Mathematik für Informatiker

Prof. Dr. Martin Henk

http://fma2.math.uni-magdeburg.de/~henk/lectures/math4inf/index.html

http://shamash.math.uni-magdeburg.de/gilgamesh

Inhaltsverzeichnis

Teil I

Mathematatik 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).

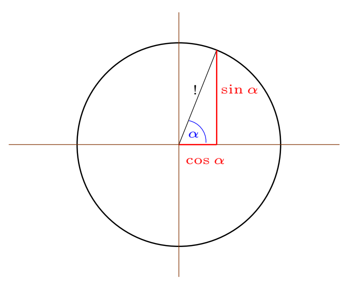


Abbildung 1: Sinus, Cosinus am Einheitskreis, d.h., am Kreis mit Mittelpunkt 0 und Radius 1.

Bemerkung 0.4 (Additionstheoreme)**.**

i) $\sin(\alpha + 2\pi) = \sin \alpha$, $\cos(\alpha + 2\pi) = \cos \alpha$,

ii) $\sin(-\alpha) = -\sin \alpha$, $\cos(-\alpha) = \cos \alpha$,

iii) $\sin^2 \alpha + \cos^2 \alpha = 1$, (Phytagoras)[1]

iv) $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$ *und* $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$.

[1] Pythagoras von Samos, 582–507

Notation 0.5. Zur Abkürzung längerer Summen oder Produkte schreibt man

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n,$$
$$\prod_{i=1}^{n} a_i = a_1 \cdot a_2 \cdot \ldots \cdot a_n.$$

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

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

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

i) $a^n = a \cdot a \cdot \ldots \cdot a$ (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

$$a^{\frac{m}{n}} = \sqrt[n]{a^m} = \left(\sqrt[n]{a}\right)^m,$$
$$a^{-\frac{m}{n}} = \frac{1}{a^{\frac{m}{n}}},\ a \neq 0.$$

Bemerkung 0.10 (Potenzgesetze)**.** *Für $x, y \geq 0$ und reelle Zahlen α, β 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) $\frac{x^\alpha}{y^\alpha} = \left(\frac{x}{y}\right)^\alpha$, $y \neq 0$,

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[2] Zahl $\mathrm{e} = 2,71828182...$, dann heißt

$$\ln a = \log_{\mathrm{e}} a$$

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) $a^{\log_b c} = c^{\log_b a}$.

[2] Leonhard Euler, 1707 – 1783

1 Elementare Logik und Mengen

Definition 1.1 (Aussagen[3]). 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. Dann lassen sich durch Negation (Verneinung), Konjunktion (`und`-Verknüpfung) und Disjunktion (`oder`-Verknüpfung) neue Aussagen bilden:

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

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

A	$\overline{A}$
w	f
f	w

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Definition 1.3 (Implikation). Seien A und B Aussagen. Die Aussage *„wenn A, dann 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 genau dann, wenn B "*, *„A dann und nur dann, wenn B "*, *„A ist äquivalent zu B"* oder *„A ist notwendig und hinreichend für B"*.

[3] Aristoteles, 384 – 322 v.Chr.

Bemerkung 1.5. *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) $(A \Rightarrow B) \Leftrightarrow (\overline{B} \Rightarrow \overline{A}) \Leftrightarrow \overline{A \wedge \overline{B}}$, iv) $(A \Leftrightarrow B) \Leftrightarrow (\overline{A} \Leftrightarrow \overline{B})$.

iii) besagt insbesondere, dass "aus A folgt B" ist genau dann wahr, wenn "aus Nicht-A folgt Nicht-B" wahr ist und dies gilt genau dann, wenn "A und Nicht-B" falsch ist.

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

$$\begin{aligned} &\text{i)} \quad ((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C), \\ &\text{ii)} \quad ((A \Leftrightarrow B) \wedge (B \Leftrightarrow C)) \Rightarrow (A \Leftrightarrow C). \end{aligned}$$

i) sagt insbesondere: wenn A die Aussage B impliziert, und wenn B die Aussage C impliziert, dann impliziert (auch) A die Aussage C.

Bemerkung 1.7. *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 *A wird die* Behauptung *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 1.5 iii)).*

iii) *Beweis einer Äquivalenz $A \Leftrightarrow B$: Man beweist die Aussageformen $A \Rightarrow B$ und $B \Rightarrow A$.*

Definition 1.8 (Mengen[4])**.** 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 1.9. 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 1.10. Mengen lassen sich auf unterschiedliche Arten beschreiben, z.B.

i) Aufzählung der Elemente, z.B., $M = \{1, 2, 3, 4, 5\}$ oder $M = \{\circ, \diamond, \star\}$.

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.,

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

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*.

[4] Georg Cantor, 1845–1918

vii) $\emptyset = \{\}$ bezeichnet die *leere Menge.*

Bemerkung 1.11 (Russellsche Antinomie[5]). *Sei M die "Menge" aller Mengen, die sich nicht selbst enthalten, d.h.,*

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

Ist $M \in M$? Falls "Ja", dann muss M die Bedingung $M \notin M$ erfüllen! Falls "Nein", dann gilt nach Definition von M aber $M \in M$! Also, im Sinne unserer Definition ist M keine Menge!

Definition 1.12 (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 1.13. *Seien $A(x)$, $x \in M$, Aussagen. Es gilt:*

$$\overline{\forall x : A(x)} \Leftrightarrow \exists\, x : \overline{A(x)} \qquad \textit{und} \qquad \overline{\exists\, x : A(x)} \Leftrightarrow \forall x : \overline{A(x)}.$$

i) besagt, dass die Aussage "für alle $x \in M$ gilt $A(x)$" genau dann falsch ist, wenn eine der Aussagen $A(x)$ falsch ist. ii) besagt, die Aussage "es gibt ein x für das $A(x)$ gilt" genau dann falsch ist, falls $A(x)$ falsch ist für alle $x \in M$.

Erklärung 1.14 (Vollständige Induktion). Sei $M = \{k, k+1, k+2, \dots\}$, wobei $k \in \mathbb{Z}$ 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 den kleinsten Wert $n = k$ gilt, man zeigt also $A(k)$.

ii) *Induktionsschritt*: Man zeigt, dass die Implikation $A(n) \Rightarrow A(n+1)$ wahr ist.

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

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

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

i) (Induktionsanfang) Wir zeigen, dass $A(0)$ wahr ist:

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

ii) (Induktionsschritt) Wir zeigen: "$A(n) \Rightarrow A(n+1)$:", d.h. $A(n+1)$ folgt aus $A(n)$.

Da $A(n)$ gilt ist

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

[5] Bertrand Russell, 1872–1970

Addition von x^{n+1} auf beiden Seiten liefert

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

Zusammenfassen auf der linken und rechten Seite führt zu

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

Dies ist gerade die Aussage $A(n + 1)$.

Oder alternativ:

ii) (Induktionsschritt)

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

Beispiel 1.16. 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 $s_1, \ldots, s_{n+1}$ bezeichnen. Sei z.B. s_2 ein Mathematikstudent. $A(n)$ angewendet auf die Gruppe $\{s_1, \ldots, s_n\}$ und $\{s_2, \ldots, s_{n+1}\}$ zeigt, dass $s_1, \ldots, s_n$ und $s_2, \ldots, s_{n+1}$ Mathematik studieren; also alle und somit gilt $A(n + 1)$.

Also gilt $A(n)$ für jede natürlich Zahl n? Wo ist der Fehler?

Definition 1.17 (Teilmenge)**.**

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

ii) 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 $A \neq B$.

iii) Eine Menge B heißt *echte Teilmenge* einer Menge A (Schreibweise: $B \subset A$), wenn $B \subseteq A$ und $B \neq A$.

Bemerkung 1.18. *Die leere Menge ist Teilmenge einer jeden Menge.*

Definition 1.19 (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.

iv) 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 1.20 (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:
$$A \cap \emptyset = \emptyset, \quad A \cup \emptyset = A, \qquad A \cup U = U, \quad A \cap U = A.$$

vii) Komplementregeln:
$$A \cap \overline{A} = \emptyset, \quad A \cup \overline{A} = U, \quad \overline{\overline{A}} = A.$$

viii) De Morgan'sche Regeln[6]:
$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

Definition 1.21 (Boolesche Algebra[7])**.** Eine Menge, in der drei Operationen, z.B. $(\cap, \cup, ^-)$ oder $(\wedge, \vee, ^-)$, definiert sind, die den Eigenschaften i)-vii) aus Satz 1.20 genügen, wird als *Boolesche Algebra* bezeichnet.

Definition 1.22 (Kartesisches Produkt[8])**.**

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* oder auch *2-Tupel*.

ii) Allgemein ist das kartesische Produkt $X_1 \times X_2 \times \ldots \times X_n$ von n-Mengen $X_1, \ldots, X_n$ definiert als
$$X_1 \times X_2 \times \cdots \times X_n = \{(x_1, x_2, \ldots, x_n) : x_i \in X_i,\ 1 \leq i \leq n\}.$$
$(x_1, x_2, \ldots, x_n) \in X_1 \times X_2 \times \ldots \times X_n$ heißt *n-Tupel*.

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

[6] Augustus De Morgan, 1806–1871
[7] George Boole, 1815–1864
[8] René Descartes, 1596–1650

Bemerkung 1.23. *Für* $(x_1, \dots, x_n), (x'_1, \dots, x'_n) \in X_1 \times \cdots \times X_n$ *gilt*

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

Definition 1.24 (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 1.25. *<2-> Seien* A, B *endliche Mengen. Dann gilt*

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

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

Definition 1.26. <4->[Potenzmenge] Sei A Menge. Die Menge $P(A) = \{X : X \subseteq A\}$ aller Teilmengen von A heißt *Potenzmenge* von A.

Lemma 1.27. *<5-> Sei* A *eine endliche Menge. Dann gilt*

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

2 Relationen und Abbildungen

Definition 2.1 (Relationen)**.**

i) Eine *binäre Relation* R zwischen zwei 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 $A_1, \ldots, A_n$ ist analog eine Teilmenge von $A_1 \times A_2 \times \cdots \times A_n$.

Definition 2.2 (Verkettung/Komposition von Relationen)**.** Sei R eine Relation zwischen den 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 insbesondere 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*

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

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. <3->[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.* $[x]_R$ ist also die Menge aller $y \in A$, die äquivalent zu x sind.

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

i) $A = \bigcup_{x \in A} [x]_R$,

ii) *Für zwei Äquivalenzklassen $[x]_R$, $[y]_R$ 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 auf $\mathbb{Z}$, und es gilt

i) Für $z \in \mathbb{Z}$ besteht $[z]_{R_m}$ aus allen Zahlen, die bei Division durch m den gleichen Rest lassen wie z. Hierbei nehmen wir stets an, dass der Rest bei Division nicht negativ ist. Man nennt $[z]_{R_m}$ auch *Restklasse*, und man bezeichnet sie einfach mit $[z]_m$.

ii) $\mathbb{Z}$ ist die Vereinigung der paarweise disjunkten Äquivalenzklassen $[0]_m, [1]_m, \ldots, [m-1]_m$.

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

i) R heißt *antisymmetrisch*, falls für alle $x, y \in A$ mit $(x, y) \in R$ und $(y, x) \in R$ gilt: $x = y$.

ii) Eine Relation R auf A, die reflexiv, antisymmetrisch und transitiv ist, heißt *Ordnung(srelation)* auf A.

iii) 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.

i) $\leq$ ist eine totale Ordnung auf $\mathbb{R}$, da für alle $x, y \in \mathbb{R}$ gilt entweder $x \leq y$ oder $y \leq x$.

ii) $\subseteq$ ist eine partielle Ordnung z.B. auf der Potenzmenge von $\{1, 2, 3\}$. Es ist keine totale Ordnung, da z.B. $\{1, 2\}$ und $\{1, 3\}$ nicht in Relation stehen.

Definition 2.11 (Funktion/Abbildung)**.** Eine *Funktion (Abbildung)* 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:

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

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)**.** <2-> Sei $f : X \to Y$ mit $x \mapsto f(x)$ eine Funktion. Die Relation (Menge)

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

heißt der *Graph* der Abbildung.

Definition 2.13 (Bildmenge, Urbildmenge)**.** Sei $f : X \to Y$ mit $x \mapsto f(x)$ eine Funktion. 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 einelementige Mengen B, also $B = \{y\}$, $y \in Y$, schreibt man auch einfach $f^{-1}(y)$.

Beispiel 2.14. Sei $f : \mathbb{R} \to \mathbb{R}$ 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 $f^{-1}(\{2, 4\}) = \{\pm\sqrt{2}, \pm 2\}$.
- $f^{-1}(\{-1\}) = f^{-1}(-1) = \emptyset$.

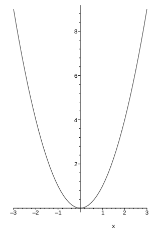


Abbildung 2: Der Graph Γ_f

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

i) 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.

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

Definition 2.17 (Verkettung (Komposition) von Funktionen)**.** Seien $f : X \to Y$ und $g : X' \to Y'$ Funktionen mit $f(X) \subseteq X'$. Die *Verkettung (Komposition) von „g nach f“* ist die Funktion $g \circ f : X \to Y'$ mit $x \mapsto (g \circ f)(x) = g(f(x))$.

Bemerkung 2.18. *Die Verkettung von Abbildungen entspricht der Verkettung von Relationen:* $\Gamma_{g \circ f} = \Gamma_g \circ \Gamma_f$.

Definition 2.19 (Identische Funktion)**.** Sei X Menge. Die Funktion $\mathrm{id}_X : X \to X$ mit $x \mapsto x$ heißt *identische Funktion* (oder auch einfach *Identität*) auf X.

Definition 2.20 (Inverse Funktion)**.** Sei $f : X \to Y$ bijektiv. Die Funktion $g : Y \to X$, die jedem $y \in Y$ das eindeutig bestimmte $x \in X$ mit $y = f(x)$ zuordnet, heißt *inverse Funktion*, und man schreibt für g auch f^{-1}.

Bemerkung 2.21. *Die inverse Abbildung entspricht der inversen Relation:* $\Gamma_{f^{-1}} = \Gamma_f^{-1}$.

Lemma 2.22. *Sei $f : X \to Y$ bijektiv. Dann gilt:*

$$\text{i) } f^{-1} \circ f = \mathrm{id}_X, \quad \textit{d.h. } (f^{-1} \circ f)(x) = x, \ \textit{für alle } x \in X,$$
$$\text{ii) } f \circ f^{-1} = \mathrm{id}_Y, \quad \textit{d.h. } (f \circ f^{-1})(y) = y, \ \textit{für alle } y \in Y.$$

Sei $g : Y \to Z$ eine weitere bijektive Abbildung. Für die Umkehrfunktion von $g \circ f : X \to Z$ gilt: $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Definition 2.23.

1. Eine Funktion der Form

$$p(x) = a_m\, x^m + a_{m-1}\, x^{m-1} + \cdots + a_1 x + a_0, \quad a_j \in \mathbb{R},$$

heißt Polynom vom Grad m (falls $a_m \neq 0$).

2. Seien $p(x), q(x)$ Polynome. Dann heißt $\frac{p(x)}{q(x)}$ *rationale Funktion.*

Definition 2.24 (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 2.25.

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

Definition 2.26 (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 2.27.

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

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

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

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

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[9]-Formel)**.**

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

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

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

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

Bemerkung 3.6.

i) *Die* Fibonacci[10]*-Zahlen (Folge): Sei $F : \mathbb{N}_0 \to \mathbb{N}_0$ mit*

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

ii) *Sei $a = (1 + \sqrt{5})/2$ (goldener Schnitt) und $b = (1 - \sqrt{5})/2$. Für die n-te Fibonacci-Zahl $F(n)$ gilt $F(n) = (1/\sqrt{5})(a^n - b^n)$.*

iii) Think recursively!

iv) *Türme von Hanoi, siehe* `http://en.wikipedia.org/wiki/Tower_of_Hanoi`

Definition 3.7 (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“.

[9] James Stirling, 1692–1770

[10] Leonardo Pisano, 1170(80)-1250

Satz 3.8.

1. $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ *für* $n \geq 1$.
2. $\binom{n}{k} = \binom{n}{n-k}$.

Satz 3.9. *<4-> Die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge beträgt* $\binom{n}{k}$.

Beispiel 3.10.

i) $\binom{4}{0} = 1$, $\binom{4}{1} = 4$, $\binom{4}{2} = 6$, $\binom{4}{3} = 4, \binom{4}{4} = 1$.

ii) 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 3.11 (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 3.12.

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

Definition 3.13 (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 3.14 (Primfaktorzerlegung)**.** *Jede natürliche Zahl größer als* 1 *lässt sich eindeutig (bis auf die Reihenfolge) als Produkt von Primzahlen schreiben.*

Satz 3.15 (Euklid[11])**.** *Es gibt unendlich viele Primzahlen.*

Bemerkung 3.16. *Der Primzahlsatz[12] sagt aus, dass*

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

Notation 3.17. 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 3.18 (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 $\mathrm{ggT}(a, b)$ bezeichnet.

Lemma 3.19. *Seien $a, b \in \mathbb{Z}$ und $m \in \mathbb{Z}$.*

i) $\mathrm{ggT}(a, b) = \mathrm{ggT}(b, a) = \mathrm{ggT}(-a, b) = \mathrm{ggT}(a, -b) = \mathrm{ggT}(-a, -b)$.

ii) $\mathrm{ggT}(a, b) = \mathrm{ggT}(a + m \cdot b, b) = \mathrm{ggT}(a, b + m \cdot a)$.

[11] Euklid, ca. 300 v.Chr.

[12] Jacques Hadamard, 1865– 1963; Charles de la Vallée Poussin, 1866–1962

Lemma 3.20 (Division mit Rest)**.** *Sei $a, b \in \mathbb{Z}$. 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 mod b“).*

Bemerkung 3.21. *Für $a, b \in \mathbb{Z}$ ist $\mathrm{ggT}(a, b) = \mathrm{ggT}(a \bmod b, b)$.*

Satz 3.22 (Euklidischer Algorithmus)**.** *Seien $a, b \in \mathbb{N}$ und sei $a \geq b$. Das folgende Verfahren, der sogenannte* Euklidische Algorithmus*, berechnet (rekursiv) $\mathrm{ggT}(a, b)$.*

(1) *Berechne $r = a \mod b$.*

(2) *Ist $r = 0$, dann $\mathrm{ggT}(a, b) = b$ und STOP.*

(3) *Rufe das Verfahren auf für $a = b$ und $b = r$, d.h. berechne $\mathrm{ggT}(b, a \bmod b)$.*

Beispiel 3.23. Berechnen von $\mathrm{ggT}(29393, 2805)$.

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

$$29393 = 10 \cdot 2805 + 1343 \tag{3.1}$$
$$2805 = 2 \cdot 1343 + 119 \tag{3.2}$$
$$1343 = 11 \cdot 119 + 34 \tag{3.3}$$
$$119 = 3 \cdot 34 + 17 \tag{3.4}$$
$$34 = 2 \cdot 17 + 0$$

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

Lemma 3.24 (Lemma von Bézout[13])**.** *Seien $a, b \in \mathbb{Z}$. Dann gibt es $x, y \in \mathbb{Z}$, so dass*

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

Bemerkung 3.25. *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. Dieses Verfahren nennt man auch* erweiterter Euklidischer Algorithmus

Beispiel 3.26 ($\mathrm{ggT}(29393, 2805)$ (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)

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

[13] Étienne Bézout; 1730 – 1783

Satz 3.27 (Modulo Rechnen)**.** *Seien* $a, b \in \mathbb{Z}$ *und* $m \in \mathbb{N}$*. Dann gilt*

1. $(a + b) \bmod m = ((a \bmod m) + (b \bmod m)) \bmod m.$
2. $(a \cdot b) \bmod m = ((a \bmod m)(b \bmod m)) \bmod m.$

Beispiel 3.28.

$$\begin{aligned}
2^{26} \bmod 27 &= (2^5 \bmod 27)^5 (2 \bmod 27) \bmod 27 = (5^5 * 2) \bmod 27 \\
&= (5^3 \bmod 27)((2 * 5 * 5) \bmod 27) \bmod 27 \\
&= ((17 * 2) \bmod 27)((5 * 5) \bmod 27) \quad \bmod 27 \\
&= ((7 * 5) \bmod 27)(5 \bmod 27) \bmod 27 \\
&= 40 \bmod 27 = 13.
\end{aligned}$$

Definition 3.29 (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“.

