigma$  zum Niveau  $1 - \alpha$  ist gegeben durch

$$\left[ \sqrt{\frac{(n-1)\bar{S}^2}{z_{1-\alpha/2}}}, \sqrt{\frac{(n-1)\bar{S}^2}{z_{\alpha/2}}} \right].$$

- v) Ersetze in den obigen Formeln  $\bar{S}^2$  durch die realisierte empirische Varianz der Stichprobe  $\bar{s}^2$  vom Umfang  $n$ .

**Satz 20.20** (Konfidenzintervall für einen Anteilswert  $p$  bei großem Stichprobenumfang  $n \geq 20$ ).

$p$  wird mittels der empirischen Anteil  $\bar{P}$  einer zufälligen Stichprobe geschätzt.

- i) Wähle eine Konfidenzniveau  $1 - \alpha$  (z.B. 0.90, 0.95 oder 0.99).
- ii) Bestimme das  $1 - \alpha/2$ -Quantil  $z_{1-\alpha/2}$  der Standardnormalverteilung.
- iii) Sei

$$g_{\pm} = \frac{n}{n + (z_{1-\alpha/2})^2} \left( \bar{P} + \frac{(z_{1-\alpha/2})^2}{2n} \pm (z_{1-\alpha/2})^2 \sqrt{\frac{\bar{P}(1-\bar{P})}{n} + \frac{(z_{1-\alpha/2})^2}{4n^2}} \right).$$

Dann ist das Konfidenzintervall zum Niveau  $1 - \alpha$  ist gegeben durch  $[g_-, g_+]$ .

- iv) Ersetze in der obigen Formel  $\bar{P}$  durch den realisierten empirischen Anteilswert  $\bar{p}$  der Stichprobe.
- v) Für  $n\bar{p}(1-\bar{p}) \geq 9$  können die Grenzen ersetzt werden durch

$$\left[ \bar{p} - (z_{1-\alpha/2})^2 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}, \bar{p} + (z_{1-\alpha/2})^2 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \right].$$

**Definition 20.21** (Hypothesen, Test).

- i) Ein statistisches Testproblem besteht aus einer *Nullhypothese*  $H_0$  und einer *Alternativhypothese*  $H_1$ , die sich gegenseitig ausschließen.
- ii) Enthalten die Hypothesen Aussagen über einen Parameter  $\Theta$ , so spricht man von einem *parametrischen Test*.
- iii) Eine Hypothese heißt *einfach*, wenn sie nur aus einem Parameterwert besteht, z.B.  $\Theta = 100$ . Ansonsten heißt sie *zusammengesetzt*, z.B.  $\Theta \neq 100$ .
- iv) Der Test basiert auf einer Zufallsstichprobe und einer Stichprobenfunktion  $T(X_1, \dots, X_n)$ , deren Verteilung unter Annahme der Nullhypothese bekannt ist.
- v) Der *Ablehnungsbereich* (oder *Verwerfungsbereich*) umfasst jene Werte von  $T(X_1, \dots, X_n)$ , die für die Alternative  $H_1$  sprechen und mit einer Wahrscheinlichkeit kleiner oder gleich  $\alpha$  auftreten.
- vi)  $\alpha$  heißt *Signifikanzniveau* des Tests. Typische Werte sind 0.1, 0.05, 0.01.
- vii) Wenn  $T(X_1, \dots, X_n)$  in den Ablehnungsbereich fällt, dann lautet die Testentscheidung „ $H_0$  zugunsten von  $H_1$  verwerfen“, ansonsten „ $H_0$  beibehalten“.

*Beispiel 20.22.*

- i) Bei der Produktion von Schokoladentafeln möchten wir durch Stichproben feststellen, ob das „erwartete“ Gewicht  $\mu$  einer produzierten Tafel dem Sollwert  $\mu_0 = 100$  Gramm entspricht.
- ii) Unsere Nullhypothese  $H_0$  ist:  $\mu = \mu_0$ . Unsere Alternativhypothese lautet  $H_1 : \mu \neq \mu_0$ .  $H_0$  ist einfach und  $H_1$  ist zusammengesetzt.
- iii) Wir ziehen nun eine Stichprobe von 10 Schokoladentafeln und erhalten ein arithmetisches Mittel  $\bar{x}$  von 98.9 Gramm. Sollen wir aufgrund der Stichprobe „ $H_0$  zugunsten von  $H_1$  verwerfen“ oder „ $H_0$  beibehalten“?
- iv) Dazu bestimmen wir: Wie groß muss die Abweichung  $c$  eines zufälligen Stichprobenmittels  $\bar{X}$  vom Erwartungswert  $\mu_0$  sein, damit es „extrem unwahrscheinlich“ ist, dass das Stichprobenmittel unter der Annahme der Gültigkeit von  $H_0$  zustande gekommen ist.
- v) „Extrem unwahrscheinlich“ wird mit dem Signifikanzniveau  $\alpha$  (hier  $\alpha = 0,05$ ) gemessen, und  $c$  ergibt sich dann aus

$$P(|\bar{X} - \mu_0| > c) = \alpha.$$

- vi) Wir nehmen an, dass  $X$  (Abfüllgewicht einer Schokolade) normalverteilt ist, mit Standardabweichung  $\sigma$ . Somit ist  $\bar{X}$  normalverteilt mit Standardabweichung  $\sigma/\sqrt{n}$ , und es ist

$$P(|\bar{X} - \mu_0| > c) = \alpha \Leftrightarrow P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq \frac{c}{\sigma/\sqrt{n}}\right) = 1 - \alpha/2.$$

- vii) Nun ist  $\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$  standardnormalverteilt, und wir können das entsprechende  $1 - \alpha/2$ -Quantil  $z_{1-\alpha/2}$  aus einer Tabelle ablesen.

viii) In dem Beispiel mit  $\sigma = 2$ ,  $\alpha = 0,05$ ,  $z_{0,975} = 1,96$  ergibt sich

$$c = 1.96 \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{2}{\sqrt{10}} = 1.24.$$

ix) Wenn also ein Stichprobenmittel  $\bar{x}$  mit

$$\bar{x} < \mu_0 - c = 98.8 \text{ oder } \bar{x} > \mu_0 + c = 101.24$$

gezogen wird, dann verwerfen wir  $H_0$  zugunsten von  $H_1$ . Ansonsten behalten wir  $H_0$ .

x) Hier ist  $\bar{x} = 98.9$  und wir behalten  $H_0$ .

**Satz 20.23** (Test für  $E(X)$  eines normalverteilten Merkmals bei bekannten  $\sigma$  (Gauß-Test)).

i) Formuliere die Hypothesen ( $\mu = E(X)$ ):

- a)  $H_0 : \mu = \mu_0$  gegen  $H_1 : \mu \neq \mu_0$ , oder
- b)  $H_0 : \mu \geq \mu_0$  gegen  $H_1 : \mu < \mu_0$ , oder
- c)  $H_0 : \mu \leq \mu_0$  gegen  $H_1 : \mu > \mu_0$

ii) Wähle das Signifikanzniveau  $\alpha$ .

iii) Ziehe eine Stichprobe vom Umfang  $n$ , berechne das arithmetische Mittel  $\bar{x}$  und den standardisierten Prüfwert

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}.$$

iv) Bzgl. der Standardnormalverteilung berechne die Quantile  $z_{1-\alpha/2}$  und  $z_{1-\alpha}$ .

v) Entscheidung:

- a)  $H_0$  ist zu verwerfen, falls  $|z| > z_{1-\alpha/2}$ .
- b)  $H_0$  ist zu verwerfen, falls  $z < -z_{1-\alpha}$ .
- c)  $H_0$  ist zu verwerfen, falls  $z > z_{1-\alpha}$ .

**Bemerkung 20.24.** Der Gauß-Test entspricht dem Vorgehen beim Berechnen von Konfidenzintervallen 20.18. Zum Beispiel verwerfen wir in a)  $H_0$  genau dann, falls  $\mu$  nicht in dem entsprechenden Konfidenzintervall zum Niveau  $1 - \alpha$  liegt. Analog können die Ergebnisse über Konfidenzintervalle in den Sätzen 20.19 und 20.20 in Form von Hypothesentests formuliert werden.

**Definition 20.25** (Empirische Verteilung, Empirischer Median). Seien  $x_1, \dots, x_n$  Realisierung einer Zufallsstichprobe eines Merkmals  $M$  (bzw. eine Sammlung von Daten).

i)

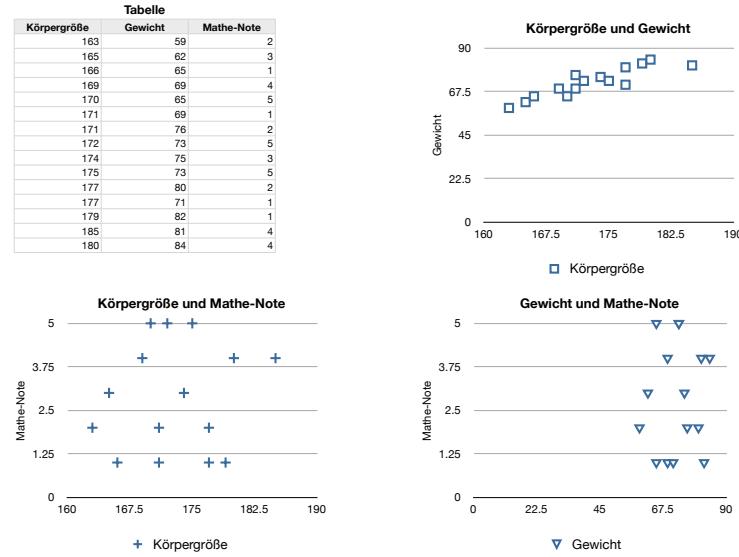
$$F : \mathbb{R} \rightarrow \mathbb{R} \quad \text{mit } F(x) = \frac{1}{n} \#\{x_i : x_i \leq x\}$$

heisst die *empirische Verteilungsfunktion* von  $M$ .

ii) Seien  $x_i$  der Größe nach geordnet, d.h., wir nehmen an  $x_1 \leq x_2 \leq \dots \leq x_n$ . Als *empirischen Median*  $\bar{x}_{1/2}$  definiert man

$$\bar{x}_{1/2} = \begin{cases} x_{\frac{n+1}{2}} : n \text{ ungerade}, \\ x_{\frac{n}{2}} : n \text{ gerade}. \end{cases}$$

**Notation 20.26** (Streudiagramme, bzw. Scatterplots).



**Definition 20.27** (Empirischer Korrelationskoeffizient). Gegeben seien die Wertepaare  $(x_1, y_1), \dots, (x_n, y_n)$ , wobei nicht alle  $x_i$  gleich sind bzw. nicht alle  $y_i$  gleich sind. Seien  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ,  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  die arithmetischen Mittelwerte, und seien

$$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad s_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

die empirischen Varianzen. Unter der *empirischen Kovarianz* (vgl. Def. 19.34) versteht man

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}),$$

und

$$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y}$$

heisst *empirischer Korrelationskoeffizient* oder *Pearson'scher Korrelationskoeffizient*.

**Satz 20.28.** Für den Pearson'schen Korrelationskoeffizient  $r_{xy}$  gilt  $|r_{xy}| \leq 1$ , und es ist  $|r_{xy}| = 1$  genau dann, wenn die Wertepaare  $(x_i, y_i)$ ,  $1 \leq i \leq n$ , auf einer Geraden liegen. Dabei ist die Gerade monoton steigend für  $r_{xy} = 1$ , und monoton fallend für  $r_{xy} = -1$ .

**Bemerkung 20.29.** Je näher  $|r_{xy}|$  an 1 liegt, umso „besser“ liegen die Punkte  $(x_i, y_i)$  um eine Gerade. Ist Steigung der Geraden positiv, dann spricht man von einer positiven (linearen) Korrelation, ansonsten von einer negativen (linearen) Korrelation. Ist  $r_{xy}$  nahe bei 0, so zeigen die Wertepaare (Stichproben) keinen linearen Zusammenhang.

**Satz 20.30.** Gegeben seien die Wertepaare  $(x_1, y_1), \dots, (x_n, y_n)$ , wobei nicht alle  $x_i$  gleich sind bzw. nicht alle  $y_i$  gleich sind. Seien  $\bar{x}, \bar{y}$  die arithmetischen Mittelwerte, seien  $s_x, s_y$  die empirischen Varianzen, und sei  $r_{xy}$  der Pearson'sche Korrelationskoeffizient.

Die Gerade  $f(x) = a \cdot x + b$  für die die quadratische Abweichung  $\sum_{i=1}^n (y_i - f(x_i))^2$  minimal wird, heisst Regressionsgerade oder Ausgleichsgerade und ist gegeben durch

$$a = r_{xy} \frac{s_y}{s_x} \text{ und } b = \bar{y} - a \bar{x}.$$

## 21 Numerische Mathematik

**Satz 21.1.** Zu  $(n+1)$  beliebigen Stützstellen  $(x_i, y_i)$ ,  $0 \leq i \leq n$ ,  $x_i \neq x_j$ ,  $i \neq j$ , gibt es genau ein Polynom  $p : \mathbb{R} \rightarrow \mathbb{R}$  vom Grad  $\leq n$  mit  $p(x_i) = y_i$ ,  $0 \leq i \leq n$ .  $p$  heißt Interpolationspolynom bzgl. den Stützstellen  $(x_i, y_i)$ ,  $0 \leq i \leq n$ .

