

# **Beweisarchiv**

April 2022

Dieses Heft steht unter der Lizenz Creative Commons CC0.

# Inhaltsverzeichnis

<b>1 Grundlagen</b>	<b>5</b>
1.1 Aussagenlogik . . . . .	5
1.2 Prädikatenlogik . . . . .	7
1.3 Mengenlehre . . . . .	11
1.3.1 Definitionen . . . . .	11
1.3.2 Rechenregeln . . . . .	11
1.4 Relationen . . . . .	14
1.5 Abbildungen . . . . .	15
1.5.1 Definitionen . . . . .	15
1.5.2 Grundlagen . . . . .	15
1.5.3 Kardinalzahlen . . . . .	21
<b>2 Analysis</b>	<b>27</b>
2.1 Folgen . . . . .	27
2.1.1 Konvergenz . . . . .	27
2.1.2 Wachstum und Landau-Symbole . . . . .	30
2.2 Stetige Funktionen . . . . .	30
2.3 Differentialrechnung . . . . .	34
2.3.1 Ableitungsregeln . . . . .	34
2.3.2 Glatte Funktionen . . . . .	36
2.3.3 Richtungsableitung . . . . .	36
2.4 Fixpunkt-Iterationen . . . . .	38
<b>3 Topologie</b>	<b>39</b>
3.1 Grundbegriffe . . . . .	39
3.1.1 Definitionen . . . . .	39
3.1.2 Elementares . . . . .	40
3.2 Metrische Räume . . . . .	41
3.2.1 Metrische Räume . . . . .	41
3.2.2 Normierte Räume . . . . .	41
3.2.3 Homöomorphismen . . . . .	42
3.3 Übungen . . . . .	42
<b>4 Lineare Algebra</b>	<b>45</b>
4.1 Matrizen . . . . .	45
4.1.1 Definitionen . . . . .	45
4.1.2 Rechenregeln . . . . .	45
4.1.3 Rechenregeln für komplexe Matrizen . . . . .	46
4.2 Eigenwerte . . . . .	46
4.2.1 Quadratische Matrizen . . . . .	48
4.3 Bilinearformen . . . . .	49
4.4 Euklidische Geometrie . . . . .	50
<b>5 Algebra</b>	<b>53</b>
5.1 Gruppentheorie . . . . .	53
5.1.1 Grundlagen . . . . .	53
5.2 Ringtheorie . . . . .	55
5.2.1 Grundlagen . . . . .	55
5.2.2 Ringhomomorphismen . . . . .	57

## Inhaltsverzeichnis

5.3	Polynomringe	57
5.3.1	Einsetzungshomomorphismus	57
5.4	Körper	58
5.4.1	Geordnete Körper	58
<b>6</b>	<b>Kombinatorik</b>	<b>61</b>
6.1	Endliche Mengen	61
6.1.1	Indikatorfunktion	61
6.2	Summen	62
<b>7</b>	<b>Wahrscheinlichkeitsrechnung</b>	<b>63</b>
7.1	Diskrete Wahrscheinlichkeitsräume	63
7.2	Allgemeine Wahrscheinlichkeitsräume	68
7.3	Stochastische Prozesse	69
7.3.1	Markow-Prozesse mit endlichem Zustandsraum	69

# 1 Grundlagen

## 1.1 Aussagenlogik

**Satz 1.1 (mp: Modus ponens).** Es gilt  $(A \Rightarrow B) \wedge A \Rightarrow B$ .

**Beweis 1 (Natürliches Schließen).**

Zu  $\{A \Rightarrow B, A\} \vdash B$ . Trivial, da eine Inferenzregel des Kalküls. Schematisch:

$$\frac{A \Rightarrow B \quad A}{B}$$

Programmterm:

$$(A \rightarrow B) \times A \rightarrow B, (f, a) \mapsto f(a). \square$$

**Beweis 2 (LEM, boolesche Algebra).** Man darf rechnen

$$\begin{aligned}(A \Rightarrow B) \wedge A \Rightarrow B &\equiv \neg((\neg A \vee B) \wedge A) \vee B \equiv \neg(\neg A \vee B) \vee \neg A \vee B \\ &\equiv \neg\varphi \vee \varphi \equiv 1,\end{aligned}$$