Bemerkung 3.30. *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 3.31. *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 3.32 (Kleiner Satz von Fermat[14])**.** *Sei* p *eine Primzahl und sei* $a \in \mathbb{N}$ *mit* $\operatorname{ggT}(a, p) = 1$*. Dann ist* $a^{p-1} \equiv 1 \bmod p$*.*

Folgerung 3.33. *Zum Testen, ob eine gegebene Zahl* $\overline{p} \in \mathbb{N}$ keine *Primzahl ist, kann man wie folgt vorgehen: Man suche sich eine Zahl* a *mit* $\operatorname{ggT}(a, \overline{p}) = 1$*. Ist dann* $a^{\overline{p}-1} \not\equiv 1 \bmod \overline{p}$*, dann ist nach dem Kleinen Satz von Fermat* $\overline{p}$ *keine Primzahl. Ist zum Beispiel* $\overline{p}$ *ungerade, so kann man* $a = 2$ *wählen.*

Beispiel 3.34. $\overline{p} = 27, a = 2$. Nach Beispiel (3.28) ist $2^{26} \equiv 13 \bmod 27$ und somit ist 27 keine Primzahl.

Notation 3.35. Sei $m \in \mathbb{N}$, $m \geq 2$.

i) Sei
$$\mathbb{Z}_m = \{[0]_m, [1]_m, \ldots, [m-1]_m\}$$
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 3.36. *Aufgrund von Satz 3.31 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* $\mathbb{Z}_m$*.*

[14]Pierre de Fermat, ca. 1605 –1665

Bemerkung 3.37. *Seien $a \in \mathbb{Z}$, $m \in \mathbb{N}$ mit $\mathrm{ggT}(a,m) = 1$. Nach Lemma 3.24 gibt es $x, y \in \mathbb{Z}$ mit $a\,x + m\,y = 1$. Das is äquivalent zu $[a]_m \odot [x]_m = [1]_m$.*

Beispiel 3.38 (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 3.39 (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 3.40.

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 3.41 (Gruppe)**.** Sei G eine nichtleere Menge, und sei $\otimes : G \times G \to G$ 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*[15]*) Gruppe*.

Beispiel 3.42.

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 $\mathrm{id}_{\{1,\ldots,n\}}$, 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 3.43 (Körper)**.** Sei $\mathbb{K}$ eine nichtleere Mengen, und seien $+, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ 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 3.44.

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

Satz 3.45. *$(\mathbb{Z}_m, \oplus, \odot)$ ist genau dann ein Körper, wenn m Primzahl ist.*

Definition 3.46 (Komplexe Zahlen)**.** Die Menge

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

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

$$\mathrm{i}^2 = \mathrm{i} \cdot \mathrm{i} = -1$$

und heißt *imaginäre Einheit*.

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

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

[15] Niels Abel, 1802–1829

Bemerkung 3.47.

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* $\mathrm{i}^2 = -1$, *dann ergeben sich Addition und Multiplikation von komplexen Zahlen unmittelbar aus den entsprechenden Operationen für reelle Zahlen.*

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

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

iii) *Weiterhin ist für* $z = x + y\,\mathrm{i} \neq 0$ $(= 0 + 0\,\mathrm{i})$

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

Satz 3.48. $(\mathbb{C}, +, \cdot)$ *ist ein Körper mit neutralem Element* 0 $(= 0 + 0\,\mathrm{i})$ *und Einselement* 1 $(= 1 + 0\,\mathrm{i})$.

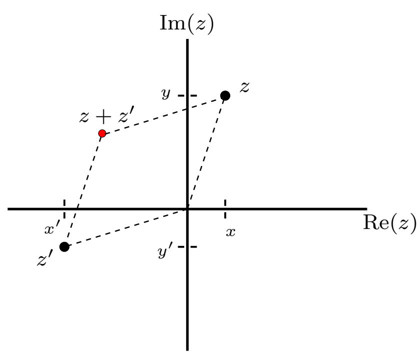


Abbildung 3: Addition von komplexen Zahlen in der Gaußsche Zahlenebene

Notation 3.49 (Gaußsche[16] Zahlenebene)**.**

Definition 3.50 (Konjugierte Zahl, Betrag)**.** Sei $z = x + y\,\mathrm{i} \in \mathbb{C}$.

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

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

Satz 3.51. *Seien* $z, z' \in \mathbb{C}$.

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

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

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

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

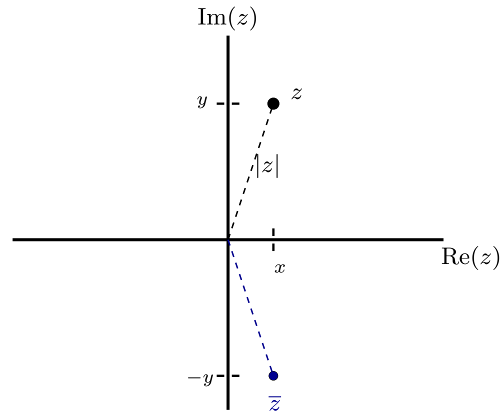


Abbildung 4: Konjugierte komplexe Zahl und Betrag von komplexen Zahlen in der Gaußsche Zahlenebene

Definition 3.52 (Polardarstellung)**.** Jede komplexe Zahl $z = x + y\,\mathrm{i} \in \mathbb{C}$, $z \neq 0$, lässt sich eindeutig in der Form

$$z = r\,(\cos\phi + \,\mathrm{i}\sin\phi) \tag{3.5}$$

darstellen. Dabei ist $r = |z|$, und ϕ 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 *Polardarstellung* von z, und (r, ϕ) nennt man *Polarkoordinaten* von z. Für die Polardarstellung benutzt man oft auch abkürzend die Eulersche Exponentialdarstellung

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

d.h. $\mathrm{e}^{\,\mathrm{i}\phi} = \cos\phi + \,\mathrm{i}\sin\phi$.[17]

Bemerkung 3.53. *Sei $z = r(\cos\phi + \,\mathrm{i}\sin\phi)$ und $z' = r'(\cos\phi' + \,\mathrm{i}\sin\phi') \in \mathbb{C}$. Dann ist (vgl. Bemerkung 0.4).*

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

bzw. in der Eulersche Exponentialdarstellung $z = r\,\mathrm{e}^{\,\mathrm{i}\phi}$ und $z' = r'\,\mathrm{e}^{\,\mathrm{i}\phi'}$ ergibt sich

$$z \cdot z' = r\,\mathrm{e}^{\,\mathrm{i}\phi}\,r'\,\mathrm{e}^{\,\mathrm{i}\phi'} = r\,r'\,\mathrm{e}^{\,\mathrm{i}(\phi+\phi')}$$

(s. Abb.5).

Lemma 3.54 (Formel von Moivre[18])**.** *Für $z = r\,(\cos\phi + \,\mathrm{i}\sin\phi) = r\,\mathrm{e}^{\,\mathrm{i}\phi}$ und $n \in \mathbb{N}$ gilt:*

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

[16] Carl Friedrich Gauß (Gauss) 1777–1855
[17] Zur schönsten Formel wurde gewählt $\mathrm{e}^{\,\mathrm{i}\pi} + 1 = 0$.
[18] Abraham de Moivre, 1667–1754

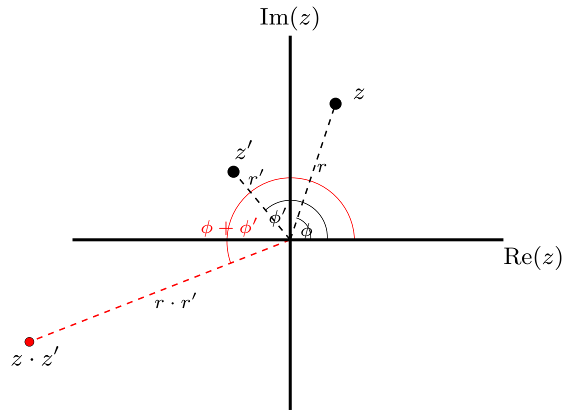


Abbildung 5: Multiplikation von komplexen Zahlen in Polardarstellung

Definition 3.55 (Einheitswurzeln)**.** Für $n \in \mathbb{N}$ heißen

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

n-te Einheitswurzeln.

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

Beispiel 3.56.

1. $n = 1$: $\omega_0^1 = 1$.
2. $n = 2$: $\omega_0^2 = 1$, $\omega_1^2 = -1$.
3. $n = 3$: $\omega_0^3 = 1$, $\omega_1^3 = -\frac{1}{2} + \frac{\sqrt{3}}{2}\,\mathrm{i}$, $\omega_2^3 = -\frac{1}{2} - \frac{\sqrt{3}}{2}\,\mathrm{i}$.
4. $n = 4$: $\omega_0^4 = 1$, $\omega_1^4 = \mathrm{i}$, $\omega_2^4 = -1$, $\omega_3^4 = -\,\mathrm{i}$.

Bemerkung 3.57. *Die n-ten Einheitswurzeln bilden in der Gaußschen Zahlenebene die Ecken eines regulären n-Ecks.*

Bemerkung 3.58. *Sei $z = r\,(\cos\phi + \mathrm{i}\,\sin\phi) = r\,\mathrm{e}^{\mathrm{i}\phi} \in \mathbb{C}$ und $n \in \mathbb{N}$. Für*

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

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

Notation 3.59. Sei $z = r\,(\cos\phi + \mathrm{i}\,\sin\phi)$. Als „(Quadrat)-Wurzel“ von z, im Zeichen $\sqrt{z}$, versteht man

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

Beispiel 3.60.

1. $\sqrt{-1} = \mathrm{i}$, $\sqrt{i} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\,\mathrm{i}$.
2. Für $x \in \mathbb{R}$, $x \leq 0$, ist $\sqrt{x} = \sqrt{|x|}\,\mathrm{i}$.

Satz 3.61 (Fundamentalsatz der Algebra)**.** *Jede Polynomfunktion $f : \mathbb{C} \to \mathbb{C}$ der Form*

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

mit $a_n \neq 0$ besitzt eine Darstellung

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

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

Beispiel 3.62.

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. $z^2 + 1 = (z - \mathrm{i})(z + \mathrm{i})$, $z^4 - 1 = (z - 1)(z - \mathrm{i})(z + 1)(z + \mathrm{i})$.

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

Satz 3.63. *Sei $f : \mathbb{C} \to \mathbb{C}$ mit $f(x) = a_n\, x^n + a_{n-1}\, x^{n-1} + \ldots + a_1\, x + a_0$, wobei $a_i \in \mathbb{R}$ und $a_n \neq 0$.*

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

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

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 $x_1, \ldots, x_n$ hat die Form

$$\begin{aligned} a_{1\,1}\,x_1 + a_{1\,2}\,x_2 &+ \cdots + a_{1\,n}\,x_n = b_1 \\ a_{2\,1}\,x_1 + a_{2\,2}\,x_2 &+ \cdots + a_{2\,n}\,x_n = b_2 \\ \vdots \qquad \vdots \qquad & \qquad\quad \vdots \quad = \vdots \\ a_{m\,1}\,x_1 + a_{m\,2}\,x_2 &+ \cdots + a_{m\,n}\,x_n = b_m. \end{aligned} \tag{4.1}$$

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

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

heißt *Lösungsmenge* des Gleichungssystems.

Beispiel 4.3.

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

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

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 $(x_1, x_2) = (0, 0)$.

iii) 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 $([1]_2, [1]_2)$.

iv) Das lineare Gleichungssystem über $\mathbb{R}$

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

besitzt die unendlich vielen Lösung $\{(-2 - 2\lambda, \lambda, -1, 3) : \lambda \in \mathbb{R}\}$.

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

i) *zwei Gleichungen vertauscht,*

ii) *das λ-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 Lösungsmnenge direkt ablesen kann.

Beispiel 4.5.

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

$$\begin{aligned} x_1 + 2\,x_2 \qquad\qquad & + \ \ x_4 = 1 \\ x_1 + 2\,x_2 + 2\,x_3 \quad & +3\,x_4 = 5 \\ 2\,x_1 + 4\,x_2 \qquad\qquad & +3\,x_4 = 5 \\ 3\,x_3 & + 2\,x_4 = 3 \end{aligned}$$

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

$$\begin{aligned} x_1 + 2\,x_2 \qquad + \ \ x_4 &= 1 \\ 2\,x_3 + 2\,x_4 &= 4 \\ x_4 &= 3 \\ 3\,x_3 + 2\,x_4 &= 3 \end{aligned}$$

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

$$\begin{aligned} x_1 + 2\,x_2 \qquad + \ \ x_4 &= 1 \\ 2\,x_3 + 2\,x_4 &= 4 \\ x_4 &= 3 \\ -x_4 &= -3 \end{aligned}$$

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

$$\begin{aligned} x_1 + 2\,x_2 \qquad + \ \ x_4 &= 1 \\ 2\,x_3 + 2\,x_4 &= 4 \\ x_4 &= 3 \\ 0 &= 0 \end{aligned}$$

4. Hieraus lassen sich die Lösungen sofort ablesen:
   - 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 $x_1 + 2\,x_2 = -2$.
   - Also ist

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

Definition 4.6 (Matrix)**.**

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

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \cdots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m\,1} & a_{m\,2} & \cdots & a_{m\,n} \end{pmatrix}, \text{ kurz } A = (a_{i\,j})_{i,j=1}^{m,n} \text{ oder } A = (a_{i\,j}),$$

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

$$\begin{aligned} a_{1\,1}\,x_1 + a_{1\,2}\,x_2 \ &+ \cdots + a_{1\,n}\,x_n = b_1 \\ a_{2\,1}\,x_1 + a_{2\,2}\,x_2 \ &+ \cdots + a_{2\,n}\,x_n = b_2 \\ \vdots \qquad \vdots \qquad & \qquad\quad \vdots \quad = \vdots \\ a_{m\,1}\,x_1 + a_{m\,2}\,x_2 &+ \cdots + a_{m\,n}\,x_n = b_m \end{aligned} \tag{4.2}$$

über $\mathbb{K}$ heißt

i)

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \cdots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m\,1} & a_{m\,2} & \cdots & a_{m\,n} \end{pmatrix} \in \mathbb{K}^{m\times n}$$

die *Koeffizientenmatrix* des Gleichungssystems.

ii) Die Matrix

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

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 λ-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:

$$\begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 3 & 5 \\ 2 & 4 & 0 & 3 & 5 \\ 0 & 0 & 3 & 2 & 3 \end{pmatrix} \overset{A_{1,2}(-1), A_{1,3}(-2)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 2 & 3 \end{pmatrix}$$

$$\overset{A_{2,4}(-3/2)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 & -3 \end{pmatrix} \overset{A_{3,4}(1), M_2(1/2)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

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:

$$A = \left(\begin{array}{cccccccccccccccccc}
0 & \cdots & 0 & 1 & * & \cdots & * & * & \cdots & * & * & \cdots & * & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & * & \cdots & * & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & * & \cdots & \cdots & \cdots & * \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 1 & * & \cdots & * \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots
\end{array}\right).$$

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

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

i) Für $i > r$, $1 \leq j \leq n$, ist $a_{i\,j} = 0$,

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

$k_1, \ldots, k_r$ 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 $A' =$*

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

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

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')

$$0\,x_1 + 0\,x_2 + \cdots + 0\,x_n = 1.$$

Also hat das Gleichungssystem keine Lösung.

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

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

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

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

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 hat Gleichungssystem genau eine Lösung.*

b) *Ist $r < n$, dann gibt es mehrere Lösungen.*

Bemerkung 4.16. *Ein Gleichungssystem über den Körper $\mathbb{R}$, bzw. über einem Körper $\mathbb{K}$ mit unendlich vielen Elementen, hat* keine, genau eine, *oder* unendlich viele *Lösungen.*

Definition 4.17 (Addition von Matrizen)**.** Seien $A, B \in \mathbb{K}^{m\times n}$ mit $A = (a_{i\,j})$ und $B = (b_{i\,j})$. Dann ist die *Addition von A und B*, geschrieben $A + B$, definiert als

$$A + B = \begin{pmatrix} a_{1\,1} + b_{1\,1} & a_{1\,2} + b_{1\,2} & \cdots & a_{1\,n} + b_{1\,n} \\ a_{2\,1} + b_{2\,1} & a_{2\,2} + b_{2\,2} & \cdots & a_{2\,n} + b_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m\,1} + b_{m\,1} & a_{m\,2} + b_{m\,2} & \cdots & a_{m\,n} + b_{m\,n} \end{pmatrix}$$

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

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

Definition 4.18 (Multiplikation mit einem Skalar)**.** Sei $A \in \mathbb{K}^{m\times n}$ und $\lambda \in \mathbb{K}$. Die *Multiplikation von A mit λ*, geschrieben $\lambda \cdot A$ (bzw. λA), ist definiert als

$$\lambda \cdot A = \begin{pmatrix} \lambda\, a_{1\,1} & \lambda\, a_{1\,2} & \cdots & \lambda\, a_{1\,n} \\ \lambda\, a_{2\,1} & \lambda\, a_{2\,2} & \cdots & \lambda\, a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda\, a_{m\,1} & \lambda\, a_{m\,2} & \cdots & \lambda\, a_{m\,n} \end{pmatrix}.$$

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

Satz 4.19. *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) $1\, A = A$,

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 4.20 (Transponierte einer Matrix)**.** Sei $A \in \mathbb{K}^{m\times n}$. Die *transponierte Matrix* von A, bezeichnet mit $A^\intercal$ (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.

$$A^\intercal = (a^\intercal_{i\,j})^{n,m}_{i,j=1} \text{ mit } a^\intercal_{i\,j} = a_{j\,i}, \quad 1 \le i \le n,\ 1 \le j \le m.$$

Bemerkung 4.21. *Seien $A, B \in \mathbb{K}^{m\times n}$ und $\lambda \in \mathbb{K}$. Dann gilt*

i) $(A + B)^\intercal = A^\intercal + B^\intercal$,

ii) $(\lambda A)^\intercal = \lambda A^\intercal$,

iii) $(A^\intercal)^\intercal = A$.

Definition 4.22 (Quadratische Matrizen)**.**

i) Eine (n,n)-Matrix $A = (a_{i\,j})$ heißt *quadratisch.* Die Elemente $a_{i\,i}$, $1 \leq i \leq n$, 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 $A = (a_{i\,j})$, bei der alle Elemente „unterhalb (oberhalb) der Diagonalen“ Null sind, d.h. für die

$$a_{i\,j} = 0 \text{ für } 1 \leq j < i \leq n \quad \Big(a_{i\,j} = 0 \text{ für } \quad 1 \leq i < j \leq n\Big)$$

gilt, heißt obere (untere) *Dreiecksmatrix.*

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

Bemerkung 4.23. *Ist $A \in \mathbb{K}^{n\times n}$ symmetrisch, dann gilt also $A^\intercal = A$.*

Definition 4.24 (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 $C = (c_{i\,j})$, deren Koeffizienten definiert sind als

$$c_{\mathbf{i}\,\mathbf{j}} = \sum_{k=1}^{n} a_{\mathbf{i}\,k}\, b_{k\,\mathbf{j}} = a_{\mathbf{i}\,1}\, b_{1\,\mathbf{j}} + a_{\mathbf{i}\,2}\, b_{2\,\mathbf{j}} + \cdots + a_{\mathbf{i}\,n}\, b_{n\,\mathbf{j}},$$

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 4.25.

i) $A(B\,C) = (A\,B)C$, *für* $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times r}$, $C \in \mathbb{K}^{r\times q}$ (Assoziativität),

ii) $(A+B)\,C = A\,C + B\,C$ *für* $A, B \in \mathbb{K}^{m\times n}$, $C \in \mathbb{K}^{n\times r}$, *bzw.*
$A(B+C) = A\,B + A\,C$ *für* $A \in \mathbb{K}^{m\times n}$, $B, C \in \mathbb{K}^{n\times r}$ *(Distributivität)*

iii) $(\lambda A)B = \lambda(A\,B) = A(\lambda B)$, *für* $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times r}$, $\lambda \in \mathbb{K}$,

iv) $(A\,B)^\intercal = B^\intercal A^\intercal$ *für* $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times r}$.