**Satz 21.2** (Lagrange'sches<sup>51</sup> Interpolationspolynom). Das Interpolationspolynom in der Lagrangeschen Form bzgl. den Stützstellen  $(x_i, y_i)$ ,  $0 \leq i \leq n$ , ist gegeben durch

$$p(x) = \sum_{i=0}^n y_i L_i(x) \quad \text{mit} \quad L_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}, \quad i = 0, \dots, n.$$

**Satz 21.3** (Neville-Algorithmus). Seien  $(x_i, y_i)$ ,  $0 \leq i \leq n$ , Stützstellen, und sei  $p_{i,j}(x)$  das Interpolationspolynom vom Grad höchstens  $j$  bzgl. den Stützstellen  $(x_k, y_k)$ ,  $k = i-j, \dots, i$ ,  $0 \leq j \leq i \leq n$ . Für  $x \in \mathbb{R}$  gilt dann

$$p_{i,j}(x) = \begin{cases} y_i, & j = 0, \\ \frac{(x-x_{i-j}) p_{i,j-1}(x) - (x-x_i) p_{i-1,j-1}(x)}{x_i - x_{i-j}}, & 1 \leq j \leq i. \end{cases}$$

**Satz 21.4** (Newton'sches<sup>52</sup> Interpolationspolynom). Das Interpolationspolynom in der Newtonschen Form bzgl. den Stützstellen  $(x_i, y_i)$ ,  $0 \leq i \leq n$ , ist gegeben durch

$$\begin{aligned} p(x) &= a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots \\ &\quad + a_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}) \\ &= \sum_{i=0}^n a_i \prod_{j=0}^{i-1} (x - x_j) \end{aligned}$$

für geeignete  $a_i \in \mathbb{R}$ . Zur Bestimmung der  $a_i$ ,  $i = 0, \dots, n$ , sei

$$\alpha_{i,j} = \begin{cases} y_i, & j = 0, \\ \frac{\alpha_{i,j-1} - \alpha_{i-1,j-1}}{x_i - x_{i-j}}, & 1 \leq j \leq i. \end{cases}$$

Dann gilt  $a_i = \alpha_{i,i}$ ,  $i = 0, \dots, n$ .

**Satz 21.5** (Horner<sup>53</sup>-Schema). Sei  $f(x)$  ein Polynom in der Form  $f(x) = \sum_{i=0}^n a_i x^i$ . Zum Auswerten von  $f(x^*)$  eignet sich die Beobachtung:

$$f(x^*) = (\dots ((a_n x^* + a_{n-1}) x^* + a_{n-2}) x^* + \dots) x^* + a_0.$$

Ist das Polynom in der Form  $f(x) = \sum_{i=0}^n a_i \prod_{j=0}^{i-1} (x - x_j)$  dann:

$$f(x^*) = (\dots (a_n(x^* - x_{n-1}) + a_{n-1})(x^* - x_{n-2}) + \dots + a_1)(x^* - x_0) + a_0.$$

**Satz 21.6.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine  $(n+1)$ -mal stetig differenzierbare Funktion  $f$ , und seien  $x_i \in [a, b]$ ,  $0 \leq i \leq n$ ,  $n+1$  verschiedene Werte. Sei  $p$  das Interpolationspolynom bzgl. den Stützstellen  $(x_i, f(x_i))$ . Dann existiert für alle  $x^* \in [a, b]$  ein  $\xi \in [a, b]$  mit

$$f(x^*) - p(x^*) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x^* - x_i).$$

---

<sup>51</sup>Joseph-Louis Lagrange, 1736–1813

<sup>52</sup>Isaac Newton, 1643–1727

<sup>53</sup>William George Horner, 1786–1837

**Satz 21.7.** Sei  $(x_0, \dots, x_n)$  die äquidistante Unterteilung des Intervalls  $[a, b]$  in  $n$  Teilintervalle, d.h.

$$x_i = a + i \frac{b-a}{n}, \quad i = 0, \dots, n.$$

Sei  $p_n$  das Interpolationspolynom vom Höchstgrad  $n$  bzgl. den Stützstellen  $(x_i, y_i)$ ,  $0 \leq i \leq n$ . Dann gilt

$$\int_a^b p_n(x) dx = \left( \frac{b-a}{n} \right) \sum_{i=0}^n \alpha_i y_i,$$

mit

$$\alpha_i = \int_0^n \prod_{j=0, j \neq i}^n \frac{t-j}{i-j} dt.$$

**Definition 21.8.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine integrierbare Funktion, sei  $(x_0, \dots, x_n)$  die äquidistante Unterteilung des Intervalls  $[a, b]$  in  $n$  Teilintervalle, und sei  $h = (b-a)/n$ . Die Formel

$$h \sum_{i=0}^n \alpha_i f(x_i)$$

mit

$$\alpha_i = \int_0^n \prod_{j=0, j \neq i}^n \frac{t-j}{i-j} dt,$$

heißt *Newton-Cotes Integrationsformel* vom Grad  $n$  für  $\int_a^b f(x) dx$ .

**Satz 21.9.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine  $(n+1)$ -mal stetig differenzierbare Funktion, sei  $(x_0, \dots, x_n)$  die äquidistante Unterteilung des Intervalls  $[a, b]$  in  $n$  Teilintervalle, und sei  $h = (b-a)/n$ . Dann gilt

$$\left| \int_a^b f(x) dx - h \sum_{i=0}^n \alpha_i f(x_i) \right| = O(h^{n+2}).$$

**Bemerkung 21.10.**

- i) Man sagt, dass die Newton-Cotes Integrationsformel ein Verfahren der Ordnung  $n+2$  ist.
- ii) Da die Güte der Newton-Cotes Integrationsformel  $n$ -Grades von der Intervalllänge  $h = (b-a)/n$  abhängt, unterteilt man ein gegebenes Intervall zunächst in kleine Teilintervalle und wendet dort jeweils die Newton-Cotes Integrationsformel vom Grad  $n$  an. Für  $n=1$  erhält man so die Trapezregel, für  $n=2$  die Simpsonregel und bei  $n=3$  die  $3/8$ -Regel.
- iii) Für  $n \geq 7$  treten negative Gewichte  $\alpha_i$ , so dass die Formeln instabil werden.

**Beispiel 21.11.**  $n=1$ : Trapezregel; sei  $f : [a, b] \rightarrow \mathbb{R}$ ,  $N > 0$ ,  $h = (b-a)/N$  und  $x_i = a + i h$ ,  $i = 0, \dots, N$ . Auf jedem Teilintervall  $[x_i, x_{i+1}]$  wenden wir die Newton-Cotes Integrationsfomel mit  $n=1$  und erhalten dort den Näherungswert

$$I_i = \frac{h}{2} (f(x_i) + f(x_{i+1})) \quad \text{für } \int_{x_i}^{x_{i+1}} f(x) dx.$$

Insgesamt erhalten wir so als Näherungswert von  $\int_a^b f(x) dx$

$$S(h) = \sum_{i=0}^{N-1} \frac{h}{2} (f(x_i) + f(x_{i+1})) = \sum_{i=0}^{(b-a)/h-1} \frac{h}{2} (f(x_i) + f(x_{i+1})).$$

Man kann nun zeigen, dass  $S(h)$  ein "Polynom in  $h^2$ " ist, mit  $S(0) = \int_a^b f(x) dx$ . Die Idee ist nun  $S(h)$  durch ein Interpolationspolynom vom Grad  $r$  bzgl. Stützstellen  $(h_i^2, S(h_i))$  zu approximieren, und dann Näherungen für  $S(0)$  zu erhalten. Die Wahl der Stützstellen führt dann zu verschiedenen Integrationsverfahren.

**Definition 21.12** (Spline). Unter einer Spline(funktion) von Grad  $k$ ,  $k \geq 1$ , bzgl. den  $n+1$  Stützstellen  $x_0 \leq \dots \leq x_n$  versteht man eine  $(k-1)$ -mal stetig differenzierbare Funktion  $S : [x_0, x_n] \rightarrow \mathbb{R}$ , die auf jedem der  $n$  Teilintervallen  $[x_i, x_{i+1}]$ ,  $i = 0, \dots, n-1$ , ein Polynom  $p_i(x)$  mit maximalem Grad  $k$  ist.

**Bemerkung 21.13.** Verwendet man eine Splinefunktion  $S(x)$  vom  $k$ -ten Grad zur Interpolation an den Stützstellen  $(x_i, y_i)$ ,  $i = 0, \dots, n$ , so hat man  $n(k+1)$  Unbekannte (die Koeffizienten der Polynome  $p_i(x)$ ) zu bestimmen. Aus der Definition erhält man die Bedingungen

- i)  $S(x_i) = y_i$ ,  $i = 0, \dots, n$ ,
- ii)  $p_i^{(j)}(x_{i+1}) = p_{i+1}^{(j)}(x_{i+1})$ ,  $j = 0, \dots, k-1$ ,  $i = 0, \dots, n-2$ .

Dies sind insgesamt  $n+1 + (n-1)k$  Bedingungen. Die verbleibenden  $(k-1)$  Bedingungen können unterschiedlich verwendet werden, z.B. kann man im Fall  $y_0 = y_n$  verlangen  $p_0^{(j)}(x_0) = p_n^{(j)}(x_n)$ ,  $j = 0, \dots, k-2$ .

*Beispiel 21.14.* Im Fall  $k=1$  erhält man

$$p_i(x) = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i), \quad i = 0, \dots, n.$$

**Bemerkung 21.15.** Häufig verwendet man kubische Splines, d.h.  $k=3$ , mit den zwei zusätzlichen Bedingungen  $p_0^{(j)}(x_0) = p_n^{(j)}(x_n) = 0$ ,  $j = 0, 1$ . In diesem Fall spricht man von natürlichen Splines (siehe z.B. <http://www.arndt-bruenner.de/mathe/scripts/kub spline.htm>).

**Definition 21.16** (Verträgliche/submultiplikative Matrixnorm (siehe Def. 7.4)). Sei  $|\cdot|$  eine Norm auf dem Vektorraum der reellen  $n \times n$  Matrizen  $\mathbb{R}^{n \times n}$ .

- i)  $|\cdot|$  heißt *submultiplikativ*, falls für alle  $A, B \in \mathbb{R}^{n \times n}$  gilt

$$|AB| \leq |A||B|.$$

- ii) Sei  $|\cdot|^*$  eine (Vektor-) Norm auf  $\mathbb{R}^n$ . Dann heißt  $|\cdot|$  *verträglich mit  $|\cdot|^*$*  falls für alle  $A \in \mathbb{R}^{n \times n}$  und  $\mathbf{x} \in \mathbb{R}^n$  gilt:

$$|A\mathbf{x}|^* \leq |A||\mathbf{x}|^*.$$

*Beispiel 21.17.*  $|A|_1 = \sum_{i,j=1}^n |a_{i,j}|$  ist submultiplikativ und verträglich mit  $|\mathbf{x}|_1 = \sum_{i=1}^n |x_i|$ .

**Notation 21.18.** Im folgenden sei  $\|\cdot\|$  die Matrixnorm

$$\|A\| = \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{0\}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

Diese Norm ist submultiplikativ und verträglich mit der Euklidischen (Vektor-) Norm  $\|\mathbf{x}\|$ ,  $\mathbf{x} \in \mathbb{R}^n$ . Des Weiteren ist

$$\|A\| = \sqrt{\lambda_n(A^\top A)},$$

wobei  $\lambda_n(A^\top A)$  der maximale Eigenwert von  $A^\top A$  ist.

**Definition 21.19** (Kondition). Für  $A \in \mathbb{R}^{n \times n}$ ,  $A$  regulär, heißt

$$\text{cond}(A) = \|A\| \|A^{-1}\|$$

die *Kondition* der Matrix  $A$ .

**Bemerkung 21.20.**

- i) Sei  $U = (u_{i,j})_{1 \leq i,j \leq n}$  eine obere Dreiecksmatrix. Dann lässt sich das Gleichungssystem  $Ux = b$ ,  $b \in \mathbb{R}^n$ , durch Rückwärtseinsetzen, d.h. beginnend mit der letzten Koordinate  $x_n$ , lösen (falls es überhaupt lösbar ist).
- ii) Analog lässt sich ein Gleichungssystem  $Lx = b$ ,  $b \in \mathbb{R}^n$ , wobei  $L$  eine untere Dreiecksmatrix ist, durch Vorwärtseinsetzen, d.h. beginnend mit der ersten Koordinate  $x_1$ , lösen (falls es überhaupt lösbar ist).
- iii) Das Lösen eines Gleichungssystems in Dreiecksform „kostet“  $\mathcal{O}(n^2)$  Operationen.
- iv) Das Inverse einer oberen (unteren) Dreiecksmatrix ist wiederum eine obere (untere) Dreiecksmatrix.

**Definition 21.21** (Permutationsmatrix). Unter einer Permutationsmatrix  $P \in \mathbb{R}^{n \times n}$  versteht man eine Matrix, die durch Vertauschung von Spalten (oder Zeilen) aus der Einheitsmatrix  $I_n$  entsteht.

**Bemerkung 21.22.** Für eine Permutationsmatrix  $P$  gilt  $P^{-1} = P^\top$ . Insbesondere ist  $P$  eine orthogonale Matrix.