wobei  $\varphi := \neg A \vee B$ .  $\square$

**Satz 1.2 (bool-cl: Kommutativgesetze).** Es gilt

$$A \wedge B \iff B \wedge A, \tag{1.1}$$

$$A \vee B \iff B \vee A. \tag{1.2}$$

**Beweis (Natürliches Schließen).**

Zu  $A \wedge B \vdash B \wedge A$ . Schematisch:

$$\frac{\frac{A \wedge B}{A} \quad \frac{A \wedge B}{B}}{B \wedge A}$$

Programmterm:

$$A \times B \rightarrow B \times A, (a, b) \mapsto (b, a).$$

Zu  $A \vee B \vdash B \vee A$ . Schematisch:

$$\frac{A \vee B \quad \frac{\overline{A}^1}{B \vee A} \quad \frac{\overline{B}^1}{B \vee A}}{B \vee A}_1$$

Programmterm:

$$A + B \rightarrow B + A, s \mapsto \mathbf{match} \, s \begin{cases} \text{inl}(a) \mapsto \text{inr}(a), \\ \text{inr}(b) \mapsto \text{inl}(b). \end{cases}$$

Vertauschen von  $A, B$  erbringt jeweils die umgekehrte Implikation.  $\square$

**Satz 1.3 (bool-dl: Distributivgesetze).** Es gilt:

$$A \wedge (B \vee C) \iff A \wedge B \vee A \wedge C, \quad (1.3)$$

$$A \vee (B \wedge C) \iff (A \vee B) \wedge (A \vee C). \quad (1.4)$$

**Beweis (Natürliches Schließen).**

Zu  $A \wedge (B \vee C) \vdash A \wedge B \vee A \wedge C$ . Programmterm:

$$A \times (B + C) \rightarrow A \times B + A \times C, (a, s) \mapsto \mathbf{match} s \begin{cases} \text{inl}(b) \mapsto \text{inl}((a, b)), \\ \text{inr}(c) \mapsto \text{inr}((a, c)). \end{cases}$$

Zu  $A \wedge B \vee A \wedge C \vdash A \wedge (B \vee C)$ . Programmterm:

$$A \times B + A \times C \rightarrow A \times (B + C), s \mapsto \mathbf{match} s \begin{cases} \text{inl}((a, b)) \mapsto (a, \text{inl}(b)), \\ \text{inr}((a, c)) \mapsto (a, \text{inr}(c)). \end{cases}$$

Zu  $A \vee (B \wedge C) \vdash (A \vee B) \wedge (A \vee C)$ . Programmterm:

$$A + B \times C \rightarrow (A + B) \times (A + C), s \mapsto \mathbf{match} s \begin{cases} \text{inl}(a) \mapsto (\text{inl}(a), \text{inl}(a)), \\ \text{inr}((b, c)) \mapsto (\text{inr}(b), \text{inr}(c)). \end{cases}$$

Zu  $(A \vee B) \wedge (A \vee C) \vdash A \vee (B \wedge C)$ . Programmterm:

$$(A + B) \times (A + C) \rightarrow A + B \times C, t \mapsto \mathbf{match} t \begin{cases} (\text{inl}(a), s) \mapsto \text{inl}(a), \\ (\text{inr}(b), \text{inl}(a)) \mapsto \text{inl}(a), \\ (\text{inr}(b), \text{inr}(c)) \mapsto \text{inr}((b, c)). \end{cases}$$

Sämtliche Teilaussagen sind bewiesen.  $\square$

**Axiom 1.4 (PE: Principle of explosion).** Eine falsche Aussage impliziert jede beliebige Aussage, kurz

$$\vdash (0 \implies A).$$

Bemerkung: Dieses Prinzip erlaubt Programmterme mit leerem Pattern matching, so dass ein Zeuge für  $0 \rightarrow A$  konstruiert werden kann.

**Axiom 1.5 (LEM: Satz vom ausgeschlossenen Dritten).**

Entweder gilt eine Aussage, oder ihre Negation gilt, kurz

$$\vdash A \vee \neg A.$$

Bemerkung: Zur Schaffung von Klarheit sollte ein Beweis die Markierung LEM bekommen, wenn transitive Abhängigkeit zu diesem Axiom besteht. Verzichtet keiner der Beweise eines Satzes auf LEM, sollte der Satz ebenfalls mit LEM markiert werden.