Bemerkung 4.26.

i) *Im Allgemeinen ist die Matrixmultiplikation* nicht *kommutativ:*

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

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

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

Zudem setzt man $A^0 = I_n$.

iv) Sei $A \in \mathbb{K}^{m\times n}$, $A = (a_{i\,j})$, 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

$$A\,x = \begin{pmatrix} \sum_{j=1}^n a_{1\,j}\,x_j \\ \sum_{j=1}^n a_{2\,j}\,x_j \\ \vdots \\ \sum_{j=1}^n a_{m\,j}\,x_j \end{pmatrix} = \begin{pmatrix} a_{1\,1}\,x_1 + a_{1\,2}\,x_2 + \dots + a_{1\,n}\,x_n \\ a_{2\,1}\,x_1 + a_{2\,2}\,x_2 + \dots + a_{2\,n}\,x_n \\ \vdots \\ a_{m\,1}\,x_1 + a_{m\,2}\,x_2 + \dots + a_{m\,n}\,x_n \end{pmatrix}$$

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 4.27 ((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 $A\,x = b$ *inhomogenes Gleichungssystem* und $A\,x = 0$ heißt (das zugehörige) *homogene Gleichungssystem.*

Satz 4.28. *Sei $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^{m\times 1}$.*

i) *Sind v, w Lösungen von $A\,x = b$, dann ist $v - w$ Lösung des zugehörigen homogenen Gleichungssystems $A\,x = \mathbf{0}$.*

ii) *Ist v Lösung von $A\,x = b$ und z Lösung von $A\,x = \mathbf{0}$, dann ist $v + z$ Lösung von $A\,x = b$.*

iii) *$A\,x = b$ ist eindeutig lösbar, genau dann wenn das zugehörige homogene Gleichungssystem $A\,x = 0$ nur die „triviale“ Lösung, d.h. $x_1 = x_2 = \dots = x_n = 0$, hat.*

Definition 4.29 (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 4.30. *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 4.31.

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

$$A\,B = \begin{pmatrix} 2\,b_{1\,1} + 4\,b_{2\,1} & 2\,b_{1\,2} + 4\,b_{2\,2} \\ -b_{1\,1} + 3\,b_{2,1} & -b_{1\,2} + 3\,b_{2\,2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

bzw.

$$\begin{aligned} 2\,b_{1\,1} + 4\,b_{2\,1} &= 1, \qquad & 2\,b_{1\,2} + 4\,b_{2\,2} &= 0,\\ -b_{1\,1} + 3\,b_{2,1} &= 0, \qquad & -b_{1\,2} + 3\,b_{2\,2} &= 1. \end{aligned}$$

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

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

Satz 4.32. *Sei $A \in \mathbb{K}^{n\times n}$ und seien $e_i \in \mathbb{K}^{n\times 1}$ mit*

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

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 System

$$A\,x = e_j.$$

Satz 4.33. *Sei $A \in \mathbb{K}^{n\times n}$ regulär. Das Gleichungssystem $A\,x = b$ hat die eindeutige Lösung $A^{-1}\,b$.*

Satz 4.34. *Seien $A, B \in \mathbb{K}^{n\times n}$ regulär.*

i) *Dann ist $A\,B$ regulär und es gilt $(A\,B)^{-1} = B^{-1}\,A^{-1}$,*

ii) $(A^{-1})^\intercal = (A^\intercal)^{-1}$,

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

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

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 4.19).

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.

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

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.

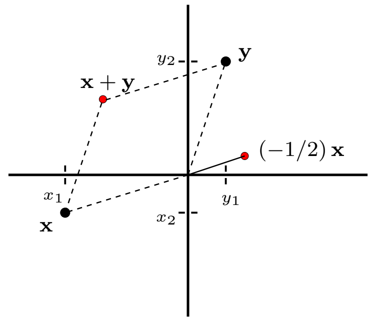


Abbildung 6: Vektoraddition und Skalarmultiplikation im $\mathbb{R}^2$

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

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

vi) Eine Funktion $f : \mathbb{K} \to \mathbb{K}$ der Form $f(x) = a_n\, x^n + a_{n-1}\, x^{n-1} + \cdots + a_1\, x + a_0$, $a_i \in \mathbb{K}$, 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 $\mathbf{a} = (a_1, \ldots, a_n)^\intercal \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\alpha \in \mathbb{R}$, heißt $H(\mathbf{a}, \alpha) = \{(x_1, \ldots, x_n)^\intercal \in \mathbb{R}^n : a_1\, x_1 + a_2\, x_2 + \cdots + a_n\, x_n = \alpha\}$* 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 (Linearkombination)**.** Sei V ein $\mathbb{K}$-Vektorraum, und seien $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$. Ein Ausdruck der Form

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

mit $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ heißt *Linearkombination* von $\mathbf{v}_1, \dots, \mathbf{v}_m$.

Definition 5.11 (Lineare Hülle)**.** Sei V ein $\mathbb{K}$-Vektorraum, und sei $M \subseteq V$ eine nicht-leere Teilmenge von V. Dann heißt

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

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

Lemma 5.12. *Sei V ein $\mathbb{K}$-Vektorraum, und sei $M \subseteq V$ eine nicht-leere Teilmenge von V. Dann ist* $\operatorname{lin} M$ *Teilraum von V.*

Bemerkung 5.13. *Sei V ein $\mathbb{K}$-Vektorraum, und seien $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$. Dann ist*

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

Beispiel 5.14.

i) $\operatorname{lin} \{ \binom{1}{0}, \binom{0}{1} \} = \mathbb{R}^2$.

ii) $\operatorname{lin} \{ \binom{1}{-1}, \binom{2}{1}, \binom{2}{-3} \} = \mathbb{R}^2$.

iii) $\operatorname{lin} \{ \binom{1}{-1}, \binom{-3}{3} \} = \{ (x_1, x_2)^\intercal \in \mathbb{R}^2 : x_1 = -x_2 \}$.

Definition 5.15 (Erzeugendensystem)**.** Sei V ein $\mathbb{K}$-Vektorraum und $M \subseteq V$, $M \neq \emptyset$. Ist $\operatorname{lin} M = V$, dann heißt M ein *Erzeugendensystem* von V.

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

Beispiel 5.16.

i) $\binom{1}{0}, \binom{0}{1}$ ist ein Erzeugendensystem von $\mathbb{R}^2$.

ii) $\binom{1}{-1}, \binom{2}{1}, \binom{2}{-3}$ ist ein Erzeugendensystem von $\mathbb{R}^2$.

iii) $\binom{1}{-1}, \binom{-3}{3}$ ist kein Erzeugendensystem von $\mathbb{R}^2$.

Definition 5.17 (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

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

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

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

Bemerkung 5.18.

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) $\mathbf{v}_1, \ldots, \mathbf{v}_m$ *sind also genau dann linear abhängig, wenn es Skalare* $\lambda_1, \ldots, \lambda_m$ *gibt, die* nicht alle 0 *sind, so dass* (5.1) *erfüllt ist.*

Beispiel 5.19.

i) $\binom{1}{0}, \binom{0}{1} \in \mathbb{R}^2$ sind linear unabhängig.

ii) $\binom{1}{-1}, \binom{2}{1}, \binom{2}{-3}$ sind linear abhängig.

iii) $\binom{1}{-1}, \binom{-3}{3}$ sind linear abhängig.

iv) $\mathbf{0}$ ist linear abhängig.

v) $1, \mathrm{i} \in \mathbb{C}$ sind linear unabhängig falls $\mathbb{C}$ als $\mathbb{R}$-Vektorraum betrachtet wird, aber linear abhängig falls $\mathbb{C}$ als $\mathbb{C}$-Vektorraum betrachtet wird.

Definition 5.20. Die Vektoren $\mathbf{e}_i \in \mathbb{K}^n$ mit

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

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

Bemerkung 5.21. *Die Einheitsvektoren* $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{K}^n$ *sind linear unabhängig und bilden ein Erzeugendensystem von* $\mathbb{K}^n$*, d.h.* $\operatorname{lin}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} = \mathbb{K}^n$.

Definition 5.22 (Basis)**.** Ein Erzeugendensystem M, welches aus linear unabhängigen Vektoren eines $\mathbb{K}$-Vektorraums besteht, heißt *Basis*.

Definition 5.23 (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 5.24.

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

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

Satz 5.26. *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. Jede Menge von mehr als n Vektoren ist linear abhängig.*

Desweiteren kann jede Menge von linear unabhängigen Vektoren zu einer Basis ergänzt werden.

Bemerkung 5.27. *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 5.28.

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 5.29. *Für* $\mathbf{a} = (a_1, \ldots, a_n)^\intercal \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ *ist die Dimension der Hyperebne* $H(\mathbf{a}, 0)$ *gleich* $n-1$, *also* $\dim_{\mathbb{R}}(H(\mathbf{a}, 0)) = n - 1$ *(siehe Bemerkung 5.9).*

6 Lineare Abbildungen und Matrizen

Definition 6.1 (Lineare Abbildung)**.** Seien V, W zwei $\mathbb{K}$-Vektorräume. Eine Abbildung $f : V \to W$ 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) $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(\mathbf{x}) = -\mathbf{x}$ ist eine lineare Abbildung (Spiegelung am Ursprung) .

ii) $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f((x_1, x_2)^\intercal) = (x_2, x_1)^\intercal$ ist eine lineare Abbildung (Spiegelung an der 45 Grad-Achse).

iii) $f : \mathbb{R}^3 \to \mathbb{R}^2$ mit $f((x_1, x_2, x_3)^\intercal) = (x_1, x_2)^\intercal$ ist eine lineare Abbildung (Orthogonale Projektion auf die (x_1, x_2)-Ebene) .

iv) $f : \mathbb{R}^2 \to \mathbb{R}^3$ mit $f((x_1, x_2)^\intercal) = (x_1, x_2, 0)^\intercal$ 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 $T_\mathbf{t} : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $f : V \to W$ eine lineare Abbildung. Seien $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ und $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$, $m \in \mathbb{N}$. 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 $f : \mathbb{K}^n \to \mathbb{K}^m$ ist genau dann linear, wenn es eine Matrix $A \in \mathbb{K}^{m\times n}$ gibt mit*

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

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

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

Beispiel 6.6.

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

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

ii) $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f((x_1, x_2)^\intercal) = (x_2, x_1)^\intercal$:

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

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

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

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

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

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

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

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

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

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 $f : \mathbb{K}^n \to \mathbb{K}^m$ 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 θ versteht man die lineare Abbildung $\mathrm{rot}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ mit

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

Definition 6.10 (Translationen)**.** Sei $\mathbf{t} \in \mathbb{R}^n$. Die Abbildung $T_\mathbf{t} : \mathbb{R}^n \to \mathbb{R}^n$ 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 θ ist gegeben durch die Abbildung $f = T_\mathbf{v} \circ \mathrm{rot}_\theta \circ T_{-\mathbf{v}}$, 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 $f : \mathbb{K}^n \to \mathbb{K}^m$ und $g : \mathbb{K}^m \to \mathbb{K}^r$ 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 $g \circ f : \mathbb{K}^n \to \mathbb{K}^r$ wieder eine lineare Abbildung mit*

$$(g \circ f)(\mathbf{x}) = (B\,A)\,\mathbf{x}.$$

Satz 6.13. *Eine lineare Abbildung $f : \mathbb{K}^n \to \mathbb{K}^n$ mit $f(\mathbf{x}) = A\,\mathbf{x}$ ist bijektiv genau dann, wenn A regulär ist. Die zugehörige Umkehrabbildung $f^{-1} : \mathbb{K}^n \to \mathbb{K}^n$ 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 $\mathrm{lin}\,\{\mathbf{e}_3\}$, $\mathrm{lin}\,\{\mathbf{e}_2\}$ oder $\mathrm{lin}\,\{\mathbf{e}_1\}$ mit Winkel θ versteht man die Abbildung(en)

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

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

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

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

iii) x_1-Achse; $\mathrm{rot}_{\mathbf{e}_1,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$\mathrm{rot}_{\mathbf{e}_1,\theta}(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

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

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

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

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

heißt *Kern* der Abbildung f, und die Menge

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

heißt *Bild* der Abbildung f.

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

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

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

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

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 $\mathrm{rg}(A)$ bezeichnet.

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

7 Normierte Vektorräume

Normierte Vektorräume

Definition 7.1 (Skalarprodukt)**.** Sei V ein $\mathbb{K}$-Vektorraum, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ 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

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

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

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

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

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

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 $\mathbb{C}^n$

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

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[19]-Schwarz[20]-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 $\mathbf{y}$ liner abhängig sind.

[19] Augustin Louis Cauchy, 1789–1857

[20] Hermann Amandus Schwarz, 1843–1921

Bemerkung 7.7. *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 7.8 (Polarkoordinaten)**.**

i) Jedes $\mathbf{x} = (x_1, x_2)^\intercal \in \mathbb{R}^2 \setminus \{0\}$ lässt sich eindeutig darstellen als (s. Definition 3.52)

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

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

ii) Jedes $\mathbf{x} = (x_1, x_2, x_3)^\intercal \in \mathbb{R}^3 \setminus \{0\}$ 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 7.9 (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 θ 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 θ um die x_3-Achse an,

iii) Man drehe die Rotationsachse wieder zurück in die Ausgangsposition.

- Man erhält so die Abbildung:

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

Definition 7.10 (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 7.11 (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 $f : V \to U$ mit

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

Ihre Matrixdarstellung ist gegeben durch

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

Definition 7.12 (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 $1 \le i, j \le n$

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

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

Bemerkung 7.13.

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*

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

Satz 7.14 (Gram[21]-Schmidt[22]-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* $1 \le k \le m$

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

Definition 7.15 (Orthogonale Matrix/Transformation)**.** Eine Matrix $U \in \mathbb{R}^{n \times n}$ heißt *orthogonale Matrix*, falls $U^\intercal = U^{-1}$, d.h

$$U^\intercal U = U \, U^\intercal = I_n.$$

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

Satz 7.16. *Sei* $U \in \mathbb{R}^{n \times n}$. *Dann sind die folgenden Aussagen äquivalent:*

i) U *orthogonal,*

ii) $U^\intercal$ *ist orthogonal,*

iii) *Die Spalten von* U *(und somit auch die Zeilen) bilden eine Orthonormalbasis des* $\mathbb{R}^n$.

[21] Jorgen Pedersen Gram, 1850–1916
[22] Ehrhart Schmidt, 1876–1959

8 Homogene Koordinaten, Quaternionen und Projektionen

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

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

Bemerkung 8.3.

i) *Sei $\mathbf{t} = (t_1, t_2)^\intercal \in \mathbb{R}^2$. Die Translation $T_\mathbf{t} : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $\mathrm{rot}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ um den Ursprung mit Winkel θ 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 $\mathbf{v} = (v_1, v_2)^\intercal$ mit Winkel θ 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 $\mathbf{t} = (t_1, t_2, t_3)^\intercal \in \mathbb{R}^3$. Die Translation $T_\mathbf{t} : \mathbb{R}^3 \to \mathbb{R}^3$ 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 θ 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 θ kann auch als Multiplikation mit der komplexen Zahl $\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$ aufgefasst werden.*

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

$$\mathbb{H} = \{a + b\,\mathrm{i} + c\,\mathrm{j} + d\,\mathrm{k} : a, b, c, d \in \mathbb{R}\},$$

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

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{i}\,\mathrm{j}\,\mathrm{k} = -1$$

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*

$$\mathrm{i}\,\mathrm{j} = \mathrm{k} = -\mathrm{j}\,\mathrm{i}, \qquad \mathrm{j}\,\mathrm{k} = \mathrm{i} = -\mathrm{k}\,\mathrm{j}, \qquad \textit{und} \qquad \mathrm{k}\,\mathrm{i} = \mathrm{j} = -\mathrm{i}\,\mathrm{k}.$$

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

Definition 8.7 (Konjugation und Norm in $\mathbb{H}$)**.** Sei $r = a + b\,\mathrm{i} + c\,\mathrm{j} + d\,\mathrm{k} \in \mathbb{H}$. Dann ist

i) $\overline{r} = a - b\,\mathrm{i} - c\,\mathrm{j} - d\,\mathrm{k}$ *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) $\|r\| = \sqrt{r\overline{r}}$,

ii) $\overline{r\,s} = \overline{s}\,\overline{r}$.

iii) $\|rs\| = \|r\|\|s\|$,

iv) *Für $r \neq 0$ ist $r^{-1} = \frac{\overline{r}}{\|r\|^2}$ (multiplikatives) Inverses zu r.*

Satz 8.9. *Sei* $\mathrm{Im}(\mathbb{H}) = \{x\,\mathrm{i} + y\,\mathrm{j} + z\,\mathrm{k} : x, y, z \in \mathbb{R}\}$ *und sei* $r = a + b\,\mathrm{i} + c\,\mathrm{j} + d\,\mathrm{k} \in \mathbb{H} \setminus \{0\}$*. Die Abbidung* $\mathbb{R}_r : \mathrm{Im}(\mathbb{H}) \to \mathrm{Im}(\mathbb{H})$ *mit* $R_r(p) = r\,p\,r^{-1}$ *ist bijketiv 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) $(R_{\pm 1}) =$ *Identität,*

iv) $R_r \circ R_s = R_{r \cdot s}$.

Satz 8.10. *Sei* $\mathrm{Im}(\mathbb{H}) = \{x\,\mathrm{i} + y\,\mathrm{j} + z\,\mathrm{k} : x, y, z \in \mathbb{R}\}$ *und sei* $r = a + b\,\mathrm{i} + c\,\mathrm{j} + d\,\mathrm{k} \in \mathbb{H} \setminus \{0\}$.

i) *Identifiziert man $p = x\,\mathrm{i} + y\,\mathrm{j} + z\,\mathrm{k} \in \mathrm{Im}(\mathbb{H})$ mit dem Vektor $(x, y, z) \in \mathbb{R}^3$, dann beschreibt $R_r : \mathbb{R}^3 \to \mathbb{R}^3$ eine Rotation, und zwar für $\|r\| = 1$ ist R_r die Rotation um $\{\lambda\,\mathbf{v} : \lambda \in \mathbb{R}\}$ mit den Winkel θ, wobei*

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

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

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

iii) *Somit lässt sich die Hintereinanderausführung von "Drehungen" $(R_r \circ R_s)(p)$ durch Quaternionenmultiplikation $rsps^{-1}r^{-1}$ 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 $Z_\pi : \mathbb{R}^n \to A$ 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* $\overline{\mathbf{a}} = (\mathbf{a}, -b)^\intercal \in \mathbb{R}^{n+1}$ *und* $\overline{A} = \{x \in \mathbb{R}^{n+1} : \langle \overline{\mathbf{a}}, \mathbf{x} \rangle = 0\}$ *die Abbildungsebene in "homogenen" Koordinaten. Sei* $\overline{\mathbf{v}} = (\mathbf{v}, 1)^\intercal \in \mathbb{R}^{n+1}$ *der Betrachtungspunkt und* $\overline{\mathbf{p}} = (\mathbf{p}, 1)^\intercal \in \mathbb{R}^{n+1}$*. Dann ist*

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

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

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

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

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

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

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

$$P_\pi(\overline{\mathbf{p}}) = \begin{pmatrix} -(\overline{a}_2\overline{v}_2 + \overline{a}_3\overline{v}_3) & \overline{a}_2\overline{v}_1 & \overline{a}_3\overline{v}_1 & \overline{a}_4\overline{v}_1 \\ \overline{a}_1\overline{v}_2 & -(\overline{a}_1\overline{v}_1 + \overline{a}_3\overline{v}_3) & \overline{a}_3\overline{v}_2 & \overline{a}_4\overline{v}_2 \\ \overline{a}_1\overline{v}_3 & \overline{a}_2\overline{v}_3 & -(\overline{a}_1\overline{v}_1 + \overline{a}_2\overline{v}_2) & \overline{a}_4\overline{v}_3 \\ 0 & 0 & 0 & -(\overline{a}_1\overline{v}_1 + \overline{a}_2\overline{v}_2 + \overline{a}_3\overline{v}_3) \end{pmatrix} \overline{p}.$$

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* $\overline{\mathbf{v}}_4 = 0$ *(allgemein* $\overline{\mathbf{v}}_{n+1} = 0$*) 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[23])**.** Sei $A = (a_{i\,j}) \in \mathbb{K}^{n\times n}$. 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 $\det(A) = a_{1\,1}$.

ii) Für $n > 1$ und einen Zeilenindex $i \in \{1, \dots, n\}$ ist

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

(Entwicklung nach der i-ten Zeile)

Beispiel 9.3.

i) Für $A = \begin{pmatrix} a_{1\,1} & a_{1\,2} \\ a_{2\,1} & a_{2\,2} \end{pmatrix}$ ist

$$\det(A) = a_{1\,1}\, a_{2\,2} - a_{1\,2}\, a_{2\,1}.$$

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

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

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

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

[23] 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

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

Bemerkung 9.7. *Wegen* $\det(A) = \det(A^\mathsf{T})$ *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* λ*-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 $A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 0 & 5 \\ 7 & 6 & 8 \end{pmatrix}$.

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

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) $\mathrm{rg}(A) = n$.

iv) $\det(A) \neq 0$.

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

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 $|\det(U)| = 1$.*

Definition 9.14 (Eigenwert, Eigenvektor)**.** Sei $A \in \mathbb{K}^{n\times n}$. $\lambda \in \mathbb{K}$ heißt *Eigenwert* der Matrix A (bzw. der linearen Abbildung $A : \mathbb{K}^n \to \mathbb{K}^n$), 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 λ.