**Satz 21.23** (LU-Faktorisierung). Sei  $A \in \mathbb{R}^{n \times n}$  regulär. Das folgende Verfahren (Gauss-Algorithmus) bestimmt eine untere Dreiecksmatrix  $L$  mit Einsen auf der Diagonalen, eine obere Dreiecksmatrix  $U$  und eine Permutationsmatrix  $P$ , so dass  $A = PLU$ :

- i) Setze  $P_0 = L_0 = I_n$  und  $U_0 = A$ , und  $k = 1$ .
- ii) Im  $k$ -ten Schritt werden durch Addition geeigneter  $\alpha_{k,l}$ -Vielfache der  $k$ -ten Zeile von  $U_{k-1}$  zu den  $l = k+1, \dots, n$ -ten Zeilen von  $U_{k-1}$  alle Einträge in der  $k$ -ten Spalte von  $U_{k-1}$  unterhalb der Diagonalen zu 0 gemacht, und man erhält  $U_k$ . Die Einträge von  $L_k$  in der  $k$ -ten Spalte unterhalb der Diagonalen sind  $-\alpha_{k,l}$ . Ist zum Ausführen einer Zeilvertauschung  $k \leftrightarrow m$ ,  $m > k$ , notwendig, dann werden zunächst die entsprechenden Zeilen von  $U_{k-1}$ , und die entsprechenden Spalten von  $P_{k-1}$  vertauscht. Bei  $L_{k-1}$  werden die ersten  $k-1$  Einträge der Zeilen  $k$  und  $m$  ausgetauscht.

**Bemerkung 21.24.**

- i) Aus Stabilitätsgründen wird oft das betragsmäßig größte Elemente der  $k$ -ten Spalte von  $U_{k-1}$  in die  $k$ -Zeile getauscht.
- ii) Zum Bestimmen der LU-Faktorisierung benötigt man  $\mathcal{O}(n^3)$  Operationen. Das Lösen eines Gleichungssystems der Form  $PLU\mathbf{x} = \mathbf{b}$  „kostet“  $\mathcal{O}(n^2)$  Operationen. Dazu berechnet man zunächst ein  $\mathbf{y} \in \mathbb{R}^n$  mit  $L\mathbf{y} = P^{-1}\mathbf{b} = P^\top\mathbf{b}$ , anschließend ein  $\mathbf{x} \in \mathbb{R}^n$  mit  $U\mathbf{x} = \mathbf{y}$ .
- iii) Um bei Computerberechnungen Ungenauigkeiten auszugleichen, kann man beim Lösen eines Gleichungssystems  $A\mathbf{x} = \mathbf{b}$  wie folgt vorgehen: Die erhaltene (erste) Lösung  $\mathbf{x}^{(0)}$  führt zu einem (möglichen) Defekt  $\mathbf{d}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$  und zu einer Defektgleichung  $A\mathbf{y}^{(0)} = \mathbf{d}^{(0)}$ . Setze  $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{y}^{(0)}$  usw. Diesen Prozess nennt man auch Nachiteration.

**Satz 21.25** (Cholesky<sup>54</sup>-Zerlegung). Zu jeder symmetrischen positiv definiten Matrix  $A \in \mathbb{R}^{n \times n}$  gibt es eine untere Dreiecksmatrix  $G = (g_{i,j})$  mit  $A = GG^\top$ . Dabei gilt für  $j = 1, \dots, n$ :

$$g_{j,j} = \sqrt{a_{j,j} - \sum_{k=1}^{j-1} g_{j,k}^2},$$

$$g_{i,j} = \frac{1}{g_{j,j}} \left( a_{i,j} - \sum_{k=1}^{j-1} g_{i,k} g_{j,k} \right), \quad i = j+1, \dots, n.$$

---

<sup>54</sup>André-Louis Cholesky, 1875-1918

**Satz 21.26** (QR-Faktorisierung). Sei  $A \in \mathbb{R}^{m \times n}$  mit  $m \geq n$  und  $\text{rg}(A) = n$ . Seien  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$  die Spaltenvektoren von  $A$ , und seien  $u_1, \dots, u_n$  die zugehörigen Gram-Schmidt orthogonalisierten Vektoren (siehe Satz 7.15), d.h.

$$\mathbf{u}_i = \mathbf{a}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{u}_j, \mathbf{a}_i \rangle}{\|\mathbf{u}_j\|^2} \mathbf{u}_j.$$

Sei  $\mathbf{q}_i = \mathbf{u}_i / \|\mathbf{u}_i\|$ ,  $Q \in \mathbb{R}^{m \times n}$  die Matrix mit Spalten  $\mathbf{q}_1, \dots, \mathbf{q}_n$ , und sei  $R \in \mathbb{R}^{n \times n}$  die oberer Dreiecksmatrix mit  $r_{i,j} = \langle \mathbf{u}_i, \mathbf{a}_j \rangle / \|\mathbf{u}_i\|$ ,  $1 \leq i \leq j \leq n$ . Dann gilt

$$A = Q R.$$

**Bemerkung 21.27.** Sei  $A \in \mathbb{R}^{m \times n}$  mit  $m \geq n$  und  $\text{rg}(A) = n$ . Zum Bestimmen der (möglichen) Lösung des überbestimmten Systems  $A \mathbf{x} = \mathbf{b}$  kann man wie folgt vorgehen:

- i) Man bestimme eine QR-Faktorisierung von  $A$ , also  $A = Q R$ . Dann gilt  $A \mathbf{x} = \mathbf{b} \Rightarrow Q R \mathbf{x} = \mathbf{b} \Rightarrow Q^\top Q R \mathbf{x} = Q^\top \mathbf{b} \Rightarrow R \mathbf{x} = Q^\top \mathbf{b}$ .
- ii) Man löse  $R \mathbf{x} = Q^\top \mathbf{b}$  und überprüfe, ob die Lösung  $\mathbf{x}^*$  auch tatsächlich Lösung von  $A \mathbf{x} = \mathbf{b}$  ist.

**Satz 21.28.** Sei  $A \in \mathbb{R}^{m \times n}$  mit  $m \geq n$ ,  $\text{rg}(A) = n$ , und sei  $b \in \mathbb{R}^m$ . Sei  $A = Q R$  eine QR-Faktorisierung von  $A$ , und sei  $\mathbf{x}^*$  die eindeutige Lösung von  $R \mathbf{x} = Q^\top \mathbf{b}$ . Dann ist  $\mathbf{x}^*$  die Lösung des linearen Ausgleichsproblems :

$$\min_{x \in \mathbb{R}^n} \|A \mathbf{x} - \mathbf{b}\|.$$

**Bemerkung 21.29.** Sei  $A \in \mathbb{R}^{m \times n}$  mit  $m \leq n$ ,  $\text{rg}(A) = m$ , und sei  $b \in \mathbb{R}^m$ . Sei  $A^\top = Q R$  eine QR-Faktorisierung von  $A^\top$ , und sei  $\mathbf{y} \in \mathbb{R}^m$  mit  $R^\top \mathbf{y} = \mathbf{b}$ . Dann ist  $Q \mathbf{y}$  Lösung von  $A \mathbf{x} = \mathbf{b}$ .

**Satz 21.30** (Jacobi<sup>55</sup>-Verfahren). Sei  $A \in \mathbb{R}^{n \times n}$ , und sei  $A = L + D + U$ . Hierbei sei  $D$  die invertierbare Diagonalmatrix mit Elementen  $a_{i,i}$ , und  $L$  ( $U$ ) ist die entsprechende untere (obere) Dreiecksmatrix mit Elementen aus  $A$  und Nullen auf der Diagonalen. Das Jacobi-Verfahren ist das folgende iterative Verfahren zum Lösen  $A \mathbf{x} = \mathbf{b}$ :

- i) Sei  $\mathbf{x}^{(0)} \in \mathbb{R}^n$ . Setze  $k = 1$ .
- ii)  $\mathbf{x}^{(k+1)} = D^{-1} (\mathbf{b} - (L + U) \mathbf{x}^{(k)})$ .
- iii) Setze  $k = k + 1$ . Gehe zu ii), solange  $\|A \mathbf{x}^{(k)} - b\|$  (oder/und auch  $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| / \|\mathbf{x}^{(k)}\|$ ) nicht klein genug ist.

**Definition 21.31** (Strikt diagonaldominant). Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *strikt diagonaldominant*, falls für alle Diagonalelemente gilt:

$$|a_{i,i}| > \sum_{j=1, j \neq i}^n |a_{i,j}|.$$

**Satz 21.32.** Das Jacobi-Verfahren konvergiert für strikt diagonaldominant Matrizen  $A$  gegen eine Lösung des Gleichungssystems  $A \mathbf{x} = \mathbf{b}$ .

**Satz 21.33** (Gauss-Seidel<sup>56</sup> Verfahren). Sei  $A \in \mathbb{R}^{n \times n}$ , und sei  $A = L + D + U$ . Hierbei sei  $D$  die invertierbare Diagonalmatrix mit Elementen  $a_{i,i}$ , und  $L$  ( $U$ ) ist die entsprechende untere (obere) Dreiecksmatrix mit Elementen aus  $A$  und Nullen auf der Diagonalen. Das Gauss-Seidel-Verfahren ist das folgende iterative Verfahren zum Lösen  $A \mathbf{x} = \mathbf{b}$ :

---

<sup>55</sup>Carl Gustav Jacob Jacobi, 1804-1851

<sup>56</sup>Philipp Ludwig Ritter von Seidel, 1821-1896

- i) Sei  $\mathbf{x}^{(0)} \in \mathbb{R}^n$ . Setze  $k = 1$ .
- ii)  $\mathbf{x}^{(k+1)} = (D + L)^{-1} (\mathbf{b} - U \mathbf{x}^{(k)})$ .
- iii) Setze  $k = k + 1$ . Gehe zu ii), solange  $\|A\mathbf{x}^{(k)} - \mathbf{b}\|$  (oder/und auch  $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$ ) nicht klein genug ist.

**Satz 21.34.** Das Gauss-Seidel-Verfahren konvergiert für strikt diagonaldominant Matrizen  $A$  gegen eine Lösung des Gleichungssystems  $A\mathbf{x} = \mathbf{b}$ .

Beispiel 21.35. Sei

$$A = \begin{pmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{pmatrix} \quad b = \begin{pmatrix} 6 \\ 25 \\ -11 \\ 15 \end{pmatrix}.$$

Die exakte Lösung ist  $(1, 2, -1, 1)^\top$ . Die folgende Tabelle gibt die auf fünf Nachkommastellen gerundeten Ergebnisse des Jacobi- und des Gauß-Seidel Verfahrens wieder:

$k$	Jacobi-Verfahren				Gauß-Seidel-Verfahren			
	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
0	0	0	0	0	0	0	0	0
1	0.6	2.2727	-1.1	1.875	0.6	2.3273	-0.98727	0.87886
2	1.0473	1.7159	-0.80523	0.88523	1.0302	2.0369	-1.0145	0.98434
3	0.93264	2.0533	-1.0493	1.1309	1.0066	2.0036	-1.0025	0.99835
4	1.0152	1.9537	-0.96811	0.97384	1.0009	2.0003	-1.0003	0.99985
5	0.98899	2.0114	-1.0103	1.0214	1.0001	2	-1	0.99999
6	1.0032	1.9922	-0.99452	0.99443	1	2	-1	1
:								
10	1.0001	1.9998	-0.99983	0.99979				
11	0.99994	2.0001	-1.0001	1.0001				
12	1	2	-0.99997	0.99996				

**Satz 21.36** (Konvergenz des Newtonverfahrens). Sei  $f : [a, b]$  eine 2-mal stetig differenzierbare Funktion, und seien  $m, M > 0$  mit  $|f'(x)| \geq m$  und  $|f''(x)| \leq M$  für alle  $x \in [a, b]$ . Sei  $x^* \in [a, b]$  eine Nullstelle von  $f$ , d.h.  $f(x^*) = 0$ , und sei  $x^{(0)} \in [a, b]$ . Für  $k \geq 0$  sei

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})} \in [a, b].$$

Dann gilt

$$|x^{(k+1)} - x^*| \leq \frac{M}{2m} |x^{(k)} - x^*|^2,$$

d.h. unter diesen Voraussetzungen konvergiert das Newton-Verfahren quadratisch gegen die Nullstelle.

**Satz 21.37** (Konvergenz des Sekantenverfahrens). Sei  $f : [a, b]$  eine 2-mal stetig differenzierbare Funktion, und seien  $m, M > 0$  mit  $|f'(x)| \geq m$  und  $|f''(x)| \leq M$  für alle  $x \in [a, b]$ . Sei  $x^* \in [a, b]$  eine Nullstelle von  $f$ , d.h.  $f(x^*) = 0$ , und seien  $x^{(0)}, x^{(1)} \in [a, b]$ . Für  $k \geq 1$  sei

$$x^{(k+1)} = x^{(k)} - f(x^{(k)}) \frac{x^k - x^{k-1}}{f(x^{(k)}) - f(x^{k-1})} \in [a, b].$$

Dann gilt

$$|x^{(k+1)} - x^*| \leq \left( \frac{M}{2m} \right)^{\tau-1} |x^{(k)} - x^*|^\tau,$$

wobei  $\tau = (1 + \sqrt{5})/2 = 1.618$  der goldene Schnitt ist. Unter den gegebenen Voraussetzungen konvergiert das Sekantenverfahren superlinear gegen die Nullstelle.