**Axiom 1.6 (Beseitigung der Doppelnegation).**

Die Doppelnegation einer Aussage  $A$  impliziert die Aussage  $A$ , kurz

$$\vdash (\neg \neg A \implies A).$$

**Satz 1.7.** Das Axiom 1.5 (LEM) zieht 1.6 nach sich.

**Beweis (Natürliches Schließen).** Zu  $\neg A \vee A \vdash (\neg \neg A \implies A)$ . Gemäß Axiom 1.4 (PE) existiert ein Zeuge  $\text{ex}(A)$  für  $0 \rightarrow A$ . Damit lässt sich der Programmterm

$$(A \rightarrow 0) + A \rightarrow (((A \rightarrow 0) \rightarrow 0) \rightarrow A), s \mapsto \mathbf{match} s \begin{cases} \text{inl} f \mapsto g \mapsto \text{ex}(A)(g(f)), \\ \text{inr} a \mapsto g \mapsto a. \end{cases}$$

konstruieren.  $\square$

## 1.2 Prädikatenlogik

**Definition 1.1 (bounded: beschränkte Quantifizierung).**

$$(\forall x \in M: P(x)) :\iff \forall x: (x \in M \implies P(x)), \quad (1.5)$$

$$(\exists x \in M: P(x)) :\iff \exists x: (x \in M \wedge P(x)). \quad (1.6)$$

**Satz 1.8 (general-dl: allgemeine Distributivgesetze).** Es gilt:

$$A \wedge (\exists x: P(x)) \iff \exists x: (A \wedge P(x)), \quad (1.7)$$

$$A \vee (\forall x: P(x)) \iff \forall x: (A \vee P(x)). \quad (1.8)$$

**Beweis (Natürliches Schließen).** Schematisch. Zur ersten Äquivalenz:

$$\frac{\frac{\frac{A \quad \overline{P(a)}^1}{\exists x: P(x)} \quad \overline{A \wedge P(a)}^1}{A \wedge P(a)}^1 \quad \frac{\frac{\overline{A \wedge P(a)}^1 \quad \overline{A \wedge P(a)}^1}{A} \quad \frac{\overline{A \wedge P(a)}^1 \quad \overline{P(a)}^1}{\exists x: P(x)}}{\frac{A \wedge \exists x: P(x)}{A \wedge \exists x: P(x)}}^1$$

Zur zweiten Äquivalenz:

$$\frac{\frac{\frac{\overline{A}^1}{A \vee P(x)} \quad \frac{\overline{A}^1}{A \vee P(x)}}{A \vee (\forall x: P(x))} \quad \frac{\frac{\overline{A}^1}{A \vee P(x)} \quad \frac{\overline{P(x)}^1}{\forall x: P(x)}}{\forall x: (A \vee P(x))}^1 \quad \frac{\frac{\overline{A}^1}{A \vee P(x)} \quad \frac{\overline{P(x)}^1}{\forall x: P(x)}}{A \vee \forall x: P(x)}^1$$

Die linken Schemata zeigen jeweils die Implikation von links nach rechts, die rechten jeweils die Implikation von rechts nach links.  $\square$

**Satz 1.9 (exists-dl: Distributivgesetz).** Es gilt:

$$(\exists x: P(x) \vee Q(x)) \iff (\exists x: P(x)) \vee (\exists x: Q(x)).$$

**Beweis (Natürliches Schließen).** Schematisch.

Die Implikation von links nach rechts:

$$\frac{\frac{\frac{\exists x: P(x) \vee Q(x)}{P(a) \vee Q(a)}^1 \quad \overline{P(a) \vee Q(a)}^1}{\overline{P(a) \vee Q(a)}^1} \quad \frac{\overline{P(a)}^2}{\exists x: P(x)} \quad \frac{\overline{Q(a)}^2}{\exists x: Q(x)}}{\frac{(\exists x: P(x)) \vee (\exists x: Q(x))}{(\exists x: P(x)) \vee (\exists x: Q(x))}}^2$$