Beispiel 9.15. 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 $(\alpha, \alpha)^\intercal$, $\alpha \neq 0$. Die Eigenvektoren zum Eigenwert -1 sind alle Vektoren der Form $(-\alpha, \alpha)^\intercal$, $\alpha \neq 0$.

Bemerkung 9.16. *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 9.17 (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 λ. Es heißt das *charakteristische Polynom* von A. Die Nullstellen $\lambda \in \mathbb{K}$ von $\chi_A(\lambda)$ sind die Eigenwerte von A.

Beispiel 9.18.

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 9.19. *Sei $A = (a_{i\,j}) \in \mathbb{K}^{n \times n}$ 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 $a_{n-1} = (-1)^{n-1}(a_{1\,1} + a_{2\,2} + \cdots + a_{n\,n})$.

Definition 9.20 (Spur)**.** Sei $A = (a_{i\,j}) \in \mathbb{K}^{n \times n}$. Die Summe der Diagonalelemente von A heißt *Spur* der Matrix A und wird mit $\mathrm{tr}(A)$ bezeichnet, d.h.

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

Bemerkung 9.21. *A und $A^\intercal$ haben das gleiche charakteristische Polynom.*

Bemerkung 9.22. *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 λ genau dann, wenn*

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

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

Definition 9.23 (Eigenraum)**.** Sei $A \in \mathbb{K}^{n \times n}$, und sei $\lambda \in \mathbb{K}$ Eigenwert von A. $\mathrm{Kern}(A - \lambda\, I_n) = \{\mathbf{u} \in \mathbb{R}^n : (A - \lambda\, I_n)\, \mathbf{u} = \mathbf{0}\}$, d.h. die Menge aller Eigenvektoren zum Eigenwert λ (plus den Nullvektor), heißt *Eigenraum* zum Eigenwert λ.

Bemerkung 9.24 (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 $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ von $\chi_A(\lambda)$.*

iii) *Für $\lambda_1, \ldots, \lambda_k$ bestimme man die Eigenräume, d.h. $\mathrm{Kern}(A - \lambda_i\, I_n) = \{\mathbf{u} \in \mathbb{R}^n : (A - \lambda_i\, I_n)\, \mathbf{u} = \mathbf{0}\}$, $1 \leq i \leq k$.*

Beispiel 9.25. Sei $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. A hat die beiden Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$.

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

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

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

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

Bemerkung 9.26. *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 λ von U*

$$|\lambda| = 1.$$

Definition 9.27 (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 9.28. *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. Dannist für $m \in \mathbb{N}$*

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

Satz 9.29. *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

Definition 9.30 (Diagonalisierbarkeit)**.** Eine Matrix $A \in \mathbb{K}^{n\times n}$ (bzw. eine lineare Abbildung $A : \mathbb{K}^n \to \mathbb{K}^n$) 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 9.31. *Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.*

Satz 9.32. *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*

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

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

Satz 9.33. *Jede symmetrische reelle 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 $U^\intercal A\, U$ eine Diagonalmatrix ist.*

Definition 9.34 (Positiv (negativ) definit)**.** Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. A heißt *positiv definit*, falls

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

Ist hingegen

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

dann heißt A *negativ definit*.

Gilt statt $>$ nur ≥ 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*.

Bemerkung 9.35. *Eine Matrix A ist genau dann positiv (semi)definit, wenn $-A$ neagtiv (semi)definit ist.*

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 $A \geq 0$ (≤ 0) sind.*

Satz 9.37 (Hurwitz[24]-Kriterium)**.** *Eine symmetrische $n \times n$-Matrix $A = (a_{i\,j}) \in \mathbb{R}^{n\times n}$ ist genau dann*

[24] Adolf Hurwitz, 1859-1919

i) *positiv definit, wenn für* $1 \leq m \leq n$ *gilt:*

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

ii) *negativ definit, wenn für* $1 \leq m \leq n$ *gilt:*

$$\det\left((a_{i\,j})_{i,j=1}^m\right) < 0 \textit{ für } m \textit{ ungerade und} \det\left((a_{i\,j})_{i,j=1}^m\right) > 0 \textit{ für } m \textit{ gerade.}$$

Definition 9.38 (Stochastische Matrix, Markov[25]-Matrix)**.** Eine Matrix $A = (a_{i\,j}) \in \mathbb{R}^{n \times n}$ heißt *stochastische Matrix* oder *Markov-Matrix* falls

i) alle Einträge nichtnegativ sind, d.h. $a_{i\,j} \geq 0$, $1 \leq i, j \leq n$, und

ii) die Summer der Einträge in jeder Spalte gleich eins ist, d.h. $\sum_{i=1}^n a_{i\,j} = 1$ 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.*

[25] Andrej Andrejewitsch Markov, 1856–1922

Teil II

Mathematik II

10 Gruppen, Ringe, Körper

Definition 10.1 (Gruppe (vergl. Def. 3.41))**.** Sei G eine nichtleere Menge, und sei $\otimes : G \times G \to G$ 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*[26]*) 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 $\mathrm{id}_{\{1,\ldots,n\}}$, 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 $\mathbb{K}$.
- $\mathrm{GL}(n, \mathbb{R}) = \{A \in \mathbb{R}^{n\times n} : \det A \neq 0\}$ 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 $(\mathbb{R}^n, +)$.
- $\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathbb{R}^{n\times n} : \det A = 1\}$ ist eine Untergruppe von $\mathrm{GL}(n, \mathbb{R})$.

[26] 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

$$< x >= \{x^k : k \in \mathbb{Z}\}.$$

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

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

Beispiel 10.10.

- $(\mathbb{Z}, +) =< 1 >$, $(\mathbb{Z}_m, +) =< [1]_m >$.
- $(\mathbb{Z}_5 \setminus \{[0]_5\}, \odot) =< [2]_5 >$ (vergl. Satz 3.45)
- $< 2 >= \{2^k : k \in \mathbb{Z}\} = \{..., \frac{1}{4}, \frac{1}{2}, 1, 2, 4, ...\}$ 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 $| < x > |$ die *Ordnung von x*.

Beispiel 10.12.

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

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

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

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[27])**.** *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.

[27] Joseph-Louis Lagrange, 1736–1813

Satz 10.17 (Euler[28])**.** *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[29], siehe Satz 3.32)**.** *Sei $a \in \mathbb{Z}$, und sei p Primzahl mit $\mathrm{ggT}(p, a) = 1$. Dann ist*

$$a^{p-1} \equiv 1 \bmod p.$$

Satz 10.20 (Chinesischer Restsatz)**.** *Seien $m_1, \dots, m_n \in \mathbb{N}$ mit $\mathrm{ggT}(m_i, m_j) = 1$ 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*

$$x \equiv a_i \bmod m_i, \quad 1 \leq i \leq n.$$

x ist modulo $m = m_1 \cdot m_2 \cdot \ldots \cdot m_n$ eindeutig bestimmt.

Bemerkung 10.21 (Bestimmen einer Lösung)**.**

- *Für $k = 1, \dots, n$ sei $M_k = m/m_k$.*
- *Dann ist $\mathrm{ggT}(M_k, m_k) = 1$ und es gibt ein $N_k \in \mathbb{Z}$ mit $N_k \cdot M_k \equiv 1 \bmod m_k$* (siehe Lemma 3.24).
- *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 $x \equiv 2 \bmod 3$, $x \equiv 3 \bmod 5$ und $x \equiv 2 \bmod 7$.

- $m_1 = 3$, $a_1 = 2$, $m_2 = 5$, $a_2 = 3$, $m_3 = 7$, $a_3 = 2$, $m = 105$.
- $M_1 = 35$, $M_2 = 21$, $M_3 = 15$.
- $N_1 = 2$, $N_2 = 1$, $N_3 = 1$.
- $x = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 233$.
- 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 $+, \cdot : R \times R \to R$ 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.

[28] Leonhard Euler, 1707–1783
[29] Pierre de Fermat, 1601–1665

Beispiel 10.24.

- Jeder Körper (vergl. Def. 3.43) 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 3.45).
- $\mathbb{Z}[\mathrm{i}] = \{a + b\,\mathrm{i} : a, b \in \mathbb{Z}\}$ 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 $a_0, a_1, \ldots, a_n \in R$. Die Abbildung $f : R \to R$ mit $f(x) = a_n \cdot x^n + a_{n-1}\, x^{n-1} + \cdots + a_1 \cdot x + a_0$ 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 $\mathrm{grad}(f)$ bezeichnet. Ist $a_{\mathrm{grad}(f)} = 1$, 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

$$\begin{aligned}
(p + q)(x) &= (2x^2 - x + 3) + (x^5 + 3x^3 - 4x^2 + 6x + 2) \\
&= x^5 + 3x^3 - 2x^2 + 5x + 5, \\
(p \cdot q)(x) &= (2x^2 - x + 3) \cdot (x^5 + 3x^3 - 4x^2 + 6x + 2) \\
&= 2x^7 - x^6 + 9x^5 - 11x^4 + 25x^3 - 14x^2 + 16x + 6.
\end{aligned}$$

Bemerkung 10.27. *Sei $(R, +, \cdot)$ Ring. Ist $(R \backslash \{0\}, \cdot)$ eine kommutative Gruppe, dann heißt $(R, +, \cdot)$ Körper (siehe Def. 3.43). Jeder Körper ist also ein Ring, aber nicht umgekehrt.*

Satz 10.28 (Polynomdivision)**.** *Sei $\mathbb{K}$ Körper und $\mathbb{K}[x]$ der zugehörige Polynomring. Für $f, g \in \mathbb{K}[x]$, $\mathrm{grad}(g) \geq 1$, gibt es $q, r \in \mathbb{K}[x]$ mit*

$$f = q \cdot g + r \text{ und } \mathrm{grad}(r) < \mathrm{grad}(g).$$

Man schreibt auch $r = f \bmod g$ (vgl. Def.3.20)

Definition 10.29 (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 $\mathrm{ggT}(f_1, f_2)$ bezeichnet.

Bemerkung 10.30. *Analog zum Euklidischen Algorithmus für Zahlen (vgl. Satz 3.22) 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. 3.25).*

Satz 10.31. *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., $f(a) = 0$.*

11 Folgen und Reihen

Definition 11.1 (Folgen)**.** Eine (reelle) *Folge* ist eine Abbildung $a : \mathbb{N}_0 \to \mathbb{R}$, $n \mapsto a_n$, 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* $a : \mathbb{N} \to \mathbb{R}$ *oder* $a : \mathbb{Z}_{\geq k} \to \mathbb{R}$ *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 $a : \mathbb{N}_0 \to \mathbb{C}$.

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 $f_0 = 1, f_1 = 1$ und für $n \geq 2$ sei $f_n = f_{n-1} + f_{n-2}$. Dies ergibt die *Fibonacci*-Folge $1, 1, 2, 3, 5, 8, \dots$ (vergl. Beispiel 3.6).
   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 abwechslendes 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 *monton 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 rellee 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 Interval $(a - \epsilon, a + \epsilon)$ liegen, d.h.

$$|a_n - a| < \epsilon \text{ für alle } n \geq n_0.$$

In diesem Fall schreibt man

$$\lim_{n \to \infty} a_n = a \quad \text{oder} \quad a_n \to a \text{ für } n \to \infty.$$

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* ϵ *ab; je kleiner* ϵ *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 Folgenglieder 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 $1, 1, 1, 1, 1, 1, 1, 1, 1, \ldots$. Man kann sie auch beschreiben als $(a_{2\,n})_{n\in\mathbb{N}_0}$.

Bemerkung 11.11. *Ein Folge komplexer Zahlen* $(a_n)_{n\in\mathbb{N}_0}$ *mit* $a_n = x_n + \mathrm{i}\, y_n$ *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* $x_n \to x$ *und* $y_n \to y$ *für* $n \to \infty$. *Der Grenzwert der Folge* $(a_n)_{n\in\mathbb{N}_0}$ *ist dann* $x + \mathrm{i}\, y$.

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 ∞, falls $\lim_{n\to\infty} 1/a_n = 0$ und ab einem Index n_0 gilt $a_n > 0$ für $n \geq n_0$. In diesem Fall schreibt man $\lim_{n\to\infty} a_n = \infty$. Man sagt auch „$(a_n)_{n\in\mathbb{N}_0}$ konvergiert gegen ∞“.

iii) Analog heißt eine Folge $(a_n)_{n\in\mathbb{N}_0}$ heißt *bestimmt divergent* gegen $-\infty$, falls $\lim_{n\to\infty} 1/a_n = 0$ und ab einem Index n_0 gilt $a_n < 0$ für $n \geq n_0$. In diesem Fall schreibt man $\lim_{n\to\infty} a_n = -\infty$. 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 ∞.

Satz 11.15 (Rechenregeln für konvergente Folge)**.** *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:*

$$\lim_{n\to\infty} \alpha\, a_n = \alpha\, a, \quad \lim_{n\to\infty} (a_n \pm b_n) = a \pm b,$$
$$\lim_{n\to\infty} (a_n \cdot b_n) = a \cdot b, \quad \lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

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*

$$\lim_{n\to\infty} a_n = \begin{cases} 0, & \text{für } k < m, \\ a_k/b_m, & \text{für } k = m, \\ \infty, & \text{für } k > m \text{ und } a_k/b_m > 0, \\ -\infty, & \text{für } k > m \text{ und } a_k/b_m < 0. \end{cases}$$

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* $\lim_{n\to\infty} a_n = 0$ *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. $\lim_{n\to\infty} a_n = 0$. 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*
$$\sum_{k=0}^{\infty}(\alpha\, a_k) = \alpha \sum_{k=0}^{\infty} a_k,$$
$$\sum_{k=0}^{\infty}(a_k \pm b_k) = \sum_{k=0}^{\infty} a_k \pm \sum_{k=0}^{\infty} b_k.$$

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
$$\sum_{k=0}^{\infty} c_k \ \textit{mit}\ c_k = \sum_{i=0}^{k} a_i\, b_{k-i}$$
(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[30]). *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.

[30] Gottfried Wilhelm Freiherr von Leibniz, 1646-1716

12 Grenzwerte und Stetigkeit von Funktionen

Definition 12.1 ((Strenge) Montonie von Funktionen)**.** Sei $D \subseteq \mathbb{R}$ und sei $f : D \to \mathbb{R}$ eine Funktion.

i) f heißt *streng monoton wachsend*, falls $f(x) > f(y)$ für alle $x, y \in D$ mit $x > y$.

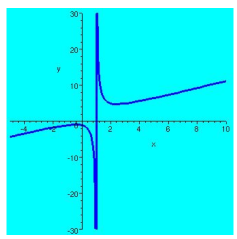
ii) f heißt *streng monoton fallend*, falls $f(x) < f(y)$ für alle $x, y \in D$ mit $x > y$.

iii) Gilt „lediglich“ $f(x) \geq f(y)$, bzw. $f(x) \leq f(y)$, dann heißt die Funktion „nur“ *monoton wachsend*, bzw. *monoton fallend*.

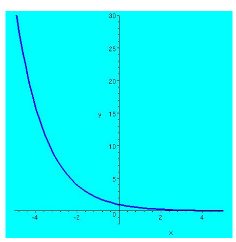
Definition 12.2 (Beschränktheit von Funktionen)**.** Eine Funktion $f : D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ heißt *beschränkt*, falls es ein $M > 0$ gibt mit $|f(x)| \leq M$ für alle $x \in D$.

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.

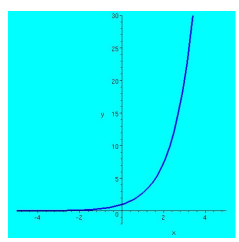
Beispiel 12.4. Sei
$$f(x) = \frac{x^3 + x^2 + x + 1}{x^2 - 1}.$$
Sie ist definiert auf der Menge $\mathbb{R} \setminus \{-1, 1\}$.



Definition 12.5 (Exponentialfunktion)**.** Sei $a > 0$. Eine Funktion der Form $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = a^x$ heißt *Exponentialfunktion zur Basis* a. Ist die Basis die *Eulersche Zahl* $\mathrm{e} = 2,7182818284...$, 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* $x \in \mathbb{R}$.

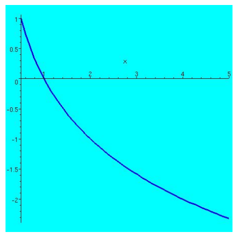
iii) *Ist $a < 1$, dann ist $f(x)$ eine streng monoton fallende Funktion. Ist $a > 1$, dann ist $f(x)$ eine streng monoton wachsende Funktion. Die Exponentialfunktion ist unbeschränkt.*

iv) $f(x + y) = f(x) \cdot f(y)$ *für alle* $x, y \in \mathbb{R}$.

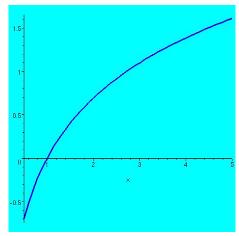
Definition 12.7 (Logarithmusfunktion)**.** Sei $a > 0$, $a \neq 1$. Die Umkehrfunktion der Exponentialfunktion $f : \mathbb{R} \to (0, \infty)$, $f(x) = a^x$, heißt *Logarithmusfunktion zur Basis* a, und wird mit

$$\log_a : (0, \infty) \to \mathbb{R}, \quad x \mapsto \log_a(x),$$

bezeichnet. Ist die Basis e, dann spricht man von dem *natürlichen Logarithmus*, der mit $\ln(x)$ bezeichnet wird.



Basis $a = 1/2$



Basis e

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, ∞)) versteht man alle Zahlen $\geq a$ (bzw. $> a$).

Definition 12.11 ((Un)gerade Funktionen)**.** Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion.

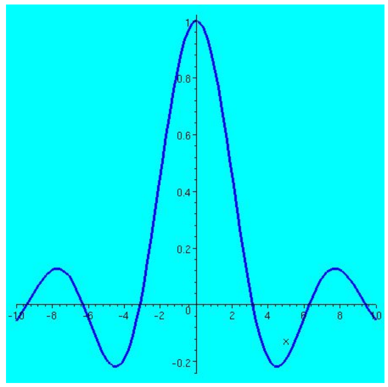
i) f heißt *gerade*, falls $f(-x) = f(x)$ für alle $x, -x \in D$.

ii) f heißt *ungerade*, falls $f(-x) = -f(x)$ für alle $x, -x \in D$.

Definition 12.12 (Periodische Funktionen)**.** Sei $f : \mathbb{R} \to \mathbb{R}$ 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. $\sin, \cos : \mathbb{R} \to \mathbb{R}$ sind periodische Funktionen mit Periode 2π. Desweiteren ist sin ungerade, und cos ist eine gerade Funktion.

Beispiel 12.14. Sei $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ mit $f(x) = \frac{\sin(x)}{x}$.



Definition 12.15 (Grenzwert einer Funktion)**.** Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion, und sei $x^\star \in \mathbb{R}$. Wenn es ein $y^\star \in \mathbb{R}$ gibt, so dass für jede Folge $(x_n)_{n\in\mathbb{N}_0}$ mit $x_n \in D \setminus \{x^\star\}$ und
$$\lim_{n\to\infty} x_n = x^\star$$
gilt, dass die Folge der Funktionswerte $(f(x_n))_{n\in\mathbb{N}_0}$ gegen $y^\star$ konvergiert, also
$$\lim_{n\to\infty} f(x_n) = y^\star,$$
dann heißt $y^\star$ *der Grenzwert von* f *für* x *gegen* $x^\star$. Man schreibt dafür
$$\lim_{x\to x^\star} f(x) = y^\star.$$
Hierbei ist auch $y^\star = \pm\infty$ erlaubt (siehe (11.13)) und man spricht dann von bestimmter Divergenz.

Beispiel 12.16.

i)
$$\lim_{x\to 0} \frac{1}{|x|} = \infty.$$

ii)
$$\lim_{x\to 5}\ 2\,x + 1 = 11.$$

iii) Sei $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ mit $f(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$. Dann existiert $\lim_{x\to 0} f(x)$ nicht, d.h. f hat keinen Grenzwert für x gegen 0.