*Beispiel 21.38.* Für  $x > 0$  betrachten wir  $f(x) = x + \ln x$ . Die folgende Tabelle enthält die auf acht Nachkommastellen gerundeten Ergebnisse des Newton- und Sekantenverfahrens:

$k$	Newton	Sekanten
0	0.5	0.5
1	0.56438239	0.6
2	0.56713899	0.5684139
3	0.56714329	0.56712028
4		0.56714331
5		0.56714329

## 22 Gewöhnliche Differentialgleichung

**Definition 22.1** (Gewöhnliche Differentialgleichung). Eine Gleichung

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0,$$

die eine reelle Funktion  $y : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}$ ,  $x \mapsto y(x)$ , mit ihren Ableitungen  $y', y'', \dots, y^{(n)}$  bis zur  $n$ -ten Ordnung verknüpft, heißt *gewöhnliche Differentialgleichung (DG) n-ter Ordnung* in *impliziter Form*. Hierbei ist  $F$  irgendeine stetige Funktion. Lässt sich diese Gleichung nach der höchsten Ableitung auflösen, also

$$y^{(n)}(x) = G(x, y(x), y'(x), \dots, y^{(n-1)}(x))$$

so spricht man von einer *gewöhnlichen Differentialgleichung n-ter Ordnung* in *expliziter Form*. Hängt  $G$  dabei nicht von  $x$  ab, so heißt die Differentialgleichung *autonom*.

**Bemerkung 22.2.**

- i) Die Lösung einer Differentialgleichung ist eine mindestens  $n$ -mal differenzierbare Funktion  $y : D \rightarrow \mathbb{R}$ , die  $F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$  für alle  $x \in D$  erfüllt.
- ii) Statt  $F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$  schreibt man abkürzend  $F(x, y, y', \dots, y^{(n)}) = 0$ .
- iii) Sucht man nach Funktionen  $y(x_1, x_2, \dots, x_n)$  die gewisse Gleichungen der partiellen Ableitungen  $\frac{\partial y}{\partial x_i \partial x_j}$  erfüllen, so spricht man von partiellen Differentialgleichungen. Zum Beispiel beschreibt die Wellengleichung

$$\frac{\partial^2 y}{\partial x^2}(x, t) - \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2}(x, t) = 0$$

die Ausbreitung von (etwa) Schallwellen.  $y(x, t)$  ist dabei die Auslenkung am Ort  $x$  zum Zeitpunkt  $t$ , und  $c$  ist die Ausbreitungsgeschwindigkeit.

**Definition 22.3** (Anfangswertproblem). Sind für die gesuchte Lösung  $y : D \rightarrow \mathbb{R}$  einer Differentialgleichung  $n$ -ter Ordnung zusätzliche  $n$  Forderungen der Form

$$y(a) = y_0, y'(a) = y_1, \dots, y^{(n-1)}(a) = y_{n-1}$$

zu erfüllen, so heißt dies *Anfangswertproblem*. Hierbei ist  $a \in D$  (etwa der Anfang eines Intervalls  $D$ ) und  $y_i \in \mathbb{R}$ ,  $0 \leq i \leq n$ .

**Satz 22.4** (Separation der Variablen). Eine allgemeine Lösung einer trennbaren (separierbaren) Differentialgleichung der Form

$$y' = f(y) g(x)$$

wobei  $f, g$  geeignete Funktionen sind, kann durch Auflösen von

$$\int \frac{dy}{f(y)} = \int g(x) dx$$

nach  $y$  erhalten werden. Sucht man eine spezielle Lösung für die Anfangswertbedingung  $y(x_0) = y_0$  dann erhält man dies durch Spezialisierung der allgemeinen Lösungen, oder durch Auflösen von

$$\int_{y_0}^{y(x)} \frac{dz}{f(z)} = \int_{x_0}^x g(s) ds$$

nach  $y$ . Man bezeichnet diesen Ansatz auch als Trennung der Variablen. Ist  $f(y_0) = 0$ , dann ist  $y(x) = y_0$  (auch) eine Lösung.

**Satz 22.5** (Picard<sup>57</sup>-Lindelöf<sup>58</sup>). Sei  $F$  eine stetig differenzierbare Funktion. Dann hat das Anfangswertproblem

$$y^{(n)} = F(x, y', \dots, y^{(n-1)})$$

mit den  $n$  Anfangsbedingungen

$$y(a) = y_0, y'(a) = y_1, \dots, y^{(n-1)}(a) = y_{n-1}$$

für jede Wahl der Anfangswerte  $y_0, \dots, y_{n-1}$  eine eindeutige Lösung, die in einem offenen Intervall um  $a$  definiert ist, d.h. die lokale Existenz und Eindeutigkeit ist gesichert.

**Satz 22.6.** Gegeben sei die autonome Differentialgleichung erster Ordnung

$$y' = f(y)$$

mit  $f$  differenzierbar und der Anfangsbedingung  $y(x_0) = y_0$ . Sie ist in einem offenen Intervall um  $x_0$  eindeutig lösbar, und es gilt

- i) Ist  $f(y_0) = 0$ , so ist die konstante Funktion  $y(x) = y_0$  die Lösung für alle  $x \in \mathbb{R}$ .
- ii) Ist  $f(y_0) < 0$ , so konvergiert  $y(x)$  (auf dem Intervall) gegen die erste Nullstelle von  $f$  links von  $y_0$  bzw. gegen  $-\infty$  falls solch eine Nullstelle nicht existiert.
- iii) Ist  $f(y_0) > 0$ , so konvergiert  $y(x)$  (auf dem Intervall) gegen die erste Nullstelle von  $f$  rechts von  $y_0$  bzw. gegen  $\infty$  falls solch eine Nullstelle nicht existiert.

**Definition 22.7** (Gleichgewichtslagen). Gegeben sei die autonome Differentialgleichung erster Ordnung

$$y' = f(y).$$

Die Nullstellen der Funktion  $f(t)$  heißen *Gleichgewichtslagen* oder *Fixpunkte* der Differentialgleichung. Ein Fixpunkt  $y_0$  heißt *asymptotisch stabil*, falls es ein offenes Intervall um  $y_0$  gibt, so dass alle Lösungen, die hier starten, gegen  $y_0$  konvergieren.

*Beispiel 22.8.* Für die logistische Wachstumsgleichung

$$y' = \mu(1 - y(x))y(x) - h$$

mit Erntemenge  $h$ ,  $h \in (0, \mu/4)$ , ist die linke Nullstelle  $y_1 = 1/2 - \sqrt{1/4 - h/\mu}$  nicht asymptotisch stabil, und die rechte Nullstelle  $y_2 = 1/2 + \sqrt{1/4 - h/\mu}$  ist asymptotisch stabil.

**Satz 22.9.** Eine Gleichgewichtslage  $y_0$  ist asymptotisch stabil, falls  $f'(y_0) < 0$ . Ist  $f'(y_0) > 0$ , dann ist  $y_0$  nicht asymptotisch stabil.

**Definition 22.10** (Lineare Differentialgleichung). Eine Differentialgleichung der Form

$$\begin{aligned} y^{(n)}(x) &= a_{n-1}(x)y^{(n-1)}(x) + a_{n-2}(x)y^{(n-2)}(x) + \dots + a_1(x)y'(x) \\ &\quad + a_0(x)y(x) + b(x) \end{aligned}$$

heißt *lineare Differentialgleichung n-ter Ordnung*.  $b(x)$  heißt auch *Störfunktion*. Ist  $b(x) = 0$  für alle  $x$ , so nennt man die Differentialgleichung *homogen*, ansonsten *inhomogen*. Hängen die Koeffizienten  $a_i(x)$ ,  $0 \leq i \leq n-1$  und  $b(x)$  nicht von  $x$  ab (sind also Konstanten), dann spricht von einer linearer Differentialgleichung mit *konstanten Koeffizienten*.

---

<sup>57</sup>Charles Émile Picard, 1856-1941

<sup>58</sup>Ernst Leonard Lindelöf, 1870-1946

**Bemerkung 22.11.** Eine lineare Differentialgleichung ist genau dann autonom, wenn sie konstante Koeffizienten besitzt.

**Satz 22.12.** Sei  $I \subset \mathbb{R}$  ein abgeschlossenes Intervall, und seien die Funktionen  $a_i, b : I \rightarrow \mathbb{R}$  stetig,  $0 \leq i \leq n-1$ . Dann hat für jede Wahl von  $a \in I$  und Anfangsbedingungen

$$y(a) = y_0, y'(a) = y_1, \dots, y^{(n-1)}(a) = y_{n-1},$$

die lineare Differentialgleichung

$$y^{(n)}(x) = a_{n-1}(x) y^{(n-1)}(x) + \dots + a_0(x) y(x) + b(x)$$

eine eindeutige Lösung  $y : I \rightarrow \mathbb{R}$ , d.h. die globale Existenz und Eindeutigkeit ist gesichert.

**Satz 22.13.** Die Lösungen einer homogenen linearen Differentialgleichungen

$$y^{(n)}(x) = a_{n-1}(x) y^{(n-1)}(x) + \dots + a_0(x) y(x) \quad (\text{hLDG})$$

bilden einen  $n$ -dimensionalen Vektorraum. Insbesondere gilt für zwei Lösungen  $y_1, y_2$  von **hLDG** und  $\alpha, \beta \in \mathbb{R}$  der DG, dass

$$\alpha y_1 + \beta y_2$$

auch eine Lösung von **hLDG** ist.

**Definition 22.14** (Wronski<sup>59</sup>-Determinante). Seien  $y_1, \dots, y_n$  hinreichend oft differenzierbare Funktionen. Die Determinante

$$W(x) = \det \begin{pmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y'_1(x) & y'_2(x) & \cdots & y'_n(x) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{pmatrix}$$

heißt Wronski-Determinante.

**Satz 22.15.** Seien  $y_i : I \rightarrow \mathbb{R}$ ,  $1 \leq i \leq n$ , hinreichend oft differenzierbare Funktionen. Gibt es ein  $x_0 \in I$  mit  $W(x_0) \neq 0$ , dann sind die Funktionen linear unabhängig.

**Satz 22.16** (vgl. Satz 4.27). Gegeben sei die inhomogene lineare Differentialgleichung

$$y^{(n)}(x) = a_{n-1}(x) y^{(n-1)}(x) + \dots + a_0(x) y(x) + b(x), \quad (\text{iLDG})$$

und es sei  $y_p$  eine feste (partikuläre) Lösung von **iLDG**. Dann ist  $\bar{y}$  eine weitere Lösung von **iLDG** genau dann, wenn  $y_p - \bar{y}$  eine Lösung der zugehörigen homogenen Differentialgleichung

$$y^{(n)}(x) = a_{n-1}(x) y^{(n-1)}(x) + \dots + a_0(x) y(x) \quad (\text{zhLDG})$$

ist. Insbesondere sind alle Lösungen von **iLDG** gegeben durch

$$\{y_p + y_h : y_h \text{ Lösung von } \text{zhLDG}\}.$$

**Satz 22.17.** Die lineare homogene Differentialgleichung 1.ter Ordnung

$$y'(x) = a_0(x) y(x)$$

hat die allgemeine Lösung

$$y(x) = c e^{\int a_0(x) dx}.$$

---

<sup>59</sup> Josef Hoené-Wronski , 1776-1853

*Beispiel 22.18.* Für  $y' = -2x y$  erhält man

$$y = c e^{\int -2x \, dx} = c e^{-x^2}.$$

**Bemerkung 22.19** (Variation der Konstanten). Gegeben sei die inhomogene lineare Differentialgleichung 1.ter Ordnung

$$y'(x) = a_0(x) y(x) + b(x).$$

Zum Bestimmen einer partikulären Lösung  $y_p$  verwendet man die Methode der Variation der Konstanten, d.h., man macht den Ansatz

$$y_p(x) = c(x) y_h(x),$$

wobei  $y_h$  eine Lösung der zugehörigen homogenen DG ist, d.h. man variiert die „Konstante  $c(x)$ “. Einsetzen in die DG führt zu

$$c'(x) y_h(x) = b(x) \quad \text{bzw. } c'(x) = \frac{b(x)}{y_h(x)},$$

und so

$$c(x) = \int b(x) e^{-\int a_0(x) \, dx} \, dx.$$

Somit ist eine partikuläre Lösung  $y_p$  der inhomogenen linearen Differentialgleichung 1.ter Ordnung

$$y'(x) = a_0(x) y(x) + b(x)$$

gegeben durch

$$y_p = \left( \int b(x) e^{-\int a_0(x) \, dx} \, dx \right) e^{\int a_0(x) \, dx}.$$

**Satz 22.20.** Die allgemeine Lösung der inhomogenen linearen Differentialgleichung 1.ter Ordnung

$$y'(x) = a_0(x) y(x) + b(x)$$

ist gegeben durch

$$\begin{aligned} y &= c e^{\int a_0(x) \, dx} + \left( \int b(x) e^{-\int a_0(x) \, dx} \, dx \right) e^{\int a_0(x) \, dx} \\ &= e^{\int a_0(x) \, dx} \left( c + \int b(x) e^{-\int a_0(x) \, dx} \, dx \right), \quad c \in \mathbb{R}. \end{aligned}$$

**Satz 22.21.** Seien  $y_1, y_2$  zwei lineare unabhängige Lösung einer homogenen linearen Differentialgleichung 2.ter Ordnung

$$y''(x) = a_1(x) y'(x) + a_0(x) y(x). \quad (\text{hDG2})$$

Dann gilt für ihre Wronski-Determinante

$$W(x) = y_1 y'_2 - y'_1 y_2 = c e^{\int a_1(x) \, dx},$$

wobei  $c$  eine Konstante ist.

**Bemerkung 22.22.** Kennt man eine (von Null verschiedene) Lösung der DG hDG2, so kann man die obige Beziehung ausnutzen um eine zweite, dazu linear unabhängige Lösung zu bestimmen.

**Bemerkung 22.23.** Gegeben sei die inhomogene lineare Differentialgleichung 2.ter Ordnung

$$y''(x) = a_1(x) y'(x) + a_0(x) y(x) + b(x), \quad (\text{ilDG2})$$

und seien  $y_1, y_2$  zwei linear unabhängige Lösungen der zugehörigen homogenen linearen DG. Zum Bestimmen einer partikulären Lösung  $y_p$  von **ilDG2** verwendet man wieder den Ansatz der Variation der Konstanten  $y_p(x) = c_1(x) y_1(x) + c_2(x) y_2(x)$ . Einsetzen in **ilDG2** liefert zusammen mit der Bedingung  $c'_1 y_1 + c'_2 y_2 = 0$  das System

$$\begin{aligned} c'_1 y_1 + c'_2 y_2 &= 0 \\ c'_1 y'_1 + c'_2 y'_2 &= b(x). \end{aligned}$$

Die Koeffizientendeterminante dieses Systems ist gleich der Wronski-Determinante, also ungleich Null, d.h. das Gleichungssystem ist eindeutig lösbar.

**Satz 22.24.** Gegeben sei die inhomogene lineare Differentialgleichung 2.ter Ordnung

$$y''(x) = a_1(x) y'(x) + a_0(x) y(x) + b(x), \quad (\text{ilDG2})$$

und seien  $y_1, y_2$  zwei linear unabhängige Lösungen der zugehörigen homogenen linearen DG. Dann ist

$$y_p(x) = - \left( \int \frac{y_2(x) b(x)}{W(x)} dx \right) y_1(x) + \left( \int \frac{y_1(x) b(x)}{W(x)} dx \right) y_2(x)$$

eine partikuläre Lösung.

**Definition 22.25** (Charakteristisches Polynom). Für eine lineare homogene Differentialgleichung  $n$ -ter Ordnung mit konstanten Koeffizienten

$$y^{(n)}(x) = a_{n-1} y^{(n-1)}(x) + a_{n-2} y^{(n-2)}(x) + \cdots + a_o y(x)$$

heißt das Polynom

$$\tau(\lambda) = \lambda^n - a_{n-1} \lambda^{n-1} - a_{n-2} \lambda^{n-2} - \cdots - a_o$$

charakteristisches Polynom der Differentialgleichung.

**Satz 22.26** (Lineare homogene Differentialgleichung mit konstanten Koeffizienten). Gegeben sei eine lineare homogene Differentialgleichung  $n$ -ter Ordnung mit konstanten Koeffizienten

$$y^{(n)}(x) = a_{n-1} y^{(n-1)}(x) + a_{n-2} y^{(n-2)}(x) + \cdots + a_o y(x). \quad (\text{IDGk})$$

Man bestimme alle verschiedenen Nullstellen des charakteristischen Polynoms  $\tau(\lambda)$  von **IDGk** einschließlich der algebraischen Vielfachheit, d.h. wie oft sie als Nullstelle auftreten. Sei  $\alpha$  eine  $k$ -fache Nullstelle.

- Ist  $\alpha$  reell, dann sind

$$e^{\alpha x}, x e^{\alpha x}, \dots, x^{k-1} e^{\alpha x}$$

linear unabhängige Lösungen von **IDGk**.

- Ist  $\alpha = \beta + \gamma i$  komplex, dann ersetze man für die beiden Nullstellen  $\alpha = \beta + \gamma i$  und  $\bar{\alpha} = \beta - \gamma i$  die Lösungsmenge durch

$$\begin{aligned} e^{\beta x} \cos \gamma x, x e^{\beta x} \cos \gamma x, \dots, x^{k-1} e^{\beta x} \cos \gamma x, \\ e^{\beta x} \sin \gamma x, x e^{\beta x} \sin \gamma x, \dots, x^{k-1} e^{\beta x} \sin \gamma x. \end{aligned}$$

Die durch dieses Vorgehen bestimmten Lösungen sind  $n$  linear unabhängige Lösungen von **IDGk**.

**Satz 22.27.** Gegeben sei eine lineare inhomogene Differentialgleichung  $n$ -ter Ordnung mit konstanten Koeffizienten

$$y^{(n)}(x) = a_{n-1} y^{(n-1)}(x) + a_{n-2} y^{(n-2)}(x) + \cdots + a_0 y(x) + b(x), \quad (\text{liDGk})$$

und sei  $\tau(\lambda)$  das charakteristische Polynom der zugehörigen homogenen Differentialgleichung.

i) Sei  $b(x) = p_m(x) e^{\alpha x}$  für ein Polynom  $p_m(x)$  vom Grad  $m$  und  $\alpha \in \mathbb{R}$ .

- Ist  $\tau(\alpha) \neq 0$ , dann machen wir den Ansatz

$$y_p(x) = \left( \sum_{i=0}^m b_i x^i \right) e^{\alpha x}.$$

- Ist  $\tau(\alpha) = 0$  mit algebraischer Vielfachheit  $k$ , dann

$$y_p(x) = x^k \left( \sum_{i=0}^m b_i x^i \right) e^{\alpha x}.$$

ii) Sei  $b(x) = p_m(x) e^{\beta x} \cos \gamma x$  oder  $b(x) = p_m(x) e^{\beta x} \sin \gamma x$  für ein Polynom  $p_m(x)$  vom Grad  $m$  und  $\beta, \gamma \in \mathbb{R}$ .

- Ist  $\tau(\beta + \gamma i) \neq 0$ , dann machen wir den Ansatz

$$y_p(x) = \left( \sum_{i=0}^m b_i x^i \right) e^{\beta x} (\mu_1 \cos \gamma x + \mu_2 \sin \gamma x).$$

- Ist  $\tau(\beta + \gamma i) = 0$  mit algebraischer Vielfachheit  $k$ , dann

$$y_p(x) = x^k \left( \sum_{i=0}^m b_i x^i \right) e^{\beta x} (\mu_1 \cos \gamma x + \mu_2 \sin \gamma x).$$

Die unbekannten Koeffizienten  $b_i, \mu_i$  können durch Einsetzen von  $y_p$  in [liDGk](#) und durch Koeffizientenvergleich bestimmt werden. Die so bestimmten  $y_p$  sind partikuläre Lösungen von [liDGk](#)

**Satz 22.28.** Sind  $y_p^1$  und  $y_p^2$  Lösungen der linearen Differentialgleichungen

$$y^{(n)}(x) = \sum_{i=0}^{n-1} a_i(x) y^{(i)}(x) + b_1(x) \quad \text{und} \quad y^{(n)}(x) = \sum_{i=0}^{n-1} a_i(x) y^{(i)}(x) + b_2(x),$$

dann ist  $y_p^1 + y_p^2$  Lösung der Differentialgleichung

$$y^{(n)}(x) = \sum_{i=0}^{n-1} a_i(x) y^{(i)}(x) + b_1(x) + b_2(x).$$

**Bemerkung 22.29.** Ein weiterer Ansatz zum Lösen von Differentialgleichungen ist der sogenannte Potenzreihenansatz, d.h. man setzt

$$y(x) = \sum_{n=0}^{\infty} \alpha_n (x - x_0)^n$$

für ein geeignetes festes  $x_0$ . Einsetzen in die Differentialgleichung kann dann durch Koeffizientenvergleich zu einer Lösung führen.

*Beispiel 22.30.* Für ein gegebenes  $m \in \mathbb{N}$  betrachten wir die DG

$$x y'' + (1-x) y' + m y = 0.$$

Substitution von  $y(x) = \sum_{n=0}^{\infty} \alpha_n (x - x_0)^n$  liefert

$$\begin{aligned} x \left( \sum_{n=2}^{\infty} n(n-1) \alpha_n (x - x_0)^{n-2} \right) + (1-x) \left( \sum_{n=1}^{\infty} n \alpha_n (x - x_0)^{n-1} \right) \\ + m \left( \sum_{n=0}^{\infty} \alpha_n (x - x_0)^n \right) = 0. \end{aligned}$$

Speziell für  $x_0 = 0$  reduziert sich dies zu

$$\sum_{n=0}^{\infty} ((n+1)^2 \alpha_{n+1} + (m-n) \alpha_n) x^n = 0.$$

Somit muss für  $n = 0, 1, \dots$  gelten

$$(n+1)^2 \alpha_{n+1} + (m-n) \alpha_n = 0,$$

was zu den Lösungen

$$y(x) = \alpha_0 \sum_{n=0}^m \frac{(-1)^n}{n!} \binom{m}{n} x^n$$

führt. Hierbei ist  $\alpha_0 \in \mathbb{R}$  beliebig. Diese Polynome heißen Laguerre<sup>60</sup>-Polynome.

**Bemerkung 22.31.** Ein einfaches numerisches Verfahren zum Lösen des Anfangswertproblems

$$y'(x) = f(x, y(x)) \text{ mit } y(x_0) = y_0$$

ist das sogenannte Eulerverfahren. Hierbei ersetzt man iterativ die gesuchte Lösung auf einem Intervall  $[x_i, x_{i+1}]$  durch die Tangente in  $(x_i, y_i)$  gegeben durch

$$t_i(x) = y_i + y'(x_i)(x - x_i) = y_i + f(x_i, y_i)(x - x_i).$$

Für vorgegebene Intervallpunkte  $x_i$ ,  $i = 0, \dots, m$ , bildet dann der Polygonzug durch die Punkte  $(x_i, y_i)$ ,  $i \geq 0$ , mit  $y_i = t_{i-1}(x_i)$  für  $i \geq 1$  eine Approximation der Lösung.

---

<sup>60</sup>Edmond Nicolas Laguerre, 1834–1886

## 23 Kurven und Flächen

**Definition 23.1** (Kurven). Eine Abbildung  $c : [a, b] \rightarrow \mathbb{R}^n$  der Form

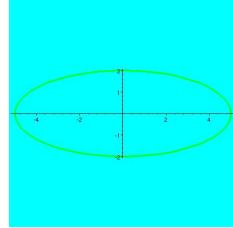
$$t \mapsto \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_n(t) \end{pmatrix}$$

heißt *parameterisierte Kurve*. Die Funktionen  $c_i : [a, b] \rightarrow \mathbb{R}$  heißen die *Komponentenfunktionen*. Sind alle Komponentenfunktionen Polynome, dann wird die Kurve als *Polynomkurve* bezeichnet.

**Bemerkung 23.2.** Eine andere Darstellungsform für eine Kurve  $c$  im  $\mathbb{R}^2$  ist die sogenannte implizite Darstellung  $c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ , wobei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine Abbildung ist. Ist dabei  $f$  ein Polynom in  $(x, y)$ , dann heißt  $c$  eine *algebraische Kurve*.

**Beispiel 23.3.** i) Seien  $a, b > 0$ . Die Ellipse  $c$  mit Achsen auf den Koordinatenachsen und den Achsen-Schnittpunkten  $\pm(a, 0)^\top, \pm(0, b)^\top$  besitzt (etwa) die Darstellungen

$$c = \left\{ \begin{pmatrix} a \cos t \\ b \sin t \end{pmatrix} : t \in [0, 2\pi] \right\} = \left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \right\}.$$



ii) Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion. Dann ist der Graph  $\{(t, f(t))^\top : t \in [a, b]\}$  das Bild der Kurve  $c(t) = (t, f(t))^\top, t \in [a, b]$ .

**Definition 23.4** (Tangentialvektor)). Sei  $c : [a, b] \rightarrow \mathbb{R}^n$  eine Kurve mit differenzierbaren Komponentenfunktionen. Dann heißt

$$\dot{c}(t) = \begin{pmatrix} c'_1(t) \\ c'_2(t) \\ \vdots \\ c'_n(t) \end{pmatrix}$$

Tangentialvektor (Geschwindigkeitsvektor) der Kurve in  $c(t)$ .

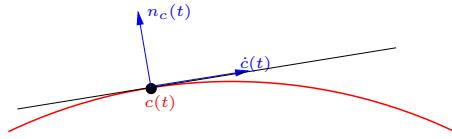
**Bemerkung 23.5.** Sei  $c : [a, b] \rightarrow \mathbb{R}^2$  eine Kurve mit differenzierbaren Komponentenfunktionen, und es sei  $t^* \in [a, b]$  mit  $\dot{c}(t^*) \neq 0$ . Der Vektor

$$n_c(t^*) = \frac{1}{\|\dot{c}(t^*)\|} \begin{pmatrix} -c'_2(t^*) \\ c'_1(t^*) \end{pmatrix}$$

heißt Normalen(einheits)vektor der Kurve in  $c(t^*)$ .