Definition 12.17 (Rechts-Linksseitiger Grenzwert)**.** Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion, und sei $x^\star \in \mathbb{R}$. Wenn es ein $y^\star \in \mathbb{R}$ gibt, so dass für jede Folge $(x_n)_{n\in\mathbb{N}_0}$ mit $x_n \in D \setminus \{x^\star\}$, $x_n > x^\star$ (!), und
$$\lim_{n\to\infty} x_n = x^\star$$

gilt, dass
$$\lim_{n\to\infty} f(x_n) = y^\star,$$
dann heißt $y^\star$ *der rechtsseitige Grenzwert von f für x gegen $x^\star$*. Man schreibt dafür
$$\lim_{x\to x^\star +} f(x) = y^\star.$$
Analog definiert man den *linksssseitigen Grenzwert von f für x gegen $x^\star$*, indem man nur Folgen $(x_n)_{n\in\mathbb{N}_0}$ mit $x_n \in D\setminus\{x^\star\}$, $x_n < x^\star$ (!), und $\lim_{n\to\infty} x_n = x^\star$ betrachet. In diesem Falle schreibt man
$$\lim_{x\to x^\star -} f(x) = y^\star.$$
Beispiel 12.18. Sei $f : \mathbb{R}\setminus\{0\} \to \mathbb{R}$ mit $f(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$. Dann gilt
$$\lim_{x\to 0+} f(x) = 1 \text{ und } \lim_{x\to 0-} f(x) = -1.$$
Bemerkung 12.19. *Es gilt*
$$\lim_{x\to x^\star} f(x) = y^\star \Leftrightarrow \lim_{x\to x^\star +} f(x) = \lim_{x\to x^\star -} f(x) = y^\star.$$
Satz 12.20 (Rechenregeln für Grenzwerte)**.** *Seien f, g Funktionen mit $\lim\limits_{x\to x^\star} f(x) = \hat{y}$ und $\lim\limits_{x\to x^\star} g(x) = \tilde{y}$. Dann gilt*

i) $\lim\limits_{x\to x^\star} \alpha\, f(x) = \alpha\,\hat{y}$ *für* $\alpha \in \mathbb{R}$.

ii) $\lim\limits_{x\to x^\star} (f(x) \pm g(x)) = \hat{y} \pm \tilde{y}$.

iii) $\lim\limits_{x\to x^\star} (f(x)\cdot g(x)) = \hat{y}\cdot\tilde{y}$.

iv) $\lim\limits_{x\to x^\star} \frac{f(x)}{g(x)} = \frac{\hat{y}}{\tilde{y}}$ *falls* $\tilde{y} \neq 0$.

Definition 12.21 (Asymptotisches Verhalten)**.** Sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion. Unter dem *asymptotischen Verhalten von f für $x \to \infty$* versteht man den Grenzwert (falls existent und $\pm\infty$ sind zugelassen)
$$\lim_{x\to\infty} f(x).$$
Analog definiert man das *asymptotischen Verhalten von f für $x \to -\infty$* als
$$\lim_{x\to-\infty} f(x).$$
Beispiel 12.22.

i)
$$\lim_{x\to\infty} \frac{1}{x} = 0 \text{ und } \lim_{x\to-\infty} \frac{1}{x} = 0$$

ii)
$$\lim_{x\to\infty} x^3 = \infty \text{ und } \lim_{x\to-\infty} x^3 = -\infty.$$

Beispiel 12.23.

i) $\lim_{x\to\infty} x^a = \infty$ für $a > 0$.

ii) $\lim_{x\to\infty} x^a = 0$ für $a < 0$.

iii) $\lim_{x\to-\infty} x^n = (-1)^n\infty$ für $n \in \mathbb{N}$.

iv) Für $a > 0$ ist $\lim_{x\to\infty} \mathrm{e}^{a\,x} = \infty$ und $\lim_{x\to-\infty} \mathrm{e}^{a\,x} = 0$.

v) $\lim_{x\to\infty} \ln(x) = \infty$ und $\lim_{x\to 0} \ln(x) = -\infty$.

Definition 12.24 (Stetigkeit)**.** Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion, und sei $x^\star \in D$. f heißt *stetig im Punkt* x^*, falls

$$\lim_{x\to x^\star} f(x) = f(x^*),$$

d.h. der Grenzwert $\lim_{x\to x^\star} f(x)$ existiert und ist gleich $f(x^\star)$. Dies bedeutet, für jede Folge $(x_n)_{n\in\mathbb{N}_0}$, $x_n \in D \setminus \{x^\star\}$, mit $\lim_{n\to\infty} x_n = x^\star$ gilt

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x^\star).$$

Ist die Funktion f in jedem Punkt ihres Definitionsbereiches D stetig, dann heißt f *stetig.*

Bemerkung 12.25. *Stetigkeit von $f : D \to \mathbb{R}$ in $x^\star$ bedeutet also, dass für x „in der Nähe von“ $x^\star$ auch $f(x)$ „nahe“ an $f(x^\star)$ ist. Genauer: Zu jedem $\delta > 0$ existiert ein $\epsilon > 0$ mit*

$$|f(x) - f(x^\star)| < \delta \textit{ für alle } x \in D \textit{ mit } |x - x^\star| < \epsilon.$$

Beispiel 12.26.

i) Jede konstante Funktion ist stetig.

ii) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x$ ist stetig.

iii) $f : \mathbb{R} \to \mathbb{R}$ 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) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = |x|$ ist stetig.

v) $f : \mathbb{R} \to \mathbb{R}$ 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 $x^\star$ stetig sind. Dann sind auch die Funktionen*

$$\begin{aligned} &\alpha\, f(x) \textit{ mit } \alpha \in \mathbb{R}, \quad f(x) \pm g(x), \quad f(x)\cdot g(x), \\ &\frac{f(x)}{g(x)}, \textit{ falls } g(x^\star) \neq 0, \end{aligned}$$

stetig in $x^\star$. Ist die Verknüpfung $f(g(x))$ der beiden Funktionen definiert, dann ist auch $f(g(x))$ stetig in $x^\star$.

Beispiel 12.29.

i) Jedes Polynom ist stetig.

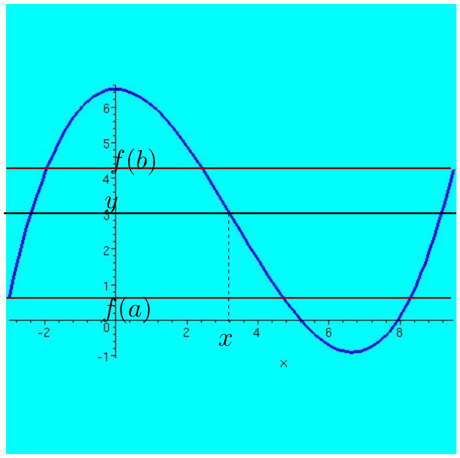
ii) Alle rationalen Funktionen sind in ihrem Definitionsbereich stetig.

iii) $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ mit $f(x) = 1/x$ ist stetig.

iv) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = |2\,x^2 - x|$ ist stetig.

v) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \mathrm{e}^{\sin(5\,\cos(x^2)+x^3)}$ ist stetig.

Satz 12.30 (Zwischenwertsatz)**.** *Sei $f : [a,b] \to \mathbb{R}$ 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[31])**.** *Jede auf einem abgeschlossenen Intervall $[a,b]$ stetige Funktion $f : [a,b] \to \mathbb{R}$ nimmt auf diesem Intervall ihr Maximum und Minimum an, d.h.*
es gibt $\tilde{x}, \hat{x} \in [a,b]$ mit

$$f(\tilde{x}) = \max\{f(x) : x \in [a,b]\}, \textit{und}$$
$$f(\hat{x}) = \min\{f(x) : x \in [a,b]\}.$$

[31] Karl Theodor Wilhelm Weierstraß 1815 - 1897

13 Differenzierbarkeit I

Definition 13.1 (Sekante)**.** Sei $D \subseteq \mathbb{R}$ und sei $f : D \to \mathbb{R}$ 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 $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion, und sei $x^\star \in D$. f heißt *differenzierbar* in $x^\star$, wenn der Grenzwert (der Sekantensteigungen)

$$\lim_{x \to x^\star} \frac{f(x) - f(x^\star)}{x - x^\star}$$

existiert. In diesem Fall wird der Grenzwert mit $f'(x^\star)$ oder auch $\frac{\mathrm{d}f}{\mathrm{d}x}(x^\star)$ bezeichnet, und heißt *Ableitung von f in $x^\star$*. Ist f in jedem Punkt $x^\star \in D$ differenzierbar, dann heißt f *differenzierbar*, und die Funktion

$$f : D \to \mathbb{R} \text{ mit } x \mapsto f'(x)$$

heißt *Ableitung von f*. Für $f'(x)$ schreibt man auch $\frac{\mathrm{d}f}{\mathrm{d}x}(x)$ bzw. einfach $\frac{\mathrm{d}f}{\mathrm{d}x}$.

Ist f' eine stetige Funktion, dann heißt f *stetig differenzierbar*.

Bemerkung 13.3. *Ist f in $x^\star$ differenzierbar, dann ist*

$$t(x) = f(x^\star) + f'(x^\star)\,(x - x^\star), \quad x \in \mathbb{R},$$

die Gleichung der Tangente durch $(x^\star, f(x^\star))$.

Beispiel 13.4.

i) $f(x) = x$ ist (stetig) differenzierbar mit $f'(x) = 1$.

ii) $f(x) = x^2$ ist (stetig) differenzierbar mit $f'(x) = 2\,x$.

iii) $f(x) = |x|$ ist nicht differenzierbar in $x^\star = 0$.

Satz 13.5. *Jede differenzierbare Funktion ist stetig (aber nicht umgekehrt).*

Satz 13.6. *Sei $D \subseteq \mathbb{R}$, und seien $f, g : D \to \mathbb{R}$ Funktionen, die in $x^\star \in D$ differenzierbar sind. Dann sind auch $\alpha\, f$, $f \pm g$, $f \cdot g$ in $x^\star$ differenzierbar, und es gilt:*

$$\begin{aligned} &\text{i)} \quad (\alpha\, f)'(x^\star) = \alpha\, f'(x^\star), \quad \text{ii)} \quad (f \pm g)'(x^\star) = f'(x^\star) \pm g'(x^\star), \\ &\text{iii)} \quad (f \cdot g)'(x^\star) = f'(x^\star) \cdot g(x^\star) + f(x^\star) \cdot g'(x^\star). \quad \text{(Produktregel)} \end{aligned}$$

Ist $g(x^\star) \neq 0$, dann ist auch $\frac{f}{g}$ in $x^\star$ differenzierbar mit

$$\text{iv)} \quad \left(\frac{f}{g}\right)'(x^\star) = \frac{f'(x^\star) \cdot g(x^\star) - f(x^\star) \cdot g'(x^\star)}{g(x^\star)^2}. \quad \text{(Quotientenregel)}$$

Ist die Verknüpfung $f \circ g$ definiert, und ist f in $g(x^\star)$ diffbar, dann ist $f \circ g$ in $x^\star$ differenzierbar mit

$$\text{v)} \quad (f \circ g)'(x^\star) = f'(g(x^\star)) \cdot g'(x^\star). \quad \text{(Kettenregel)}$$

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 $f : D \to \mathbb{R}$ 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 $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine stetige, streng monotone Funktion, und sei $\tilde{D} = f(D)$. Dann existiert die Umkehrfunktion $f^{-1} : \tilde{D} \to \mathbb{R}$. Ist f in $x^\star$ differenzierbar mit $f'(x^\star) \neq 0$, so ist f^{-1} in $y^\star = f(x^\star)$ differenzierbar, und es gilt*

$$\left(f^{-1}\right)'(y^\star) = \frac{1}{f'(x^\star)} = \frac{1}{f'(f^{-1}(y^\star))}$$

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. Aus der Ableitung von $\ln(x)$ folgt

$$1 = \ln{}'(1) = \lim_{x\to 1} \frac{\ln(x) - \ln(1)}{x - 1} = \lim_{x\to 1} \frac{\ln(x)}{x - 1}.$$

Speziell für die Folge $x_n = 1 + 1/n$ mit $\lim_{n\to\infty} x_n = 1$ folgt daraus

$$1 = \lim_{n\to\infty} n \ln\left(1 + \frac{1}{n}\right)$$

und so

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = \mathrm{e}.$$

Satz 13.12 (Regel von de l'Hospital[32])**.** *Sei $D \subseteq \mathbb{R}$, $f, g : D \to \mathbb{R}$ differenzierbar, und sei $x^\star \in D$ ($\pm\infty$ zugelassen). Gilt*

$$\lim_{x\to x^\star} f(x) = \lim_{x\to x^\star} g(x) = 0,$$

oder

$$\lim_{x\to x^\star} |f(x)| = \lim_{x\to x^\star} |g(x)| = \infty,$$

und existiert $\lim_{x\to x^\star} \frac{f'(x)}{g'(x)}$ ($\pm\infty$ zugelassen), dann gilt

$$\lim_{x\to x^\star} \frac{f(x)}{g(x)} = \lim_{x\to x^\star} \frac{f'(x)}{g'(x)}.$$

[32] Guillaume de l'Hospital, 1661-1704

Beispiel 13.13.

1. $\lim_{x\to 0}\frac{\sin(x)}{x} = \lim_{x\to 0}\frac{\cos(x)}{1} = 1.$
2. $\lim_{x\to\infty} x^2\,\mathrm{e}^{-x} = \lim_{x\to\infty}\frac{x^2}{\mathrm{e}^x} = \lim_{x\to\infty}\frac{2\,x}{\mathrm{e}^x} = \lim_{x\to\infty}\frac{2}{\mathrm{e}^x} = 0.$
3. $\lim_{x\to 0+} x\ln(x) = \lim_{x\to 0+}\frac{\ln(x)}{1/x} = \lim_{x\to 0+}\frac{1/x}{-1/x^2} = \lim_{x\to 0+}(-x) = 0.$

Definition 13.14 (Höhere Ableitungen)**.** Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine differenzierbare Funktion. Ist die Ableitung $f' : D \to \mathbb{R}$ in $x^\star \in D$ differenzierbar, dann heißt

$$(f')'(x^\star)$$

die *zweite* Ableitung von f in $x^\star$ und wird mit $f''(x^\star)$ bezeichnet.

Analog definiert man rekursiv die 3-te, 4-te usw. Ableitung in $x^\star$. Allgemein wird n-te Ableitung von f in $x^\star$ mit $f^{(n)}(x^\star)$ bezeichnet. Die Funktion selbst wird auch als 0-te Ableitung $f^{(0)}$ bezeichnet. Eine andere Bezeichung für die n-Ableitung ist $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$.

Definition 13.15 (Konvexe/konkave Funktionen)**.** Sei $D \subseteq \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ 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 $f : D \to \mathbb{R}$ 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 $f : D \to \mathbb{R}$ eine Funktion. $x^\star \in D$ heißt

i) *lokales (relatives) Maximum*, falls ein $\epsilon > 0$ existiert, so dass

$$f(x) \leq f(x^\star) \text{ für alle } x \in (x^\star - \epsilon, x^\star + \epsilon) \cap D.$$

ii) *lokales (relatives) Minimum*, falls ein $\epsilon > 0$ existiert, so dass

$$f(x) \geq f(x^\star) \text{ für alle } x \in (x^\star - \epsilon, x^\star + \epsilon) \cap D.$$

iii) *globales Maximum*, falls

$$f(x) \leq f(x^\star) \text{ für alle } x \in D.$$

iv) *globales Minimum*, falls

$$f(x) \geq f(x^\star) \text{ für alle } x \in D.$$

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* $f : [a, b] \to \mathbb{R}$ *differenzierbar, und sei* $x^\star \in (a, b)$*. Ist* $x^\star$ *ein lokales Minimum oder Maximum, dann gilt* $f'(x^\star) = 0$*.*

ii) *Sei* $f : [a, b] \to \mathbb{R}$ *zweimal differenzierbar, und sei* $x^\star \in (a, b)$*.*

   a) *Ist* $f'(x^\star) = 0$ *und* $f''(x^\star) < 0$*, dann ist* $x^\star$ *ein lokales Maximum.*

   b) *Ist* $f'(x^\star) = 0$ *und* $f''(x^\star) > 0$*, dann ist* $x^\star$ *ein lokales Minimum.*

Satz 13.19. *Sei* $f : [a, b] \to \mathbb{R}$ n*-mal differenzierbar, und sei* $x^\star \in (a, b)$*. Sei* $f'(x^\star) = 0$ *und* $f^{(n)}(x^\star)$ *die erste Ableitung, die an der Stelle* $x^\star$ *nicht gleich* 0 *ist, d.h.,* $f^{(i)}(x^\star) = 0$*,* $1 \le i \le n-1$ *und* $f^{(n)}(x^\star) \neq 0$*.*

Ist n *gerade und*

i) $f^{(n)}(x^\star) < 0$*, dann ist* $x^\star$ *ein lokales Maximum.*

ii) $f^{(n)}(x^\star) > 0$*, dann ist* $x^\star$ *ein lokales Minimum.*

Definition 13.20 (Wendepunkte, Sattelpunkte)**.** Sei $f : [a, b] \to \mathbb{R}$ eine Funktion. $x^\star \in (a, b)$ heißt *Wendepunkt*, wenn es ein $\epsilon > 0$ gibt, so dass f in $[x^\star - \epsilon, x^\star]$ konkav und in $[x^\star, x^\star + \epsilon]$ konvex ist (oder umgekehrt).

Ist f in $x^\star$ differenzierbar, und gilt zudem $f'(x^\star) = 0$, dann wird der Wendepunkt auch *Sattelpunkt* genannt.

Satz 13.21 (Kriterium für Wendepunkte)**.**

i) *Sei* $f : [a, b] \to \mathbb{R}$ *zweimal differenzierbar, und sei* $x^\star \in (a, b)$*. Ist* $x^\star$ *ein Wendepunkt, dann gilt* $f''(x^\star) = 0$*.*

ii) *Sei* $f : [a, b] \to \mathbb{R}$ *dreimal differenzierbar, und sei* $x^\star \in (a, b)$*. Ist* $f''(x^\star) = 0$ *und* $f'''(x^\star) \neq 0$*, dann ist* $x^\star$ *ein Wendepunkt.*

Satz 13.22. *Sei* $f : [a, b] \to \mathbb{R}$ n*-mal differenzierbar, und sei* $x^\star \in (a, b)$*. Sei* $f'(x^\star) = 0$ *und* $f^{(n)}(x^\star)$ *die erste Ableitung, die an der Stelle* $x^\star$ *nicht gleich* 0 *ist, d.h.,* $f^{(i)}(x^\star) = 0$*,* $1 \le i \le n-1$ *und* $f^{(n)}(x^\star) \neq 0$*.*

Ist n *ungerade, dann ist* $x^\star$ *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 $f : [0, 1] \to \mathbb{R}$ 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 $f : [-1, 1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1 + x, & -1 \le x \le 0, \\ 1 - x + x^2, & 0 \le x \le 1. \end{cases}$$

Satz 13.27 (Satz von Rolle[33])**.** *Sei* $f : [a, b] \to \mathbb{R}$ *eine differenzierbare Funktion, mit* $f(a) = f(b)$*. Dann gibt es ein* $x^\star \in (a, b)$ *mit* $f'(x^\star) = 0$*.*

Satz 13.28 (Mittelwertsatz der Differentialrechnung)**.** *Sei* $f : [a, b] \to \mathbb{R}$ *eine differenzierbare Funktion. Dann gibt es ein* $x^\star \in (a, b)$ *mit*

$$f'(x^\star) = \frac{f(b) - f(a)}{b - a}.$$

[33] Michel Rolle, 1652–1719

Satz 13.29. *Sei $f : [a, b] \to \mathbb{R}$ eine differenzierbare Funktion. mit $f'(x) = 0$ für alle $x \in (a, b)$. Dann ist f konstant.*

Satz 13.30. *Sei $f : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion, und für eine Konstante c gelte $f'(x) = c\,f(x)$ für alle $x \in \mathbb{R}$. Dann ist $f(x) = f(0)\mathrm{e}^{c\,x}$.*

14 Taylor- und Potenzreihen

Definition 14.1 (Taylorpolynom[34])**.** Sei $D \subseteq \mathbb{R}$, und sei $f : D \to \mathbb{R}$ eine Funktion, die an der Stelle $x^\star \in D$ mindestens n-mal differenzierbar ist. Dann heißt

$$\begin{aligned} T_n(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(x^\star)}{k!} (x - x^\star)^k \\ &= f(x^\star) + f'(x^\star)(x - x^\star) + \frac{f''(x^\star)}{2} (x - x^\star)^2 + \cdots + \frac{f^{(n)}(x^\star)}{n!} (x - x^\star)^n \end{aligned}$$

das zu f gehörige *Taylorpolynom vom Grad* n an der Stelle $x^\star$, bzw. mit *Entwicklungspunkt* $x^\star$. Man schreibt daher auch manchmal $T_n(x^\star; x)$.

Bemerkung 14.2. *$T_1(x)$ ist gerade die Tangente an der Stelle $x^\star$ (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 Entwicklungspunkt $x^\star = 0$ 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 $x^\star = 0$ 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 $f : D \to \mathbb{R}$ eine Funktion, die an der Stelle $x^\star \in D$ mindestens n-mal differenzierbar ist. Der „Fehler"

$$R_n(x) = f(x) - T_n(x^\star; x)$$

heißt *Restglied*.