**Bemerkung 23.6.**

- i) Der Normalenvektor ist senkrecht zum Tangentialvektor, der eine Tangentialgerade an die Kurve erzeugt.



- ii) Sei  $c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$  eine implizite Darstellung einer Kurve, wobei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  partiell differenzierbar ist, und es sei  $\text{grad } f(x^*, y^*) \neq 0$ . Dann ist – bis auf das Vorzeichen – der Normalen(einheits)vektor in  $(x^*, y^*)$  gegeben durch

$$\frac{1}{\|\text{grad } f(x^*, y^*)\|} \text{grad } f(x^*, y^*).$$

**Definition 23.7** (Länge einer Kurve). Sei  $c : [a, b] \rightarrow \mathbb{R}^n$  eine Kurve mit stetig differenzierbaren Komponentenfunktionen. Dann heißt

$$L(c) = \int_a^b \sqrt{\sum_{i=1}^n [c'_i(t)]^2} dt = \int_a^b \|\dot{c}(t)\| dt$$

Länge der Kurve  $c$  von  $a$  nach  $b$ . Die Funktion  $L_c : [a, b] \rightarrow \mathbb{R}$  mit

$$L_c(l) = \int_a^l \|\dot{c}(t)\| dt$$

heißt Längenfunktion der Kurve  $C$ .

**Definition 23.8** (Reparameterisierung). Seien  $c : [a, b] \rightarrow \mathbb{R}^n$  eine Kurve, und sei  $[\hat{a}, \hat{b}] \subset \mathbb{R}$ . Sei weiterhin  $h : [\hat{a}, \hat{b}] \rightarrow [a, b]$  eine bijektive, differenzierbare Funktion mit  $h'(t) > 0$  (bzw.  $h'(t) < 0$ ) für alle  $t \in [a, b]$ . Dann heißt die Kurve  $d : [\hat{a}, \hat{b}] \rightarrow \mathbb{R}^n$  mit  $d(t) = c(h(t))$  eine Reparameterisierung von  $c$ .

**Definition 23.9** (Bernsteinpolynome<sup>61</sup>). Für  $0 \leq i \leq m$  heißt das Polynom  $b_{i,m} : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$b_{i,m}(t) = \binom{m}{i} (1-t)^{m-i} t^i = \frac{m!}{(m-i)! i!} (1-t)^{m-i} t^i$$

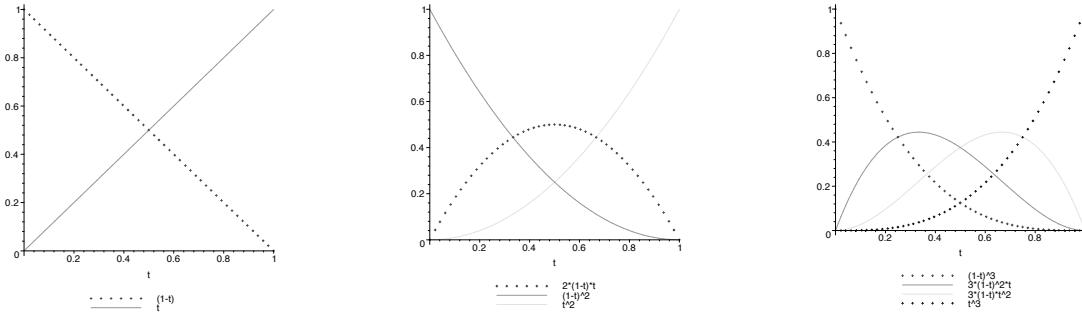
$i$ -tes Bernsteinpolynom vom Grad  $m$ .

Beispiel 23.10.

$$\begin{aligned} b_{0,m} &= (1-t)^m, & b_{1,m} &= m(1-t)^{m-1} t, \\ b_{m,m} &= t^m, & b_{m-1,m} &= m(1-t)t^{m-1}. \end{aligned}$$

---

<sup>61</sup>Sergei Natanowitsch Bernstein, 1880–1968



Lineare (Grad 1), quadratische (Grad 2) und kubische (Grad 3) Bernsteinpolynome

**Satz 23.11** (Eigenschaften von Bernsteinpolynomen). Für  $i > m$ ,  $i < 0$  oder  $m < 0$  setzen wir  $b_{i,m}(t) = 0$ .

- i)  $b_{i,m}(t) \geq 0$  für  $t \in [0, 1]$ ,
- ii)  $\sum_{i=0}^m b_{i,m}(t) = 1$ ,
- iii)  $b_{i,m}(t) = (1-t) \cdot b_{i,m-1}(t) + t \cdot b_{i-1,m-1}(t)$ ,
- iv)  $b'_{i,m}(t) = m(b_{i-1,m-1}(t) - b_{i,m-1}(t))$ ,
- v)  $b_{i,m}(\alpha t) = \sum_{j=i}^m b_{i,j}(\alpha) b_{j,m}(t)$  für  $\alpha \in \mathbb{R}$ .

**Definition 23.12** (Bézierkurven<sup>62</sup>). Seien  $c_0, \dots, c_m \in \mathbb{R}^n$ . Die Bézierkurve  $B : [0, 1] \rightarrow \mathbb{R}$  von Grad  $m$  bzgl. den Kontrollpunkten  $c_0, \dots, c_m$  ist definiert durch

$$B(t) = \sum_{i=0}^m b_{i,m}(t) c_i,$$

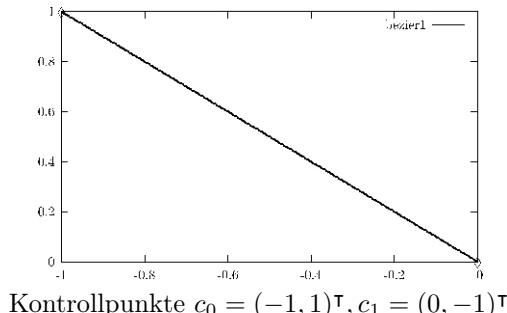
wobei  $b_{i,m}(t)$  die Bernsteinpolynome sind.

Im Falle  $n = 2$  heißt der Polygonzug  $c_0, c_1, \dots, c_m$  auch Kontrollpolygonzug.

*Beispiel 23.13.*

- Lineare Bézierkurve (Grad 1) mit zwei Kontrollpunkten  $c_0$  und  $c_1$  ist die Strecke

$$B(t) = (1-t) c_0 + t c_1, \quad t \in [0, 1].$$



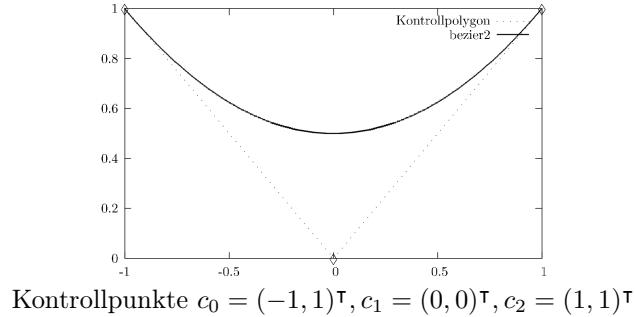
Kontrollpunkte  $c_0 = (-1, 1)^\top, c_1 = (0, -1)^\top$

---

<sup>62</sup>Pierre Étienne Bézier, 1910 – 1999

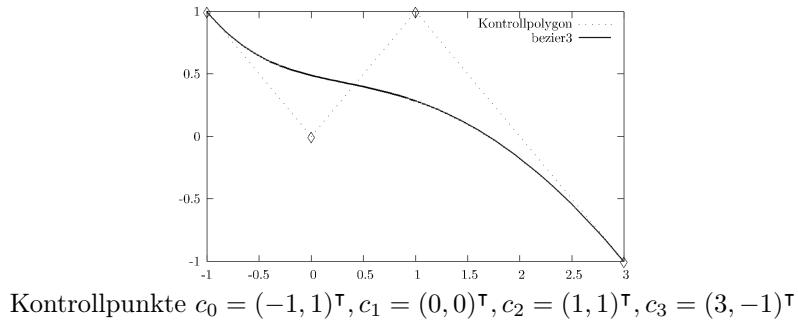
- *Quadratische Bézierkurve* (Grad 2) mit drei Kontrollpunkten  $c_0, c_1$  und  $c_2$  ist die Kurve

$$B(t) = (1-t)^2 c_0 + 2(1-t)t c_1 + t^2 c_2, \quad t \in [0, 1].$$

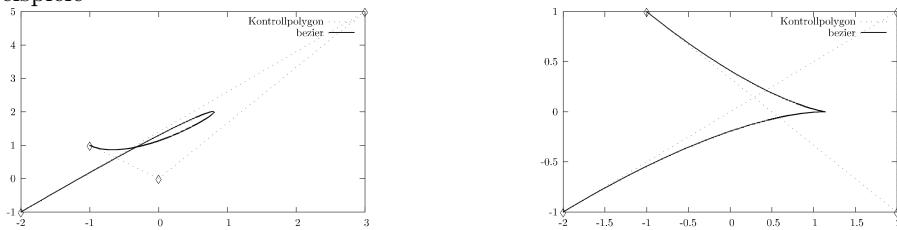


- *Kubische Bézierkurve* (Grad 3) mit vier Kontrollpunkten  $c_0, c_1, c_2$  und  $c_3$  ist die Kurve

$$B(t) = (1-t)^3 c_0 + 3(1-t)^2 t c_1 + 3(1-t)t^2 c_2 + t^3 c_3, \quad t \in [0, 1].$$



- weitere Beispiele



**Satz 23.14.** Jede Polynomkurve  $c(t)$ ,  $t \in [0, 1]$ , vom Grad  $m$  lässt sich als Bézierkurve vom Grad  $m$  darstellen, d.h., sei

$$c(t) = \sum_{i=0}^m a_i t^i,$$

für geeignete  $a_i \in \mathbb{R}^n$ , dann gibt es Kontrollpunkte  $c_i \in \mathbb{R}^n$ ,  $0 \leq i \leq m$ , so dass  $c(t) = \sum_{i=0}^m b_{i,m}(t) c_i$ . Für die Kontrollpunkte  $c_i$  gilt dabei

$$c_i = \sum_{j=0}^i \frac{\binom{i}{j}}{\binom{m}{j}} a_j.$$

**Satz 23.15** (Eigenschaften von Bézierkurven). Sei  $B(t)$  eine Bézierkurve vom Grad  $m$ . Dann gilt

- i)  $B(0) = c_0, B(1) = c_m,$
- ii)  $\dot{B}(0) = m \cdot (c_1 - c_0), \dot{B}(1) = m(c_m - c_{m-1}),$
- iii)

$$B(t) \in \left\{ \sum_{i=0}^m \lambda_i c_i : \sum_{i=0}^m \lambda_i = 1, \quad \lambda_i \geq 0 \right\}, \quad t \in [0, 1],$$

d.h. das Bild der Bézierkurve liegt in der konvexen Hülle der Kontrollpunkte. Dies ist die kleinste konvexe Menge, die die Kontrollpunkte enthält. Insbesondere ist die Bézierkurve (fast) eine Gerade wenn alle Kontrollpunkte (fast) auf einer Geraden liegen.

- iv) Sei  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine affine Abbildung (siehe Definition 6.8). Dann gilt

$$T(B(t)) = \sum_{i=0}^m b_{i,m}(t) T(c_i).$$

**Satz 23.16** (De Casteljau<sup>63</sup> - Algorithmus). Sei  $B(t)$  eine Bézierkurve vom Grad  $m$  bzgl. den Kontrollpunkten  $c_0, \dots, c_m$ . Sei  $t^* \in [0, 1]$ ,  $c_i^{(0)} = c_i$  für  $0 \leq i \leq m$ , und für  $1 \leq j \leq m$  sei

$$c_i^{(j)} = (1 - t^*) \cdot c_i^{(j-1)} + t^* \cdot c_{i+1}^{(j-1)}, \quad \text{für } 0 \leq i \leq m-j.$$

Dann gilt:

$$c_i^{(j)} = \sum_{k=0}^j b_{k,j}(t^*) c_{i+k}.$$

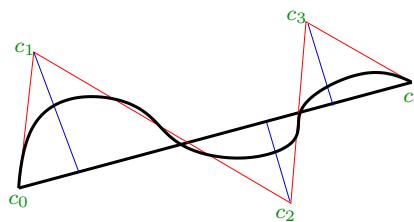
Insbesondere ist  $B(t^*) = c_0^{(m)}$ .

**Satz 23.17** (Unterteilung von Bézierkurven). Sei  $B(t)$  eine Bézierkurve vom Grad  $m$  bzgl. den Kontrollpunkten  $c_0, \dots, c_m \in \mathbb{R}^n$ . Sei  $\alpha \in [0, 1]$  und sei  $B_l(t) = B(t)$  für  $t \in [0, \alpha]$  und  $B_r(t) = B(t)$  für  $t \in [\alpha, 1]$ .  $B_l(t)$  und  $B_r(t)$  sind Polynomkurven vom Grad  $m$  und daher als Bézierkurven über dem Intervall  $[0, 1]$  darstellbar.

Sind  $c_i^{(j)}$  die Punkte des “de Casteljau” Algorithmus von  $B(t)$  für den Wert  $t^* = \alpha$ , so ist  $B_l(t)$  die Bézierkurve vom Grad  $m$  bzgl. den Kontrollpunkten  $c_0^{(j)}, j = 0, \dots, m$ , und  $B_r(t)$  ist die Bézierkurve vom Grad  $m$  bzgl. den Kontrollpunkten  $c_j^{(m-j)}, j = 0, \dots, m$ .

**Bemerkung 23.18.** Eine Anwendung von Bézierkurven ist das “Rendering” von Kurven, um sie auszugeben (Drucker, Bildschirm, etc.). Da sich lineare Kurven (Strecken) einfach ausgeben lassen und die Auflösung des Ausgabemediums nicht beliebig fein sein kann, wird zunächst festgelegt, wann eine Bézierkurve  $B(t)$  vom Grad  $m$  mit Kontrollpunkten  $c_0, \dots, c_m$  fast linear ist:

Dazu gebe man eine Toleranz  $\epsilon > 0$  vor und berechne den maximalen Abstand eines Punktes  $c_i, i = 0, \dots, m$ , von der Geraden durch die beiden Endpunkte  $c_0$  und  $c_m$ . Ist dieser Abstand nicht größer als die vorgegebene Toleranz, dann ist die Kurve fast linear.




---

<sup>63</sup>Paul de Casteljau, 1930 –

Sei  $c(t) : [a, b] \rightarrow \mathbb{R}^2$  eine Polynomkurve vom Grad  $m$ . Zum "Plotten" von  $c(t)$  bietet sich folgendes rekursives Verfahren an:

- (0) Stelle  $c(t)$  als Bézierkurve  $B(t)$  vom Grad  $m$  über  $[0, 1]$  dar. Ist  $B(t)$  *fast linear*, gehe zu (3).
- (1) Unterteile  $B(t)$  in zwei Bézierkurven  $B_l(t), B_r(t)$  bzgl. dem Wert  $\alpha = 1/2$ .
- (2)
  - Ist  $B_l(t)$  *fast linear* gehe zu (3), sonst wende (1) auf  $B_l(t)$  an.
  - Ist  $B_r(t)$  *fast linear* gehe zu (3), sonst wende (1) auf  $B_r(t)$  an.
- (3) Man gebe die Strecke zwischen den Kontrollpunkten  $c_0$  und  $c_m$  aus.

**Definition 23.19** ((2)-Fläche). Sei  $U \subset \mathbb{R}^2$ . Eine Abbildung  $S : U \rightarrow \mathbb{R}^3$  der Form

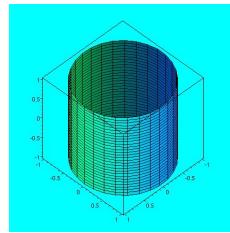
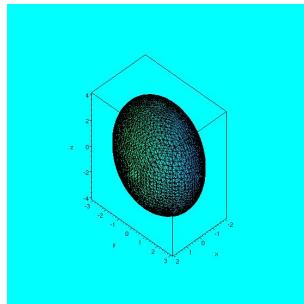
$$S(s, t) = \begin{pmatrix} x_1(s, t) \\ x_2(s, t) \\ x_3(s, t) \end{pmatrix}$$

heißt *2-Fläche*. Die Funktionen  $x_i : U \rightarrow \mathbb{R}$  heißen die *Komponentenfunktionen*. Sind alle Komponentenfunktionen Polynome, dann wird die Fläche als *Polynomfläche* bezeichnet.

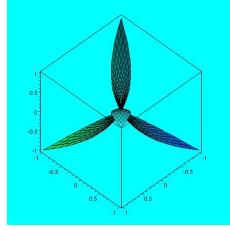
**Bemerkung 23.20.** Eine andere (und üblichere) Darstellungsform für eine Fläche  $c$  im  $\mathbb{R}^3$  ist auch hier die implizite Darstellung  $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : f(x_1, x_2, x_3) = 0\}$ , wobei  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  eine Abbildung ist. Ist dabei  $f$  ein Polynom in  $(x_1, x_2, x_3)$ , dann heißt  $S$  eine *algebraische Fläche*.

*Beispiel 23.21.* Seien  $a, b, c > 0$ . Das Ellipsoid  $S$  mit Achsen auf den Koordinatenachsen und den Achsen-Schnittpunkten  $\pm(a, 0, 0)^T, \pm(0, b, 0)^T, \pm(0, 0, c)^T$  besitzt (etwa) die Darstellungen

$$\begin{aligned} S(s, t) &= \{(a \sin s \cos t, b \sin s \sin t, c \cos s)^T : s \in [0, \pi], t \in [0, 2\pi]\} \\ &= \left\{ (x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0 \right\}. \end{aligned}$$



Zylinder:  $S(s, t) = (\cos(s), \sin(s), t)^T$ ,  $s \in [0, 2\pi]$ ,  $t \in [-1, 1]$ .



Propeller:  $S(s, t) = (s^2 t, s t^2, s t)^T, s \in [-1, 1], t \in [-1, 1]$ .

**Definition 23.22** (Tangentialebene). Sei  $S : U \rightarrow \mathbb{R}^3, U \subset \mathbb{R}^2$ , eine Fläche. Für  $(s^*, t^*) \in U$  betrachten wir die beiden Kurven  $S_{s^*}(t) = S(s^*, t)$  und  $S_{t^*}(s) = S(s, t^*)$ .

Sind die beiden Tangentialvektoren  $\dot{S}_{s^*}(t^*)$  und  $\dot{S}_{t^*}(s^*)$  linear unabhängig, dann heißt

$$S(s^*, t^*) + \text{lin} \{ \dot{S}_{s^*}(t^*), \dot{S}_{t^*}(s^*) \}$$

Tangentialebene von  $S$  in  $S(s^*, t^*)$ . Der Vektor

$$n_S(s^*, t^*) = \frac{\dot{S}_{s^*}(t^*) \times \dot{S}_{t^*}(s^*)}{\| \dot{S}_{s^*}(t^*) \times \dot{S}_{t^*}(s^*) \|}$$

Normalen(einheits)vektor der Fläche  $S$  in  $S(s^*, t^*)$ .

**Bemerkung 23.23.** Sei  $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : f(x_1, x_2, x_3) = 0\}$  eine implizite Darstellung einer Fläche, wobei  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  partiell differenzierbar ist, und es sei  $\text{grad } f(x_1^*, x_2^*, x_3^*) \neq 0$ . Dann ist – bis auf das Vorzeichen – der Normalen(einheits)vektor in  $(x_1^*, x_2^*, x_3^*)$  gegeben durch

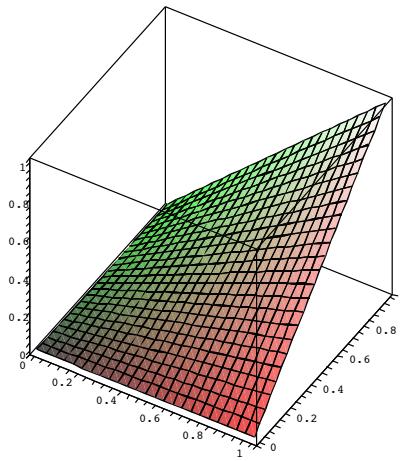
$$\frac{1}{\| \text{grad } f(x_1^*, x_2^*, x_3^*) \|} \text{grad } f(x_1^*, x_2^*, x_3^*).$$

**Definition 23.24** (Bézierflächen). Seien  $c_{i,j}, i = 0, \dots, m, j = 0, \dots, p$ , (Kontrollpunkte) im  $\mathbb{R}^3$ . Dann heißt

$$S(s, t) = \sum_{i=0}^m \sum_{j=0}^p b_{i,m}(s) b_{j,p}(t) c_{i,j}, \quad (s, t) \in [0, 1] \times [0, 1],$$

Bézierfläche.

*Beispiel 23.25.* Sei  $m = p = 1$ , und  $c_{0,0} = (0, 0, 0)^T, c_{0,1} = (1, 0, 0)^T, c_{1,0} = (0, 1, 0)^T, c_{1,1} = (1, 1, 1)^T$ .



**Satz 23.26** (Eigenschaften von Bézierflächen). *Sei  $S(s, t)$  eine Bézierfläche mit Kontrollpunkten  $c_{i,j} \in \mathbb{R}^3$ ,  $i = 0, \dots, m$ ,  $j = 0, \dots, p$ . Dann gilt:*

- i)  $S(s, t) \in \text{conv}\{c_{0,0}, \dots, c_{m,p}\}$  für  $s, t \in [0, 1]$ , d.h.  $S(s, t)$  ist enthalten in dem Polytop, welches von den Kontrollpunkten erzeugt wird. Allgemein ist für  $a_1, \dots, a_s \in \mathbb{R}^n$

$$\text{conv}\{a_1, \dots, a_s\} = \left\{ \sum_{i=0}^s \lambda_i a_i : \lambda_i \geq 0, \sum_{i=0}^s \lambda_i = 1 \right\}$$

die konvexe Hülle der Punkte  $a_1, \dots, a_s$ , das ist die kleinste konvexe Menge, die diese Punkte enthält.

- ii)  $S(0, 0) = c_{0,0}$ ,  $S(1, 0) = c_{m,0}$ ,  $S(0, 1) = c_{0,p}$  und  $S(1, 1) = c_{m,p}$ .

- iii) Sei  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  eine affine Transformation. Dann gilt

$$T(S(s, t)) = \sum_{i=0}^m \sum_{j=0}^p b_{i,m}(s) b_{j,p}(t) T(c_{i,j}).$$

# Index

- $A^\intercal$ , 31
- $I_n$ , 32
- $S_n$ , 16
- $\mathbb{C}$ , 21
- Cov, 106
- $\mathbb{K}^M$ , 36
- $\mathbb{K}^n$ , 35
- $\mathbb{K}^{m \times n}$ , 28
- $\mathbb{N}$ , 8
- $\mathbb{N}_0$ , 14
- $\mathbb{Q}$ , 8
- $\mathbb{R}$ , 8
- $\mathbb{R}_{\geq 0}$ , 13
- $\mathbb{Z}$ , 8
- $\mathbb{Z}_m$ , 19
- mod, 18
- cos, 5
- det(), 49
- $\emptyset$ , 8
- ggT, 18
- i, 21
- Im(), 21
- $\in$ , 8
- $GL(n, \mathbb{R})$ , 56
- $SL(n, \mathbb{R})$ , 56
- $\prod$ , 6
- Re(), 21
- sin, 5
- $\subset$ , 10
- $\subseteq$ , 10
- $\sum$ , 6
- $a \equiv b \pmod m$ , 19
- p-Quantil, 110
- Äquivalenz, 7
- Äquivalenzklasse, 12
- Äquivalenzrelation, 12
  
- Abbildung, 13
  - affin, 41
  - bijektiv, 13
  - Bild, 42
  - Graph, 13
  - injektiv, 13
  - invers, 14
  - Kern, 42
  - Komposition, 14
  - linear, 40
  - surjektiv, 13
  - Verkettung, 14
  
- Ablehnungsbereich, 113
- Ableitung, 72, 93
  - höhere partielle, 93
  - partiell, 92
- Ableitungen
  - hohere, 74
- Absorption, 11
- Affine Abbildung, 41
- Allquantor, 9
- Alternativhypothese, 113
- Assoziativität, 10
- asymptotisch stabil, 125
- asymptotisches Verhalten, 69
- Ausgleichsproblem, 121
- Aussagen, 7
  
- Bézierflächen, 137
- Basis, 38
- Bernsteinpolynome, 132
- Betrag, 5
- Bijektiv, 13
- Bild, 13, 42
- Bildmenge, 13
- Binomischer Lehrsatz, 17
- Boolesche Algebra, 11
- Bézierskurven, 133
  
- Cauchy-Schwarz-Ungleichung, 43
- Charakteristisches Polynom, 51
- Chinesischer Restsatz, 58
- Cholesky-Zerlegung, 121
  
- De Casteljau - Algorithmus, 135
- De Morgan'sche Regeln, 11
- Definitionsbereich, 13
- Determinante, 49
- Dichte, 104
- Differential, 93
- Differentialgleichung
  - autonom, 124
  - gewöhnlich, 124
  - homogen, 125
  - inhomogen, 125
  - linear, 125
  - partiell, 124
  - trennbar, 124
- differenzierbar, 72
  - stetig, 72
- Differenzierbarkeit, 72, 93

Dimension, 38  
 Dimensionsformel, 42  
 Distributivgesetz, 21  
 Distributivität, 10  
 Division mit Rest, 18  
 Dreiecksungleichung, 44  
 Eigenraum, 52  
 Eigenvektor, 51  
 Eigenwert, 51  
 Einbettung, 40  
 Einheitsvektoren, 38  
 Einheitswurzeln, 23  
 Einselement, 21  
 Elementarereignis, 101  
 empirische Kovarianz, 115  
 empirische Varianz, 110  
 empirische Verteilungsfunktion, 114  
 empirischen Median, 114  
 empirischer Anteil, 110  
 empirischer Korrelationskoeffizient, 115  
 Entwicklungspunkt, 77  
 Entwicklungssatz von Laplace, 49  
 Ereignis, 101
 

- disjunkt, 101
- sicheres, 101
- unmögliches, 101

 Ereignisse
 

- unabhängig, 103
- erwartungstreu, 111

 Erwartungswert, 105  
 Erzeugendensystem, 37  
 Euklidischer Algorithmus, 18  
 Eulerverfahren, 130  
 Existenzquantor, 9  
 Exponentialfunktion, 66  
 Exponentialverteilung, 108  
 Extremum
 