Satz 14.5 (Taylor)**.** *Sei $D \subseteq \mathbb{R}$, und sei $f : D \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion. Dann gilt:*

$$|R_n(x)| = |f(x) - T_n(x^\star; x)| \le \frac{M}{(n+1)!} |x - x^\star|^{n+1},$$

wobei M so gewählt ist, dass $|f^{(n+1)}(x)| \le M$ für alle $x \in D$.

Beispiel 14.6. Sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \sin(x)$, $x^\star = 0$ und $n = 11$. Dann ist $f^{(12)}(x) = \sin(x)$, und somit gilt die Fehlerabschätzung

$$|R_{11}(1/2)| = |\sin(1/2) - T_{11}(0; 1/2)| \le \frac{1}{(12)!} |1/2|^{12} < 10^{-12}.$$

[34] Brook Taylor, 1685-1731

Definition 14.7 (Landausymbole[35])**.** Sei $D \subseteq \mathbb{R}$, seien $f, g : D \to \mathbb{R}$, und sei $x^\star \in D$.

i) f heißt *von Ordnung groß O von g für x gegen $x^\star$* falls es Konstanten $M > 0$, $\delta > 0$ gibt, so dass für $x \in (x^\star - \delta, x^\star + \delta)$ gilt
$$\left|\frac{f(x)}{g(x)}\right| \leq M.$$
Man schreibt dann $f = \mathcal{O}(g)$ für $x \to x^\star$. In diesem Fall wächst f höchstens so schnell wie g. Die Bedingung ist sicherlich erfüllt, falls $\lim_{x\to x^\star} |\frac{f(x)}{g(x)}| \leq M$ gilt. Ist $x^\star = \pm\infty$, dann schreibt man $f = \mathcal{O}(g)$ für $x \to x^\star$ falls
$$\lim_{x\to x^\star}\left|\frac{f(x)}{g(x)}\right| \leq M.$$

ii) f heißt *von Ordnung klein o von g für x gegen $x^\star$* falls
$$\lim_{x\to x^\star}\left|\frac{f(x)}{g(x)}\right| = 0.$$
Man schreibt dann $f = o(g)$ für $x \to x^\star$. In diesem Fall wächst f (echt) langsamer als g.

Bemerkung 14.8. *Mit dieser Notation ist also für x gegen $x^\star$*
$$f(x) = T_n(x^\star; x) + \mathcal{O}((x - x^\star)^{n+1}).$$

Definition 14.9 (Taylorreihe)**.** Sei $D \subseteq \mathbb{R}$, und sei $f : D \to \mathbb{R}$ eine beliebig oft differenzierbare Funktion. Sei $x^\star \in D$. Die Reihe
$$T_\infty(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x^\star)}{k!}\,(x - x^\star)^k$$
heißt *Taylorreihe* mit Entwicklungspunkt $x^\star$. Man schreibt auch manchmal $T_\infty(x^\star; x)$.

Beispiel 14.10. Sei $f(x) = \frac{1}{1-x}$ und $x^\star = 0$. 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 $x^\star \in \mathbb{R}$ (bzw. $\in \mathbb{C}$). Für $x \in \mathbb{R}$ (bzw. $\in \mathbb{C}$) heißt die Reihe
$$\sum_{k=0}^{\infty} a_k\,(x - x^\star)^k$$
Potenzreihe mit Entwicklungspunkt $x^\star$.

Bemerkung 14.12. *Jede Taylorreihe ist eine Potenzreihe.*

Lemma 14.13. *Konvergiert die Potenzreihe $\sum_{k=0}^{\infty} a_k\,(x - x^\star)^k$ für ein $x_1 \in \mathbb{R}$, dann konvergiert sie (absolut) für alle $x \in \mathbb{R}$ mit*
$$|x - x^\star| < |x_1 - x^\star|.$$

[35] Edmund Georg Hermann Landau, 1877-1938

Definition 14.14 (Supremum/Infimum)**.** Sei $A \subseteq \mathbb{R}$. Ein Element $u \in \mathbb{R}$ heißt *obere Schranke* von A, falls $a \leq u$ für alle $a \in A$. Ein Element $l \in \mathbb{R}$ heißt *untere Schranke* von A, falls $l \leq a$ 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, $x^\star \in \mathbb{R}$, und $\sum_{k=0}^{\infty} a_k \, (x - x^\star)^k$ die zugehörige Potenzreihe.

$$C = \{x \in \mathbb{R} : \sum_{k=0}^{\infty} a_k \, (x - x^\star)^k \text{ konvergiert } \}$$

heißt *Konvergenzbereich* der Potenzreihe, und

$$\rho = \sup\{|x - x^\star| : x \in C\}$$

heißt der *Konvergenzradius* der Potenzreihe.

Beispiel 14.17. Für die Potenzreihe $\sum_{k=0}^{\infty} x^k$ (Entwicklungspunkt $x^\star = 0$) ist

$$C = \{x \in \mathbb{R} : |x| < 1\} \text{ und } \rho = 1.$$

Satz 14.18. *Sei $\sum_{k=0}^{\infty} a_k \, (x - x^\star)^k$ eine Potenzreihe mit Konvergenzradius ρ. Dann ist die Reihe*

i) *absolut konvergent für alle $x \in \mathbb{R}$ mit $|x - x^\star| < \rho$, und*

ii) *divergent für alle $x \in \mathbb{R}$ mit $|x - x^\star| > \rho$.*

Satz 14.19. *Sei $\sum_{k=0}^{\infty} a_k \, (x - x^\star)^k$ eine Potenzreihe mit Konvergenzradius ρ, und es sei $a_k \neq 0$ für fast alle $k \in \mathbb{N}_0$. Existiert der Grenzwert $\lim_{k \to \infty} |a_k / a_{k+1}|$ (∞ zugelassen), dann ist*

$$\rho = \lim_{k \to \infty} \left| \frac{a_k}{a_{k+1}} \right|.$$

Beispiel 14.20. Für die Potenzreihe $\sum_{k=1}^{\infty} \frac{1}{k} \, x^k$, also $a_k = 1/k$, ist

$$\lim_{k \to \infty} \left| \frac{a_k}{a_{k+1}} \right| = \lim_{k \to \infty} \frac{k+1}{k} = 1.$$

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* $\sum_{k=0}^{\infty} a_k\,(x-x^\star)^k$ *eine Potenzreihe mit Konvergenzradius* ρ. *Existiert der Grenzwert* $\lim_{k\to\infty}\sqrt[k]{|a_k|}$ *(*∞ *zugelassen), dann gilt*

$$\rho = \frac{1}{\lim_{k\to\infty}\sqrt[k]{|a_k|}}.$$

Ist der Grenzwert 0*, dann ist* $\rho=\infty$*, und ist der Grenzwert* ∞ *(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

$$\lim_{k\to\infty}\sqrt[k]{|a_k|} = \lim_{k\to\infty} k = \infty.$$

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* $n\to\infty$*.*

Satz 14.24. *Einige Taylorreihen und ihr Konvergenzradien* ρ*:*

$$\text{i) } \frac{1}{1-x} = \sum_{k=0}^{\infty} x^k, \quad x^\star = 0, \rho = 1,$$

$$\text{ii) } \ln(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(x-1)^k, \quad x^\star = 1, \rho = 1,$$

$$\text{iii) } \mathrm{e}^x = \sum_{k=0}^{\infty} \frac{1}{k!}x^k, \quad x^\star = 0, \rho = \infty,$$

$$\text{iv) } \sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k+1)!}x^{2\,k+1}, \quad x^\star = 0, \rho = \infty,$$

$$\text{v) } \cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k)!}x^{2\,k}, \quad x^\star = 0, \rho = \infty,$$

Satz 14.25. *Eine Potenzreihe* $g(x) = \sum_{k=0}^{\infty} a_k\,(x-x^\star)^k$ *ist innerhalb ihres Konvergenzbereiches beliebig oft stetig differenzierbar. Die Ableitungen können durch* gliedweises Differenzieren *erhalten werden, d.h.*

$$g'(x) = \sum_{k=1}^{\infty} k\,a_k\,(x-x^\star)^{k-1} = \sum_{k=0}^{\infty}(k+1)\,a_{k+1}\,(x-x^\star)^k,$$

und der Konvergenzbereich von $g'(x)$ *ist gleich dem Konvergenzbereich von* $g(x)$*.*

Beispiel 14.26.

$$\cos(x) = (\sin(x))' = \left(\sum_{k=0}^{\infty}\frac{(-1)^k}{(2\,k+1)!}x^{2\,k+1}\right)' = \sum_{k=0}^{\infty}(2\,k+1)\,\frac{(-1)^k}{(2\,k+1)!}x^{2\,k} = \sum_{k=0}^{\infty}\frac{(-1)^k}{(2\,k)!}x^{2\,k}.$$

15 Integralrechnung I

Definition 15.1 (Zerlegung)**.** Sei $[a,b] \subset \mathbb{R}$ ein Intervall. Eine Menge $Z = \{x_0, x_1, \ldots, x_n\}$ mit $a = x_0 < x_1 < x_2 \cdots < x_n = b$ heißt *Zerlegung* von $[a,b]$ bzgl. den Stützstellen $x_0, x_1, \ldots, x_n$.

Definition 15.2 (Ober- Untersumme)**.** Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion, und sei $Z = \{x_0, \ldots, x_n\}$ eine Zerlegung von $[a,b]$. Dann heißt

1. $U_f(Z) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \cdot \inf\{f(x) : x \in [x_i, x_{i+1}]\}$ *Untersumme* 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 $f : [a,b] \to \mathbb{R}$ 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*
$$U_f(Z_2) \geq U_f(Z_1) \quad \textit{und} \quad O_f(Z_2) \leq O_f(Z_1).$$
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 Untersummen nach oben beschränkt, und Obersummen nach unten.*

Definition 15.5 ((Riemann[36])-Integral)**.** Sei $f : [a,b] \to \mathbb{R}$ 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
$$\int_a^b f(x)\,\mathrm{d}x$$
bezeichnet. Abkürzend schreibt man auch nur $\int_a^b f$.

Notation 15.6. Für ein Integral $\int_a^b f(x)\,\mathrm{d}x$ heißt die Funktion f der *Integrand*, x heißt die *Integrationsvariable* und a, b heißen die *Integrationsgrenzen.*

Satz 15.7. *Seien $f, g : [a,b] \to \mathbb{R}$ integrierbar über $[a,b]$, und sei $\alpha \in \mathbb{R}$. Dann gilt*

i) *αf und $f + g$ sind integrierbar, und es ist:*
$$\int_a^b \alpha f = \alpha \int_a^b f \ \textit{und} \ \int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

ii) *$f \cdot g$ ist in integrierbar.*

[36] Georg Friedrich Bernhard Riemann, 1826-1866

Bemerkung 15.8. *Seien $f, g : [a, b] \to \mathbb{R}$ 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 $f : [a, b] \to \mathbb{R}$ 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 $f : [a, b] \to \mathbb{R}$ eine Funktion.*

i) *Ist f monoton, dann ist f integrierbar.*

ii) *Ist f stetig, dann ist f integrierbar.*

Definition 15.11. Sei $f : [a, b] \to \mathbb{R}$ integrierbar. Dann setzen wir

$$\int_c^c f(x)\,\mathrm{d}x = 0 \text{ und } \int_d^c f(x)\,\mathrm{d}x = -\int_c^d f(x)\,\mathrm{d}x$$

für für $a \leq c \leq d \leq b$.

Definition 15.12 (Stammfunktion)**.** Seien $f, F : [a, b] \to \mathbb{R}$ zwei Funktion. Ist F differenzierbar mit $F' = f$, dann heißt F Stammfunktion von f.

Satz 15.13 (Hauptsatz der Differential- und Integralrechnung)**.** *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Die Funktion $F : [a, b] \to \mathbb{R}$ mit*

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

ist Stammfunktion von f

Lemma 15.14. *Sei $F : [a, b] \to \mathbb{R}$ eine Stammfunktion der Funktion $f : [a, b] \to \mathbb{R}$. Dann ist $G : [a, b] \to \mathbb{R}$ 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 $f : [a, b] \to \mathbb{R}$ stetig, und sei F eine Stammfunktion von f. Dann gilt:*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

Notation 15.16. Sei F Stammfunktion von $f : [a, b] \to \mathbb{R}$. Dann setzt man

$$F(x)\Big|_a^b = F(b) - F(a).$$

Definition 15.17 (Bestimmtes und Unbestimmtes Integral)**.** Sei $f : [a, b] \to \mathbb{R}$ integrierbar. Eine Stammfunktion F von f heißt auch *unbestimmtes Integral* von f und wird mit

$$\int f(x)\,\mathrm{d}x$$

(bzw. $\int f$) bezeichnet. Manchmal versteht man unter dem unbestimmten Integral auch die Menge aller Stammfunktionen, und schreibt (relativ unpräzise)

$$\int f(x)\,\mathrm{d}x = F(x) + C,$$

wobei F eine Stammfunktion ist, und C deutet die Menge aller Konstanten an.

Hingegen nennt man das Integral $\int_a^b f(x)\,\mathrm{d}x$ mit gegebenen Integrationsgrenzen a, b auch *bestimmtes Integral.*

Beispiel 15.18. Für $n \in \mathbb{N}_0$ ist

$$\int x^n\,\mathrm{d}x = \frac{x^{n+1}}{n+1} \quad (+C).$$

Satz 15.19. *Sei $f(x) = \sum_{k=0}^{\infty} a_k\,(x - x^\star)^k$ eine Potenzreihe mit Konvergenzradius ρ. Dann ist auch die Funktion $F(x)$ die durch gliedweises Integrieren der Potenzreihe entsteht, d.h.*

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1}\,(x - x^\star)^{k+1},$$

konvergent in $\{x \in \mathbb{R} : |x - x^\star| < \rho\}$, 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}, $x \neq 0$	$\ln(\lvert x\rvert)$
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 $g : [a, b] \to \mathbb{R}$ stetig differenzierbar, und $f : g([a, b]) \to \mathbb{R}$ eine stetige Funktion mit Stammfunktion F. Dann ist $F \circ g$ Stammfunktion von $(f \circ g) \cdot g'(x)$ und es gilt*

$$\int_a^b f(g(x)) \cdot g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(y)\,\mathrm{d}y = F(g(b)) - F(g(a)).$$

Hier wurde $y = g(x)$ substituiert, und mit der „informellen“ Schreibweise

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}g}{\mathrm{d}x} = g'(x) \ \textit{bzw.}\ \mathrm{d}y = g'(x)\mathrm{d}x$$

kann man sich auch (für das unbestimmte Integral) merken

$$\int f(g(x)) \cdot g'(x)\,\mathrm{d}x = \int f(y)\,\mathrm{d}y.$$

Beispiele 15.22. $\int_{1/\pi}^{2/\pi} -\frac{\sin(1/x)}{x^2}\,\mathrm{d}x$. Sei $g : [1/\pi, 2/\pi] \to \mathbb{R}$ 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

$$\begin{aligned}\int_{1/\pi}^{2/\pi} -\frac{\sin(1/x)}{x^2}\,\mathrm{d}x &= \int_{1/\pi}^{2/\pi} f(g(x))\cdot g'(x)\,\mathrm{d}x\\ &= \int_{g(1/\pi)}^{g(2/\pi)} f(y)\,\mathrm{d}y = \int_{\pi}^{\pi/2} \sin y\,\mathrm{d}y\\ &= -\int_{\pi/2}^{\pi} \sin y\,\mathrm{d}y = \cos(y)\Big|_{\pi/2}^{\pi}\\ &= \cos(\pi) - \cos(\pi/2) = -1.\end{aligned}$$

Bemerkung 15.23.

$$\int \frac{f'(x)}{f(x)} = \ln(|f(x)|) + C.$$

Beispiel 15.24.

$$\int_a^b \tan x\,\mathrm{d}x = \int_a^b \frac{\sin x}{\cos x}\,\mathrm{d}x = -\ln(|\cos x|)\Big|_a^b.$$

Satz 15.25 (Partielle Integration)**.** *Es sei $g : [a,b] \to \mathbb{R}$ stetig differenzierbar und $f : [a,b] \to \mathbb{R}$ stetig mit Stammfunktion F. Dann ist $F\cdot g$ Stammfunktion von $f\cdot g + F\cdot g'$ und es gilt*

$$\int_a^b f(x)\cdot g(x)\,\mathrm{d}x = (F(x)\cdot g(x))\Big|_a^b - \int_a^b F(x)\cdot g'(x)\,\mathrm{d}x.$$

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 $\int_a^b \cos(x)\cdot x\,\mathrm{d}x$. Mit $g(x) = x$, $f(x) = \cos(x)$ und so $g'(x) = 1$, $F(x) = \sin(x)$ erhalten wir

$$\begin{aligned}\int_a^b \cos(x)\cdot x\,\mathrm{d}x &= \int_a^b f(x)\cdot g(x)\,\mathrm{d}x\\ &= (F(x)\cdot g(x))\Big|_a^b - \int_a^b F(x)\cdot g'(x)\,\mathrm{d}x\\ &= (\sin(x)\cdot x)\Big|_a^b - \int_a^b \sin(x)\,\mathrm{d}x\\ &= (\sin(b)\,b - \sin(a)\,a) - (-\cos(b) + \cos(a))\\ &= \sin(b)\,b - \sin(a)\,a + \cos(b) - \cos(a).\end{aligned}$$

   Insbesondere ist $\sin(x)\,x + \cos(x)$ (eine) Stammfunktion von $\cos(x)\,x$.

2. Wir betrachten $\int_a^b \ln(x)\,\mathrm{d}x$. Mit $g(x) = \ln(x)$, $f(x) = 1$ und so $g'(x) = 1/x$, $F(x) = x$ erhalten wir

$$\begin{aligned}\int_a^b \ln(x)\,\mathrm{d}x &= \int_a^b f(x)\cdot g(x)\,\mathrm{d}x\\ &= (F(x)\cdot g(x))\Big|_a^b - \int_a^b F(x)\cdot g'(x)\,\mathrm{d}x\\ &= (x\,\ln(x))\Big|_a^b - \int_a^b x\cdot\frac{1}{x}\,\mathrm{d}x\\ &= (b\,\ln(b) - a\,\ln(a)) - (b-a).\end{aligned}$$

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 $f:(a,b)\to\mathbb{R}$ eine Funktion.

1. Ist f auf jedem Teilintervall $[c,d]\subset[c,b)$, also $c \le d < b$, integrierbar und existiert der Grenzwert
$$\lim_{d\to b}\int_c^d f(x)\,\mathrm{d}x$$
so bezeichnet man ihn mit $\int_c^b f(x)\,\mathrm{d}x$.

2. Ist f auf jedem Teilintervall $[d,c]\subset(a,c]$, also $a < d \le cb$, integrierbar und existiert der Grenzwert
$$\lim_{d\to a}\int_d^c f(x)\,\mathrm{d}x$$
so bezeichnet man ihn mit $\int_a^c f(x)\,\mathrm{d}x$.

3. Existieren die Grenzwerte
$$\lim_{d\to b}\int_c^d f(x)\,\mathrm{d}x \text{ und } \lim_{d\to a}\int_d^c f(x)\,\mathrm{d}x$$
so setzt man
$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

Beispiel 15.28.

1. Für $s>1$ betrachten wir $\int_1^\infty \frac{1}{x^s}\,\mathrm{d}x$. Für $d\in[1,\infty)$ ist
$$\int_1^d \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{1-s}\frac{1}{x^{s-1}}\bigg|_1^d = \frac{1}{1-s}\left(\frac{1}{d^{s-1}}-1\right).$$
Wegen
$$\lim_{d\to\infty}\frac{1}{1-s}\left(\frac{1}{d^{s-1}}-1\right) = \frac{1}{1-s}\cdot(-1) = \frac{1}{s-1}$$
$$\int_1^\infty \frac{1}{x^s}\,\mathrm{d}x = \lim_{d\to\infty}\int_1^d \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{s-1}.$$
Für $0\le s\le 1$ existiert der obige Grenzwert nicht.

2. Für $0\le s\le 1$ betrachten wir $\int_0^1 \frac{1}{x^s}\,\mathrm{d}x$. Für $d\in(0,1]$ ist
$$\int_d^1 \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{1-s}\frac{1}{x^{s-1}}\bigg|_d^1 = \frac{1}{1-s}\left(1-\frac{1}{d^{s-1}}\right).$$
Wegen
$$\lim_{d\to 0}\frac{1}{1-s}\left(1-\frac{1}{d^{s-1}}\right) = \frac{1}{1-s}\cdot(1) = \frac{1}{1-s}$$
ist
$$\int_0^1 \frac{1}{x^s}\,\mathrm{d}x = \lim_{d\to 0}\int_d^1 \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{1-s}.$$
Für $1<s$ existiert der obige Grenzwert nicht.

Satz 15.29 (Konvergenzkriterium für Integrale)**.**

i) *Sei* $g : (a, b) \to \mathbb{R}$ *mit* $g(x) \geq 0$*. Existiert das Integral* $\int_a^b g(x)\,\mathrm{d}x$ *und ist* $|f(x)| \leq g(x)$ *für alle* $x \in (a, b)$*, dann existiert auch* $\int_a^b f(x)\,\mathrm{d}x$*. Man sagt, das* g *eine (konvergente) Majorante von* f *ist.*

ii) *Sei* $g : (a, b) \to \mathbb{R}$ *mit* $g(x) \geq 0$*. Existiert das Integral* $\int_a^b g(x)\,\mathrm{d}x$ *nicht und ist* $f(x) \geq g(x)$ *für alle* $x \in (a, b)$*, dann existiert auch* $\int_a^b f(x)\,\mathrm{d}x$ *nicht. Man sagt, das* g *eine (divergente) Minorante von* f *ist.*

Satz 15.30 (Integralvergleichskriterium für Reihen)**.** *Sei* $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ *eine nicht-negative monoton fallende Funktion. Die Reihe* $\sum_{n=1}^{\infty} f(n)$ *konvergiert genau dann, wenn* $\int_1^{\infty} f(x)\,\mathrm{d}x$ *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 $T_n : \mathbb{R} \to \mathbb{R}$ 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 $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$ heißt *trigonometrisches Polynom* vom Grad n.

Bemerkung 16.2. *T_n ist periodisch mit Periode T, d.h.*

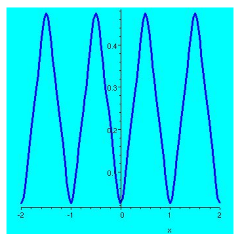
$$T_n(x + T) = T_n(x) \text{ für alle } x \in \mathbb{R}.$$

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 Peride $T = 1$ 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[37], Fourierkoeffizienten)**.** Sei $T > 0$ und $f : [-T/2, T/2] \to \mathbb{R}$ eine integrierbare Funktion, und sei $n \in \mathbb{N}$.

$$a_k = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos(k\,\omega\,t)\,\mathrm{d}t, \quad k = 0, \ldots, n$$
$$b_k = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin(k\,\omega\,t)\,\mathrm{d}t, \quad k = 1, \ldots, n$$

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:*

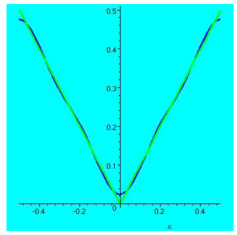
$$\int_{-T/2}^{T/2} \cos(k\,\omega\,x)\,\cos(l\,\omega\,x)\,\mathrm{d}x = \int_{-T/2}^{T/2} \sin(k\,\omega\,x)\,\sin(l\,\omega\,x)\,\mathrm{d}x = \begin{cases} T/2 & k = l \\ 0 & k \neq l \end{cases},$$
$$\int_{-T/2}^{T/2} \cos(k\,\omega\,x)\,\sin(l\,\omega\,x)\,\mathrm{d}x = 0,$$
$$\int_{-T/2}^{T/2} \cos(k\,\omega\,x)\,\mathrm{d}x = \int_{-T/2}^{T/2} \sin(k\,\omega\,x)\,\mathrm{d}x = 0.$$

[37] Jean Baptiste Joseph Fourier, 1768-1830

Satz 16.6. *Ist $f(x)$ eine gerade Funktion, so ist $b_k = 0$, $k = 1, \ldots, n$, und ist $f(x)$ eine ungerade Funktion so ist $a_k = 0$, $k = 0, \ldots, n$.*

Beispiel 16.7. Das Fourierpolynom 3-ten Grades der Funktion $f : [-1/2, 1/2] \to \mathbb{R}$ mit $f(x) = |x|$ ist das Polynom (siehe (16.3))

$$F_3(x) = \frac{1}{4} - \frac{2}{\pi^2}\cos(2\,\pi\,x) - \frac{2}{9\pi^2}\cos(6\pi\,x).$$



$F_3(x)$

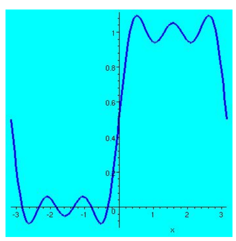
Definition 16.8 (Fourierreihe)**.** Sei $T > 0$, $\omega = 2\pi/T$, und sei $f : [-T/2, T/2] \to \mathbb{R}$ eine integrierbare Funktion mit Fourierkoeffizienten a_k, b_k. Dann heißt

$$F_\infty(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\left(a_k\,\cos(k\,\omega\,x) + b_k\,\sin(k\,\omega\,x)\right)$$

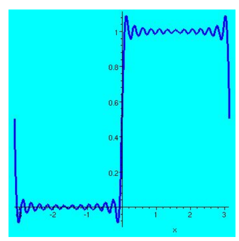
die *Fourierreihe* von f.

Beispiel 16.9. Sei $f : [-\pi, \pi] \to \mathbb{R}$ mit $f(t) = 1$ für $t > 0$ und $f(t) = 0$ für $t \leq 0$. Dann ist

$$F_\infty(x) = \frac{1}{2} + \sum_{k=1}^{\infty}\frac{1-(-1)^k}{k\,\pi}\sin(k\,x).$$



$F_5(x)$



$F_{25}(x)$

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

$$\langle f, g\rangle = \int_a^b f(x)\,g(x)\,\mathrm{d}x, \text{ und } \|f\| = \sqrt{\langle f, f\rangle}.$$

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*

$$\frac{1}{\sqrt{2\,\pi}}, \frac{\cos(k\,x)}{\sqrt{\pi}}, \frac{\sin(k\,x)}{\sqrt{\pi}}, \; k \in \mathbb{N}$$

bilden ein Orthonormalsystem (siehe (7.12) und Satz 16.5) in $C[-\pi, \pi]$. Insbesondere gilt für $g(x) = \frac{c_0}{2} + \sum_{k=1}^{n}(c_k\,\cos(k\,x) + d_k\,\sin(k\,x))$

$$\|g\|^2 = \pi\left(\frac{c_0^2}{2} + \sum_{k=1}^{n}\left(c_k^2 + d_k^2\right)\right).$$

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) $\lim_{n\to\infty} \|f - F_n\|^2 = 0$ *(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

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n),$$

wobei D eine Teilmenge des $\mathbb{R}^n$ ist. Statt $(x_1, \ldots, x_n)$ wird auch oft $\mathbf{x}$ geschrieben, und so $f(\mathbf{x})$ anstelle von $f(x_1, \ldots, x_n)$.

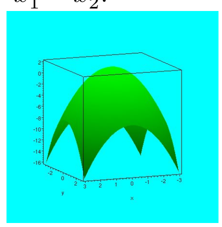
Bemerkung 17.2. *Für eine Funktion $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ ist der Graph*

$$\{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in D\}$$

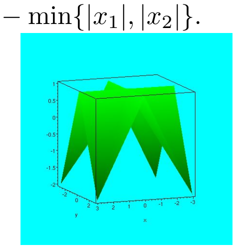
ein 3-dimensionales Gebilde (Gebirge). I.A. für $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ ein $(n+1)$-dimensionales Gebilde.

Beispiel 17.3.

i) $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(\mathbf{x}) = f(x_1, x_2) = 2 - x_1^2 - x_2^2$.



ii) $f : \mathbb{R}^2 \to \mathbb{R}$ 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* $\mathbf{x}^\star \in \mathbb{R}^n$, bezeichnet als $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star$, falls der Abstand zwischen $\mathbf{x}_k$ und $\mathbf{x}^\star$

$$\|\mathbf{x}_k - \mathbf{x}^\star\| = \sqrt{(\mathbf{x}_{k,1} - \mathbf{x}_1^\star)^2 + (\mathbf{x}_{k,2} - \mathbf{x}_2^\star)^2 + \cdots + (\mathbf{x}_{k,n} - \mathbf{x}_n^\star)^2}$$

gegen 0 konvergiert, d.h.

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star \Leftrightarrow \lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}^\star\| = 0.$$

Man schreibt auch $\mathbf{x}_k \to \mathbf{x}^\star$.

Bemerkung 17.5. $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star$ *ist also gleichbedeutend damit, dass jede Koordinate $\mathbf{x}_{k,i}$ gegen $\mathbf{x}_i^\star$ konvergiert, $0 \leq i \leq n$. Insbesondere übertragen sich die Eigenschaften von Folgen aus Kapitel 11 in „kanonischer Weise“ auf Folgen von Vektoren (s. z.B. Satz 11.12, Satz 11.15).*

Definition 17.6 (Grenzwert einer Funktion)**.** Sei $D \subseteq \mathbb{R}^n$, $f : D \to \mathbb{R}$ eine Funktion, und sei $\mathbf{x}^\star \in \mathbb{R}^n$. Wenn es ein $y^\star \in \mathbb{R}$ gibt, so dass für jede Folge $(\mathbf{x}_k)_{k \in \mathbb{N}_0}$ mit $\mathbf{x}_k \in D \setminus \{\mathbf{x}^\star\}$ und

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star$$

gilt, dass die Folge der Funktionswerte $(f(\mathbf{x}_k))_{k\in\mathbb{N}_0}$ gegen $y^\star$ konvergiert, also

$$\lim_{k\to\infty} f(\mathbf{x}_k) = y^\star,$$

dann heißt $y^\star$ *der Grenzwert von f für $\mathbf{x}$ gegen $\mathbf{x}^\star$*. Man schreibt dafür

$$\lim_{\mathbf{x}\to\mathbf{x}^\star} f(\mathbf{x}) = y^\star.$$

Hierbei ist auch $y^\star = \pm\infty$ erlaubt, und man spricht dann von bestimmter Divergenz (vgl. Def. 12.15).

Definition 17.7 (Stetigkeit)**.** Sei $D \subseteq \mathbb{R}^n$, $f : D \to \mathbb{R}$ eine Funktion, und sei $\mathbf{x}^\star \in D$. f heißt *stetig im Punkt (an der Stelle)* $\mathbf{x}^*$, falls

$$\lim_{\mathbf{x}\to\mathbf{x}^\star} f(\mathbf{x}) = f(\mathbf{x}^*),$$

d.h. der Grenzwert $\lim_{\mathbf{x}\to\mathbf{x}^\star} f(\mathbf{x})$ existiert und ist gleich $f(\mathbf{x}^\star)$. Dies bedeutet, für jede Folge $(\mathbf{x}_k)_{k\in\mathbb{N}_0}$, $\mathbf{x}_k \in D \setminus \{\mathbf{x}^\star\}$, mit $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star$ gilt

$$\lim_{k\to\infty} f(\mathbf{x}_k) = f(\lim_{k\to\infty} \mathbf{x}_k) = f(\mathbf{x}^\star).$$

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 $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}$ Funktionen, die in $\mathbf{x}^\star \in \mathbb{R}^n$ stetig sind. Dann sind auch die Funktionen*

$$\begin{aligned}&\alpha f(\mathbf{x}) \text{ mit } \alpha\in\mathbb{R}, \quad f(\mathbf{x}) \pm g(\mathbf{x}), \quad f(\mathbf{x})\cdot g(\mathbf{x}),\\ &\frac{f(\mathbf{x})}{g(\mathbf{x})}, \text{ falls } g(\mathbf{x}^\star) \neq 0,\end{aligned}$$

stetig in $\mathbf{x}^\star$.

Definition 17.10 (Vektorwertige Funktion)**.** Sei $D \subset \mathbb{R}^n$. Unter einer (vektorwertigen) Funktion $\mathbf{f} : D \to \mathbb{R}^m$ versteht man die Abbildung $\mathbf{x} \mapsto (\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x}), \ldots, \mathbf{f}_m(\mathbf{x}))$, wobei $\mathbf{f}_i : D \to \mathbb{R}$, $1 \leq i \leq m$, reellwertige Funktionen sind.

Bemerkung 17.11.

i) *Eine Matrix $A \in \mathbb{R}^{m\times n}$ beschreibt eine lineare Abbildung von $\mathbb{R}^n \to \mathbb{R}^m$ (vgl. Satz 6.5).*

ii) *Im Falle $n = 1$ nennt man f auch (parameterisierte) Kurven.*

Definition 17.12 (Grenzwert einer Funktion)**.** Sei $D \subseteq \mathbb{R}^n$, $\mathbf{f} : D \to \mathbb{R}^m$ eine Funktion, und sei $\mathbf{x}^\star \in \mathbb{R}^n$. Wenn es ein $\mathbf{y}^\star \in \mathbb{R}^m$ gibt, so dass für jede Folge $(\mathbf{x}_k)_{k\in\mathbb{N}_0}$ mit $\mathbf{x}_k \in D \setminus \{\mathbf{x}^\star\}$ und

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star$$

gilt, dass die Folge der Funktionswerte $(\mathbf{f}(\mathbf{x}_k))_{k\in\mathbb{N}_0}$ gegen $\mathbf{y}^\star$ konvergiert, also

$$\lim_{k\to\infty} \mathbf{f}(\mathbf{x}_k) = \mathbf{y}^\star,$$

dann heißt $\mathbf{y}^\star$ *der Grenzwert von $\mathbf{f}$ für $\mathbf{x}$ gegen $\mathbf{x}^\star$*. Man schreibt dafür

$$\lim_{\mathbf{x}\to\mathbf{x}^\star} \mathbf{f}(\mathbf{x}) = \mathbf{y}^\star.$$

Mit anderen Worten: $\mathbf{y}^\star$ ist der Grenzwert von $\mathbf{f}$ für $\mathbf{x}$ gegen $\mathbf{x}^\star$ genau dann, wenn $\mathbf{y}_i^\star$ der Grenzwert von $\mathbf{f}_i$ für $\mathbf{x}$ gegen $\mathbf{x}^\star$ und für $1 \leq i \leq m$ ist (vgl. Def. 17.6).

Definition 17.13 (Stetigkeit)**.** Sei $D \subseteq \mathbb{R}^n$, $\mathbf{f} : D \to \mathbb{R}^m$ eine Funktion, und sei $\mathbf{x}^\star \in D$. $\mathbf{f}$ heißt *stetig im Punkt (an der Stelle)* $\mathbf{x}^*$, falls

$$\lim_{\mathbf{x}\to\mathbf{x}^\star} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^*),$$

d.h. der Grenzwert $\lim_{\mathbf{x}\to\mathbf{x}^\star} \mathbf{f}(\mathbf{x})$ existiert und ist gleich $\mathbf{f}(\mathbf{x}^\star)$. Dies bedeutet, für jede Folge $(\mathbf{x}_k)_{k\in\mathbb{N}_0}$, $\mathbf{x}_k \in D \setminus \{\mathbf{x}^\star\}$, mit $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^\star$ gilt

$$\lim_{k\to\infty} \mathbf{f}(\mathbf{x}_k) = \mathbf{f}(\lim_{k\to\infty} \mathbf{x}_k) = \mathbf{f}(\mathbf{x}^\star).$$

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 $\mathbf{x}^\star$ stetig, wenn jede Funktion $\mathbf{f}_i : D \to \mathbb{R}$, $1 \le i \le n$, in $\mathbf{x}^\star$ stetig ist (vgl. Def. 17.7).

Definition 17.14 (ε-Umgebung, offene Menge)**.**

i) Für $\mathbf{x}^\star \in \mathbb{R}^n$ und $\varepsilon > 0$ heißt

$$B(x^\star, \varepsilon) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}^\star\| < \varepsilon\}$$

ε-Umgebung von $\mathbf{x}^\star$. Geometrisch ist dies die „offene Kugel mit Radius ε und Mittelpunkt $\mathbf{x}^\star$“.

ii) Eine Menge $D \subseteq \mathbb{R}^n$ heißt *offen*, falls es für alle $\mathbf{x}^\star \in D$ ein $\varepsilon > 0$ gibt mit $B(x^\star, \varepsilon) \subseteq D$.

Bemerkung 17.15. *Um nun (zunächst) für Funktionen $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbf{x}^\star$ durch eine Hyperebene approximieren (s. Bem. 5.9).*

Definition 17.16 (Partielle Ableitung(en))**.** Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, D offen, eine Funktion und $\mathbf{x}^\star \in D$.

i) Unter der *partiellen Ableitung* $\frac{\partial f}{\partial x_i}(\mathbf{x}^\star)$ *von f nach x_i* in einem Punkt $\mathbf{x}^\star$ 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

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^\star) = \lim_{h\to 0} \frac{f(\mathbf{x}^\star + h\,\mathbf{e}_i) - f(\mathbf{x}^\star)}{h}$$

existieren muss, wobei $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^\intercal$ den i-ten Einheitsvektor bezeichnet (s. Def. 5.20).

ii) Ist f in jedem Punkt $\mathbf{x}^\star \in D$ partiell nach x_i differenzierbar, dann heißt f *partiell nach x_i differenzierbar*, und die Funktion

$$\frac{\partial f}{\partial x_i} : D \to \mathbb{R} \text{ mit } x \mapsto \frac{\partial f}{\partial x_i}(\mathbf{x})$$

heißt *partielle Ableitung von f nach x_i*.

iii) Ist f partiell nach x_i differenzierbar für $1 \le i \le n$, 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 $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, D offen, eine partiell differenzierbare Funktion. Dann heißt der Vektor

$$\operatorname{grad} f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)$$

Gradient von f im Punkt $\mathbf{x}$.

Bemerkung 17.19. *Seien $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}$, D offen, partiell differenzierbare Funktionen. Dann gilt die Produktregel (vgl. Satz 13.6 iii))*

$$\operatorname{grad}(f \cdot g)(\mathbf{x}) = g(\mathbf{x}) \cdot \operatorname{grad} f(\mathbf{x}) + f(\mathbf{x}) \cdot \operatorname{grad} g(\mathbf{x}).$$

Definition 17.20 (Höhere partielle Ableitungen)**.** Sei $f : D \subset \mathbb{R}^n \to \mathbb{R}$, D offen, eine partiell differenzierbare Funktion. Sind alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} : D \to \mathbb{R}$ selbst wieder partiell differenzierbar, so heißt f zweimal partiell differenzierbar, und man schreibt statt $\frac{\partial \frac{\partial f}{\partial x_i}}{\partial x_j}$

$$\frac{\partial^2 f}{\partial x_j \partial x_i}$$

für die partielle Ableitung nach x_j der partielle Ableitung nach x_i.

Entsprechend sind höhere (k-te) partielle Ableitung definiert

$$\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}}.$$

Satz 17.21 (Schwarz[38])**.** *Ist $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, 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 $\mathbf{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, D offen, eine Abbildung. $\mathbf{f}$ heißt in $\mathbf{x}^\star \in D$ *(total) differenzierbar*, falls es eine Matrix $A(\mathbf{x}^\star) \in \mathbb{R}^{m \times n}$ gibt, so dass

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^\star) + A(\mathbf{x}^\star)\,(\mathbf{x} - \mathbf{x}^\star) + r(\mathbf{x})$$

mit

$$\lim_{\mathbf{x} \to \mathbf{x}^\star} \frac{r(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}^\star\|} = 0.$$

$A(\mathbf{x}^\star)$ heißt die *Ableitung oder das Differential von* $\mathbf{f}$ in $\mathbf{x}^\star$.