- global, 74, 94
- lokal, 74, 94

 Fakultät, 16  
 Fibonacci Zahlen, 16  
 Fixpunkte, 125  
 Fläche, 136
 

- algebraisch, 136

 Folgen, 61
 

- Alternierend, 61
- Beschränkt, 61
- bestimmt divergent, 62
- divergent, 62
- Glieder, 61

 Grenzwert, 61  
 Heron'sche, 63  
 Konvergent, 61  
 Monoton, 61  
 Partialsummen, 63  
 Rekursiv, 61  
 Teil, 62  
 Formel von Moivre, 23  
 Fourierreihe, 88  
 Fundamentalsatz der Algebra, 24  
 Funktion, 13
 

- Grenzwert, 68, 90, 91
- mehrerer Variablen, 90
- vektorwertig, 91

 Funktional-Matrix, 93  
 Funktionen
 

- gerade, 67
- konkav, 74
- konvex, 74
- periodisch, 68
- stetig, 70, 91, 92
- ungerade, 67

 Funktionswert, 13  
 Gammafunktion, 110  
 Gaus'sche Glockenkurve, 108  
 Gaus-Test, 114  
 Gauss-Seidel-Verfahren, 121  
 Gausverteilung, 108  
 Gauß'sche Algorithmus, 27  
 Gauß'sche ganzen Zahlen, 59  
 Gauß'sches Eliminationsverfahren, 27  
 Gegeneignis, 101  
 Geometrische Reihe, 9, 64  
 Geschwindigkeitsvektor, 131  
 Gewöhnliche Differentialgleichung, 124  
 Gleichgewichtslagen, 125  
 Gleichungssystem
 

- homogenes, 33
- inhomogenes, 33

 Gleichverteilung, 107
 

- stetig, 108

 goldene Schnitt, 122  
 Goldener Schnitt, 17  
 Grad, 59  
 Gradient, 93  
 Gram-Schmidt-Orthogonalisierung, 45  
 Graph, 13  
 Grenzwert, 61
 

- Vektoren, 90

 Gruppe, 20, 56
 

- abelsche, 20, 56

Index, 57  
 kommutative, 20, 56  
 Ordnung, 57  
 Unter, 56  
 zyklische, 57  
  
 Harmonische Reihe, 63  
 Hesse-Matrix, 94  
 Homogene Koordinaten, 46  
 Horner-Schema, 117  
 Hurwitz-Kriterium, 53  
 Hyperebene, 36  
 Hypergeometrische Verteilung, 107  
 Hypothese  
     einfach, 113  
     zusammengesetzt, 113  
  
 Ideale, 59  
 Idempotenz, 10  
 identisch verteilt, 107  
 Imaginärteil, 21  
 Imaginare Einheit, 21  
 Implikation, 7  
 implizite Darstellung, 131  
 indefinit, 53, 95  
 Index, 57  
 Infimum, 79  
 Injektiv, 13  
 Integral, 81  
     im  $\mathbb{R}^n$ , 96  
     Majorante, 86  
     Minorante, 86  
     uneigentlich, 85  
 Integrale  
     Vektorwertige, 99  
 Integralvergleichskriterium, 86  
 Integrand, 81  
 Integrationsgrenzen, 81  
 Integrationsvariable, 81  
 integrierbar, 81  
     im  $\mathbb{R}^n$ , 96  
 Interpolationspolynom, 117  
     Lagrange, 117  
     Newton, 117  
 Intervall, 67  
     abgeschlossen, 67  
     offene, 67  
 Inverse Abbildung, 14  
 Inverse Relation, 12  
 Inverses Element, 20, 56  
 irreduzible Polynome, 60  
 Irrtumswahrscheinlichkeit, 111  
  
 Jacobi-Matrix, 93  
 Jacobi-Verfahren, 121  
  
 Körper, 21  
 Kartesisches Produkt, 11  
 Kern, 42  
 Kettenregel, 72  
 Kettregeln, 94  
 Koeffizientenmatrix, 28  
 Kommutativität, 10  
 Komplementregeln, 11  
 Komplexe Zahlen, 21  
 Kondition, 119  
 Konfidenzintervall, 111  
 Konfidenzniveau, 111  
 Kongruente Zahlen, 19  
 Kongruente Polynome, 60  
 Konjugierte Zahl, 22  
 konsistent, 111  
 konsistent im quadratischen Mittel, 111  
 Konvergenz, 61  
 Konvergenzbereich, Konvergenzradius, 79  
 Korper, 59  
 Korrelation  
     negativ, 115  
     positiv, 115  
 Korrelationskoeffizient  
     empirisch, 115  
     Pearson'scher Korrelationskoeffizient, 115  
 Kovarianz, 106  
     empirisch, 115  
 Kugelkoordinaten, 99  
 Kurve, 91  
     algebraisch, 131  
     Länge, 132  
     parameterisierte, 131  
     Reparameterisierung, 132  
  
 Länge, 43  
 Lösungsmenge, 26  
 Lagrange-Multiplikator, 95  
 Laguerre-Polynome, 130  
 Landausymbole, 78  
 Laplace-Experiment, 101  
 Lemma von Bézout, 18  
 Linear abhängig, 37  
 Linear unabhängig, 37  
 Lineare Abbildung, 40  
 Lineare Differentialgleichung, 125  
 Lineare Hülle, 37  
 Lineares Gleichungssystem  
     erweiterte Matrix, 28

Lineares Gleichungssystem, 26  
 Koeffizientenmatrix, 28  
 Linearkombination, 37  
 Logarithmus  
     natürlicher, 6  
 Logarithmusfunktion, 67  
 LU-Faktorisierung, 120  
 Langenfunktion, 132  
 Mächtigkeit  
     endlicher Mengen, 11  
 Majorante, 65  
 Majorantenkriterium, 65  
 Matrix, 28  
     Diagonal-, 32  
     Diagonalelemente einer, 32  
     diagonalsierbar, 53  
     Dreiecks-, 32  
     Einheits-, 32  
     Funktional, 93  
     Hesse, 94  
     indefinit, 95  
     inverse, 33  
     Jacobi, 93  
     Kondition, 119  
     Rang, 42  
     reguläre, 33  
     Spur, 52  
     symmetrische, 32  
     Transponierte, 31  
 Matrixnorm  
     submultiplikativ, 119  
     vertraglich, 119  
 Matrizen  
     ähnlich, 53  
     Addition, 31  
     Multiplikation, 32  
     orthogonale, 45  
     Produkt, 32  
     quadratisch, 32  
     Skalarmultiplikation, 31  
 Maximum, 79  
     global, 74, 94  
     lokal, 74, 94  
 Median, 110  
     empirischen, 114  
 Menge  
     gleichmächtig, 14  
     offen, 92  
 Mengen, 8  
     Differenz, 10  
     Durchschnitt, 10  
     Vereinigung, 10  
 Minimum, 79  
     global, 74, 94  
     lokal, 74, 94  
 Minorante, 65  
 Mittelwertsatz der Differentialrechnung, 75  
 nachiteration, 120  
 Nebenklassen, 57  
 Negativ definit, 53  
 Negativ semi-definit, 53  
 Neutrales Element, 20, 56  
 Newton-Cotes Integrationsformel, 118  
 Norm, 43, 44  
     Normalbereich, 97  
 Normalenvektor, 131, 137  
 Normalverteilung, 108  
     normiertes Polynom, 59  
 Nullfolge, 63  
 Nullhypothese, 113  
 Nullmatrix, 28  
 Nullstellen, 24  
 Obersumme, 81  
     im  $\mathbb{R}^n$ , 96  
 Ordnung, 13  
     partiell, 13  
     total, 13  
 Ordnungsrelation, 13  
 Orthogonale Projektion, 45  
 Orthogonale Matrix, 45  
 Orthogonale Transformation, 45  
 Orthogonale Vektoren, 44  
 Orthonormalbasis, 45  
 Orthonormalsystem, 45  
 Parallelprojektion, 48  
 parametrischen Test, 113  
 Partialsummen, 63  
     partielle Ableitung, 92  
     Partielle Integration, 84  
 Pearson'scher Korrelationskoeffizient, 115  
 Permutation, 16  
 Permutationsmatrix, 120  
 Polardarstellung, 23  
 Polarkoordinaten, 23, 44, 99  
 Polynom, 24, 36, 59  
     Grad, 36, 59  
     normiert, 59  
     trigonometrisches, 87  
 Polynomdivision, 59  
 Polynome

- irreduzible, 60
- Laguerre, 130
- Polynomfunktion, 59
- Polynomkurve, 131
- Polynomring, 59
- Positiv definit, 53
- Positiv semi-definit, 53
- Potenzmenge, 11
- Potenzreihen, 78
- Primzahl, 17
- Primzahlsatz, 17
- Produktregel, 72
- Projektion
  - orthogonal, 40
- QR-Faktorisierung, 121
- Quader, 96
- Quadratische Matrizen, 32
- Quantil, 110
- Quaternionen, 47
- Quotientenkriterium, 65
- Quotientenregel, 72
- Rationale Funktionen, 66
- Realteil, 21
- Realtionen
  - Komposition, 12
- Regressionsgerade, 116
- Reihe
  - absolut konvergent, 64
  - alternierend, 64
  - geometrisch, 64
  - harmonisch, 63
  - Majorantenkriterium, 65
  - Quotientenkriterium, 65
- Reihen, 63
  - divergent, 63
  - konvergent, 63
- Rekursionssatz, 16
- rekursive Vorschrift, 16
- Relation
  - antisymmetrisch, 13
- Relationen, 12
  - binär, 12
  - inverse, 12
  - reflexiv, 12
  - symmetrisch, 12
  - transitiv, 12
- Restglied, 77
- Restklassen, 12
- Restklassenring, 60
- Riemann integrierbar, 81
  - im  $\mathbb{R}^n$ , 96
- Riemann-Integral, 81
  - im  $\mathbb{R}^n$ , 96
- Ring, 58
  - Einselement, 58
  - kommutativ, 58
  - Polynom, 59
- Rotation, 41
- Sattelpunkt, 75
- Scatterplot, 115
- Schiefkörper, 47
- Schwerpunkt, 99
- Schätzer, 110
- Schätzfunktion, 110
- Schätzungen, 110
- Sekante, 72
  - Steigung, 72
- Signifikanzniveau, 113
- Skalare, 26
- Skalarmultiplikation, 35
- Skalarprodukt
  - Standard, 43
- Spaltenindex, 28
- Spiegelung, 40
- Spline, 119
  - kubisch, 119
  - natürlich, 119
- Spur, 52
- Stammfunktion, 82
- Standardabweichung, 105
- Standardnormalverteilung, 108
- stetig differenzierbar, 72
- Stetigkeit, 70, 91, 92
- Stichprobenfunktion, 110
- Stichprobenvariablen, 110
- Stirling-Formel, 16
- Streudiagramm, 115
- strikt diagonaldominant, 121
- Stufenform, 29
- Stützstellen, 117
- Substitutionsregel, 83
- Supremum, 79
- Surjektiv, 13
- Tangentialvektor, 131
- Taylorpolynom, 77
- Taylorreihe, 78
- Teilfolge, 62
- Teilmenge, 10
- Teilraum, 36
- Teilsummen, 63

Transformationssatz, 98  
 Translation, 40  
 Translationen, 41  
 Transponierte einer Matrix, 31  
 trigonometrisches Polynom, 87  
 Umgebung, 92  
 Umkehrfunktion  
     Ableitung, 73  
 Uneigentliche Integrale, 85  
 Untergruppe, 56  
 Untergruppenkriterium, 56  
 Untersumme, 81  
     im  $\mathbb{R}^n$ , 96  
 Untervektorraum, 36  
 Urbildmenge, 13  
 Urnenmodelle, 102  
 Varianz, 105  
 Vektoraddition, 35  
 Vektoren, 35  
 Vektorraum, 35  
     Dimension, 38  
     endlichdimensional, 38  
     unendlichdimensional, 38  
     Unter-, 36  
 Verkettung von Abbildungen, 14  
 Verknüpfungsregeln für Mengen, 10  
 Verteilung, 104  
      $\chi^2$ , 110  
     hypergeometrisch, 107  
 Verteilungsfunktion, 104  
     empirisch, 114  
 Vertrauensintervall, 111  
 Vertrauensniveau, 111  
 Verwerfungsbereich, 113  
 Vielfachheit  
     algebraisch, 53  
     geometrisch, 53  
 Vollständige Induktion, 9  
 Volumen, 96  
 Wahrscheinlichkeit, 101, 102  
     bedingte, 103  
 Wahrscheinlichkeitsraum, 102  
 Wendepunkt, 75  
 Wertebereich, 13  
 Zeilenindex, 28  
 Zeilenstufenform, 29  
 Zentralprojektion, 48  
 Zerlegung, 81  
     im  $\mathbb{R}^n$ , 96  
 Zufallsexperiment, 101  
 Zufallsstichprobe, 110  
 Zufallsvariable, 103  
     diskret, 104  
     standardisiert, 106  
     stetige, 104  
     unabhängig, 105  
 Zyklische Gruppe, 57  
 Zylinderkoordinaten, 99