Definition 17.23 (Jacobi[39]-Matrix)**.** Sei $\mathbf{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, D offen, eine partiell differenzierbare Abbildung, d.h. alle Koordinaten-Funktionen $\mathbf{f}_i$ sind partiell differenzierbar, und sei $\mathbf{x}^\star \in D$. Die $(m \times n)$-Matrix

$$J_{\mathbf{f}}(\mathbf{x}^\star) = \begin{pmatrix} \frac{\partial \mathbf{f}_1}{\partial x_1}(\mathbf{x}^\star) & \ldots & \frac{\partial \mathbf{f}_1}{\partial x_n}(\mathbf{x}^\star) \\ \frac{\partial \mathbf{f}_2}{\partial x_1}(\mathbf{x}^\star) & \ldots & \frac{\partial \mathbf{f}_2}{\partial x_n}(\mathbf{x}^\star) \\ \vdots & \ldots & \vdots \\ \frac{\partial \mathbf{f}_m}{\partial x_1}(\mathbf{x}^\star) & \ldots & \frac{\partial \mathbf{f}_m}{\partial x_n}(\mathbf{x}^\star) \end{pmatrix}$$

heißt *Jacobi-Matrix* oder auch *Funktional-Matrix.*

[38] Hermann Amandus Schwarz, 1843-1921
[39] Carl Gustav Jacob Jacobi, 1804-1851

Bemerkung 17.24. *Ist $m = 1$, dann ist $J_{\mathbf{f}}(\mathbf{x}^\star) = \operatorname{grad}\mathbf{f}(\mathbf{x}^\star)$.*

Satz 17.25. *Sei $\mathbf{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, D offen, und sei $\mathbf{x}^\star \in D$.*

i) *Ist $\mathbf{f}$ in $\mathbf{x}^\star$ differenzierbar mit Ableitung $A(\mathbf{x}^\star)$. Dann ist $\mathbf{f}$ in $\mathbf{x}^\star$ stetig und es ist $J_{\mathbf{f}}(\mathbf{x}^\star) = A(\mathbf{x}^\star)$.*

ii) *Ist $\mathbf{f}_i$ in $\mathbf{x}^\star$ stetig partiell differenzierbar für $1 \le i \le m$, dann ist $\mathbf{f}$ in $\mathbf{x}^\star$ differenzierbar mit Ableitung $J_{\mathbf{f}}(\mathbf{x}^\star)$.*

Korollar 17.26. *Ist $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, D offen, stetig partiell differenzierbar, dann ist f stetig.*

Satz 17.27 (Kettenregel (vgl. Satz 13.6 v)))**.** *Sei $\mathbf{g} : C \subseteq \mathbb{R}^n \to \mathbb{R}^m$, C offen, in $x^\star \in C$ differenzierbar, und sei $\mathbf{f} : D \subseteq \mathbb{R}^m \to \mathbb{R}^k$, D offen, mit $\mathbf{g}(C) \subseteq D$, und es sei $\mathbf{f}$ in $\mathbf{g}(\mathbf{x}^\star)$ differenzierbar. Dann ist die Verknüpfung $\mathbf{f} \circ \mathbf{g} : C \to \mathbb{R}^k$ in $\mathbf{x}^\star$ differenzierbar, und es gilt:*

$$J_{\mathbf{f}\circ\mathbf{g}}(\mathbf{x}^\star) = J_{\mathbf{f}}(\mathbf{g}(\mathbf{x}^\star)) \cdot J_{\mathbf{g}}(\mathbf{x}^\star).$$

Definition 17.28 (Lokale Extrema (vgl. Def. 13.17))**.** Sei $D \subseteq \mathbb{R}^n$ und $f : D \to \mathbb{R}$ eine Funktion. $\mathbf{x}^\star \in D$ heißt

i) *lokales Maximum*, falls ein $\epsilon > 0$ existiert, so dass

$$f(\mathbf{x}) \le f(\mathbf{x}^\star) \text{ für alle } \mathbf{x} \in B(\mathbf{x}^\star, \epsilon) \cap D.$$

ii) *lokales Minimum*, falls ein $\epsilon > 0$ existiert, so dass

$$f(\mathbf{x}) \ge f(\mathbf{x}^\star) \text{ für alle } \mathbf{x} \in B(\mathbf{x}^\star, \epsilon) \cap D.$$

iii) *globales Maximum*, falls

$$f(\mathbf{x}) \le f(\mathbf{x}^\star) \text{ für alle } \mathbf{x} \in D.$$

iv) *globales Minimum*, falls

$$f(\mathbf{x}) \ge f(\mathbf{x}^\star) \text{ für alle } \mathbf{x} \in D.$$

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 $f : D \to \mathbb{R}$ partiell differenzierbar. Besitzt f in $\mathbf{x}^\star \in D$ ein lokales Extremum, dann ist $\operatorname{grad} f(\mathbf{x}^\star) = \mathbf{0}$.*

Definition 17.30 (Hesse[40]-Matrix)**.** Sei $D \subseteq \mathbb{R}^n$ offen, und sei $f : D \to \mathbb{R}$ eine 2-mal stetig partiell differenzierbare Funktion. Für $\mathbf{x}^\star \in D$ heißt die symmetrische $(n \times n)$-Matrix

$$H_f(\mathbf{x}^\star) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}^\star) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^\star) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}^\star) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^\star) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}^\star) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}^\star) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}^\star) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}^\star) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}^\star) \end{pmatrix}$$

Hesse-Matrix von f in $\mathbf{x}^\star$.

Satz 17.31 ((vgl. Satz 13.18 ii), iii) und Satz 9.37))**.** *Sei $D \subset \mathbb{R}^n$ offen, und sei $f : D \to \mathbb{R}$ eine 2-mal stetig partiell differenzierbare Funktion. Sei $\mathbf{x}^\star \in D$ mit $\operatorname{grad} f(\mathbf{x}^\star) = \mathbf{0}$.*

[40] Otto Hesse, 1811-1874

i) *Ist $H_f(\mathbf{x}^\star)$ positiv definit, dann ist $\mathbf{x}^\star$ ein lokales Minimum.*

ii) *Ist $H_f(\mathbf{x}^\star)$ negativ definit, dann ist $\mathbf{x}^\star$ ein lokales Maximum.*

iii) *Ist $H_f(\mathbf{x}^\star)$ weder positiv noch negativ semi-definit (d.h. indefinit), dann ist $\mathbf{x}^\star$ kein lokales Extremum, sondern ein Sattelpunkt.*

iv) *Ist $H_f(\mathbf{x}^\star)$ negativ semi-definit oder positiv-definit, dann ist keine Aussage möglich.*

Bemerkung 17.32. *Eine 2×2-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 abgeschossenen Intervallen vorstellen.*

Satz 17.35 (vgl. Satz 12.31)**.** *Jede stetige Funktion $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ nimmt auf einer kompakten Menge D ihr Maximum und Minimum an, d.h. es gibt $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in D$ mit*

$$\begin{aligned} f(\tilde{\mathbf{x}}) &= \max\{f(\mathbf{x}) : x \in D\}, \textit{und} \\ f(\hat{\mathbf{x}}) &= \min\{f(\mathbf{x}) : x \in D\}. \end{aligned}$$

Satz 17.36 (Lokale Extrema unter Nebenbedingungen)**.** *Sei $D \subseteq \mathbb{R}^n$ offen, und seien $f, g : D \to \mathbb{R}$ stetig partiell differenzierbare Funktionen. Sei $M = \{x \in D : g(x) = 0\}$ und sei $\mathbf{x}^\star \in M$ mit $\operatorname{grad} g(\mathbf{x}^\star) \neq \mathbf{0}$. Ferner sei $\mathbf{x}^\star$ ein lokales Maximum (Minimum) von f unter der Nebenbedingung $g(\mathbf{x}) = 0$, d.h. es gibt ein $\epsilon > 0$, so dass gilt*

$$f(\mathbf{x}) \leq f(\mathbf{x}^\star) \quad (\textit{bzw. } f(\mathbf{x}) \geq f(\mathbf{x}^\star)) \;\; \textit{für alle } \mathbf{x} \in B(\mathbf{x}^\star, \epsilon) \cap M.$$

Dann gibt es ein $\lambda \in \mathbb{R}$ mit

$$\operatorname{grad} f(\mathbf{x}^\star) = \lambda \operatorname{grad} g(\mathbf{x}^\star).$$

Man nennt λ einen Lagrange[41]-Multiplikator.

[41] 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

$$Q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] = \{\mathbf{x} \in \mathbb{R}^n : a_i \leq x_i \leq b_1,\ 1 \leq i \leq n\}.$$

Quader (mit Kantenlängen $b_i - a_i$, $1 \leq i \leq n$). Das *Volumen* $\mathrm{vol}(Q)$ des Quaders ist gegeben durch

$$\mathrm{vol}(Q) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \ldots \cdot (b_n - a_n).$$

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 $\mathrm{vol}(Q)$ ist der Flächeninhalt von Q.*

Definition 18.3 (Zerlegung im $\mathbb{R}^n$)**.** Sei $Q \subset \mathbb{R}^n$ ein Quader. Eine Menge $Z = \{Q_1, \ldots, Q_m\}$ von Quadern $Q_i \subseteq Q$ heißt *Zerlegung* von Q, falls

$$Q = \bigcup_{i=1}^{m} Q_i,$$

und je zwei verschiedene Quader 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 $f : Q \to \mathbb{R}$ eine beschränkte Funktion. Sei $Z = \{Q_1, \ldots, Q_m\}$ eine Zerlegung von Q. Dann heißt

i) $U_f(Z) = \sum_{i=1}^{m} \mathrm{vol}(Q_i) \cdot \inf\{f(x) : x \in Q_i\}$ *Untersumme* von f bzgl. Z.

ii) $O_f(Z) = \sum_{i=1}^{m} \mathrm{vol}(Q_i) \cdot \sup\{f(x) : x \in Q_i\}$ *Obersumme* von f bzgl. Z.

(vgl. Def. 15.2)

Definition 18.5 ((Riemann[42])-Integral im $\mathbb{R}^n$)**.** Sei $Q \subset \mathbb{R}^n$ ein Quader, und sei $f : Q \to \mathbb{R}$ eine beschränkte Funktion.

i) $U_f = \sup\{U_f(Z) : Z \text{ Zerlegung von } Q\}$ heißt *(Riemann'sches) Unterintegral* von f über Q.

ii) $O_f = \inf\{O_f(Z) : Z \text{ Zerlegung von } Q\}$ 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

$$\int_Q f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

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 $f : D \subset \mathbb{R}^n \to \mathbb{R}$ eine beschränkte Funktion auf der beschränkten Menge $D \subset \mathbb{R}^n$. Wähle einen Quader Q mit $D \subseteq Q$ und betrachte*

$$\tilde{f} : Q \to \mathbb{R} \text{ mit } \tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & : \mathbf{x} \in D, \\ 0 & : \mathbf{x} \in Q \setminus D. \end{cases}$$

Dann setzt man

$$\int_D f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_Q \tilde{f}(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

falls die rechte Seite existiert.

[42] 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)

$$\mathrm{vol}(D) = \int_D 1 \,\mathrm{d}\mathbf{x}$$

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. Quader).*

Satz 18.10 (Fubini[43])**.**

i) *$n = 2$: Sei $Q = [a_1, b_1] \times [a_2, b_2]$ und $f : Q \to \mathbb{R}$ integrierbar. Dann gilt*

$$\int_Q f(\mathbf{x}) \,\mathrm{d}\mathbf{x} = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) \,\mathrm{d}x_1 \right) \mathrm{d}x_2 = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) \,\mathrm{d}x_2 \right) \mathrm{d}x_1.$$

ii) *$n \geq 2$: $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ und $f : Q \to \mathbb{R}$ integrierbar. Dann gilt*

$$\int_Q f(\mathbf{x}) \,\mathrm{d}\mathbf{x} = \int_{a_n}^{b_n} \left(\int_{a_{n-1}}^{b_{n-1}} \left(\cdots \left(\int_{a_1}^{b_1} f(x_1, x_2, \ldots, x_n) \mathrm{d}x_1 \right) \cdots \right) \mathrm{d}x_{n-1} \right) \mathrm{d}x_n,$$

wobei man auf der rechten Seite auch jede andere Reihenfolge beim Integrieren verwenden darf.

Beispiel 18.11. Sei $f : Q = [0, 1] \times [0, 2] \to \mathbb{R}$ mit $f(x_1, x_2) = (2 - x_1\, x_2 + x_2)$

$$\begin{aligned}\int_Q f(\mathbf{x}) \,\mathrm{d}\mathbf{x} &= \int_0^1 \left(\int_0^2 f(x_1, x_2) \,\mathrm{d}x_2 \right) \mathrm{d}x_1 = \int_0^1 \left(\int_0^2 (2 - x_1\, x_2 + x_2) \,\mathrm{d}x_2 \right) \mathrm{d}x_1 \\ &= \int_0^1 \left(2\, x_2 - \frac{1}{2} x_1\, x_2^2 + \frac{1}{2}\, x_2^2 \right) \Big|_0^2 \mathrm{d}x_1 = \int_0^1 (6 - 2\, x_1) \,\mathrm{d}x_1 = \left(6\, x_1 - x_1^2\right) \Big|_0^1 = 5.\end{aligned}$$

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 $\alpha_1, \beta_1 : [a, b] \to \mathbb{R}$ stetige Funktionen. Entsprechend definiert man einen Normalbereich bzgl. der y-Achse, bzw. Normalbereiche in Dimensionen ≥ 3. Z.B. ist

$$D = \left\{ (x_1, x_2, x_3)^\intercal \in \mathbb{R}^3 : a \leq x_1 \leq b, \alpha_1(x_1) \leq x_2 \leq \beta_1(x_1), \alpha_2(x_1, x_2) \leq x_3 \leq \beta_2(x_1, x_2) \right\}$$

ein Normalbereich in $\mathbb{R}^3$. Hierbei sind zusätzlich $\alpha_2, \beta_2 : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ stetige Funktionen.

Satz 18.13. *Sei $D \subset \mathbb{R}^2$ ein Normalbereich, und sei $f : D \to \mathbb{R}$ eine integrierbare Funktion auf D. Dann gilt:*

$$\int_D f(\mathbf{x}) \,\mathrm{d}\,\mathbf{x} = \int_a^b \left(\int_{\alpha_1(x_1)}^{\beta_1(x_1)} f(x_1, x_2) \,\mathrm{d}\,x_2 \right) \mathrm{d}\,x_1.$$

Eine analoge Aussage gilt für Normalbereiche in Dimension ≥ 3. Mit der Notation aus Definition 18.12 gilt z.B. für $f : D \subset \mathbb{R}^3 \to \mathbb{R}$:

$$\int_D f(\mathbf{x}) \,\mathrm{d}\,\mathbf{x} = \int_a^b \left(\int_{\alpha_1(x_1)}^{\beta_1(x_1)} \left(\int_{\alpha_2(x_1, x_2)}^{\beta_2(x_1, x_2)} f(x_1, x_2, x_3) \,\mathrm{d}\,x_3 \right) \mathrm{d}\,x_2 \right) \mathrm{d}\,x_1.$$

[43] 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 $f : T \to \mathbb{R}$ mit $f(\mathbf{x}) = 2 - x_1^2 - x_2^2$.

$$\begin{aligned}
\int_T f(\mathbf{x})\,\mathrm{d}\mathbf{x} &= \int_0^1 \left(\int_0^{1-x_1} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1 = \int_0^1 \left(\int_0^{1-x_1} \left(2 - x_1^2 - x_2^2\right)\mathrm{d}x_2\right)\mathrm{d}x_1 \\
&= \int_0^1 \left(2\,x_2 - x_2\,x_1^2 - \frac{1}{3}\,x_2^3\right)\Big|_0^{1-x_1}\mathrm{d}x_1 = \int_0^1 \left(2\,(1-x_1) - (1-x_1)\,x_1^2 - \frac{1}{3}(1-x_1)^3\right)\mathrm{d}x_1 \\
&= \int_0^1 \left(\frac{5}{3} - x_1 - 2\,x_1^2 + \frac{4}{3}\,x_1^3\right)\mathrm{d}x_1 = \left(\frac{5}{3}x_1 - \frac{1}{2}x_1^2 - \frac{2}{3}\,x_1^3 + \frac{1}{3}\,x_1^4\right)\Big|_0^1 = \frac{5}{6}.
\end{aligned}$$

Satz 18.15. *Sei $f : D \to \mathbb{R}$ integrierbar. Sei $A \in \mathbb{R}^{n\times n}$ mit $\det A \neq 0$, und sei $t \in \mathbb{R}^n$. Sei $\widetilde{D} \subset \mathbb{R}^n$ mit $D = A\,\widetilde{D} + t = \{A\,x + t : x \in \widetilde{D}\}$. Dann gilt*

$$\int_D f(\mathbf{x})\,\mathrm{d}\mathbf{x} = |\det A| \cdot \int_{\widetilde{D}} f(A\,\mathbf{x} + t)\,\mathrm{d}\mathbf{x}.$$

Insbesondere ist $\mathrm{vol}(D) = |\det A|\,\mathrm{vol}(\widetilde{D})$.

Korollar 18.16. *Sei $\widetilde{D} \subset \mathbb{R}^n$ beschränkt, $\lambda \in \mathbb{R}$, $t \in \mathbb{R}^n$ und*

$$\lambda\,\widetilde{D} + t = \{\lambda\,\mathbf{x} + t : \mathbf{x} \in \widetilde{D}\}.$$

Dann gilt

$$\mathrm{vol}(\lambda\,\widetilde{D} + t) = |\lambda|^n\,\mathrm{vol}(\widetilde{D}).$$

Satz 18.17 (Transformationssatz)**.** *Sei $f : D \to \mathbb{R}$ Riemann-integrierbar. Sei $\mathbf{T} : \widetilde{D} \to D$ eine bijektive Abbildung von einer Menge $\widetilde{D} \subset \mathbb{R}^n$ nach D, die stetig partiell differenzierbar ist und deren Umkehrabbildung ebenfalls stetig partiell differenzierbar ist. Dann gilt*

$$\int_D f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\widetilde{D}} f(\,\mathbf{T}(\mathbf{x})\,) \cdot |\det J_{\mathbf{T}}(\mathbf{x})|\,\mathrm{d}\mathbf{x}.$$

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)- Zylinderkoordianten)**.**

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 7.8).

Beispiel 18.20. Volumen der Kugel $D = B(\mathbf{0}, 1) = \{\mathbf{x} \in \mathbb{R}^3 : \|x\| \leq 1\}$. Sei $\widetilde{D} = [0,1] \times [0, 2\pi) \times [0, \pi]$ und $\mathbf{T} : \widetilde{D} \to B(\mathbf{0}, 1)$ 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

$$\begin{aligned} \operatorname{vol}(B(\mathbf{0},1)) &= \int_{B(\mathbf{0},1)} 1\,\mathrm{d}\mathbf{x} = \int_{\widetilde{D}} 1\,|\det J_{\mathbf{T}}(\mathbf{x})|\,\mathrm{d}\mathbf{x} = \int_{\widetilde{D}} |r^2\,\sin\vartheta|\,\mathrm{d}\mathbf{x} \\ &= \int_0^1 \left(\int_0^{2\pi} \left(\int_0^{\pi} |r^2\,\sin\vartheta|\,\mathrm{d}\,\vartheta \right) \mathrm{d}\,\phi \right) \mathrm{d}\,r. \end{aligned}$$

$$\begin{aligned} &= \int_0^1 \left(\int_0^{2\pi} \left(r^2(-\cos\vartheta)\Big|_0^{\pi} \right) \mathrm{d}\,\phi \right) \mathrm{d}\,r = \int_0^1 \left(\int_0^{2\pi} \left(2\,r^2\right) \mathrm{d}\,\phi \right) \mathrm{d}\,r = \int_0^1 \left(4\,\pi\,r^2\right) \mathrm{d}\,r \\ &= \left(\frac{4}{3}\,\pi\,r^3 \right)\Big|_0^1 = \frac{4}{3}\pi. \end{aligned}$$

Definition 18.21 (Vektorwertige Integrale)**.** Sei $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$, und es sei jede einzelne Koordinatenfunktion $\mathbf{f}_i : D \to \mathbb{R}$ Riemann-integrierbar, $1 \leq i \leq m$. Dann ist

$$\int_D \mathbf{f}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \begin{pmatrix} \int_D \mathbf{f}_1(\mathbf{x})\,\mathrm{d}\mathbf{x} \\ \int_D \mathbf{f}_2(\mathbf{x})\,\mathrm{d}\mathbf{x} \\ \vdots \\ \int_D \mathbf{f}_m(\mathbf{x})\,\mathrm{d}\mathbf{x} \end{pmatrix}.$$

Definition 18.22 (Schwerpunkt)**.** Sei $D \subset \mathbb{R}^n$ mit $\operatorname{vol}(D) > 0$. Dann heißt der Punkt

$$\frac{1}{\operatorname{vol}(D)} \int_D \mathbf{x}\,\mathrm{d}\mathbf{x}$$

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 $\operatorname{vol}(T) = 1/2$, und für die erste Koordinate des Schwerpunktes $(x_1^\star, x_2^\star)$ ergibt sich

$$\begin{aligned} x_1^\star &= \frac{1}{\frac{1}{2}} \int_T x_1 \mathrm{d}\mathbf{x} = 2 \left(\int_0^1 \left(\int_0^{1-x_1} x_1\,\mathrm{d}x_2 \right) \mathrm{d}x_1 \right) \\ &= 2 \left(\int_0^1 \left(x_1\,(1 - x_1)\right) \mathrm{d}x_1 \right) = 2 \left(\int_0^1 \left(x_1 - x_1^2\right) \mathrm{d}x_1 \right) \\ &= 2 \left(\frac{1}{2} x_1^2 - \frac{1}{3} x_1^3 \right)\Big|_0^1 = \frac{1}{3}. \end{aligned}$$

Ebenso findet man für die 2.te Koordinate $x_2^\star = \frac{1}{3}$.

Index