Die Implikation von rechts nach links:

$$\frac{\frac{\frac{\overline{(\exists x: P(x)) \vee (\exists x: Q(x))}}{\overline{P(a) \vee Q(a)}}^1 \quad \overline{P(a)}^2}{\overline{P(a) \vee Q(a)}}^2 \quad \frac{\frac{\overline{(\exists x: P(x)) \vee (\exists x: Q(x))}}{\overline{P(b) \vee Q(b)}}^1 \quad \overline{Q(b)}^3}{\overline{P(b) \vee Q(b)}}^3}{\overline{P(a) \vee Q(a)}^2 \quad \overline{P(b) \vee Q(b)}^3}^1$$

Man beachte, dass die Zeugen  $a, b$  unterschiedlich sein können.  $\square$

**Satz 1.10 (exists-asym-dl: asymmetrisches Distributivgesetz).** Es gilt:

$$(\exists x: P(x) \wedge Q(x)) \implies (\exists x: P(x)) \wedge (\exists x: Q(x)).$$

**Beweis (Natürliches Schließen).** Schematisch:

$$\frac{\frac{\frac{P(a) \wedge Q(a)}{P(a)}^1 \quad \frac{P(a) \wedge Q(a)}{Q(a)}^1}{\exists x: P(x) \quad \exists x: Q(x)} \quad \frac{(\exists x: P(x)) \wedge (\exists x: Q(x))}{(\exists x: P(x)) \wedge (\exists x: Q(x))}^1$$

In Worten: Weil aufgrund der Prämisse ein Zeuge  $a$  mit sowohl  $P(a)$  als auch  $Q(a)$  vorliegt, dürfen wir schließen, dass die Existenzaussagen  $\exists x: P(x)$  und  $\exists x: Q(x)$  erfüllt sind.  $\square$

**Satz 1.11 (dm-g1: 1. allgemeines de Morgansches Gesetz).** Es gilt

$$(\neg \exists x: P(x)) \iff (\forall x: \neg P(x)).$$

**Beweis.** Unter Spezialisierung von Satz 1.13 findet sich die äquivalente Umformung

$$\neg \exists x: P(x) \equiv (\exists x: P(x)) \Rightarrow 0 \equiv \forall x: (P(x) \Rightarrow 0) \equiv \forall x: \neg P(x). \quad \square$$

**Satz 1.12 (dm-g2: 2. allgemeines de Morgansches Gesetz).** Es gilt

$$(\neg \forall x: P(x)) \iff (\exists x: \neg P(x)).$$

**Beweis (LEM).** Nutzung von LEM und Satz 1.11 gestattet die äquivalente Umformung

$$\neg \forall x: P(x) \equiv \neg \forall x: \neg \neg P(x) \equiv \neg \neg \exists x: \neg P(x) \equiv \exists x: \neg P(x). \quad \square$$

**Satz 1.13.** Es gilt:

$$(\forall x: (P(x) \implies A)) \iff ((\exists x: P(x)) \implies A).$$

**Beweis 1 (Natürliches Schließen).** Schematisch:

$$\frac{\frac{\forall x: (P(x) \implies A) \quad \frac{\overline{\exists x: P(x)}}{P(a)}^1 \quad \frac{\overline{P(a)}}{P(a)}^2}{P(a) \implies A} \quad \frac{A}{(\exists x: P(x)) \implies A}^1 \quad \frac{\frac{(\exists x: P(x)) \implies A \quad \frac{\overline{P(x)}}{\exists x: P(x)}^1}{A} \quad \frac{A}{P(x) \implies A}^1}{\forall x: (P(x) \implies A)}$$

Die linke Seite zeigt die Implikation von links nach rechts, die rechte die Implikation von rechts nach links.  $\square$

**Beweis 2 (LEM, boolesche Algebra).** Unter Nutzung von Satz 1.8 (general-dl) und Satz 1.11 (dm-g1) gilt

$$\begin{aligned} \forall x: (P(x) \implies A) &\equiv \forall x: (\neg P(x) \vee A) \equiv (\forall x: \neg P(x)) \vee A \\ &\equiv \neg(\exists x: P(x)) \vee A \equiv (\exists x: P(x)) \Rightarrow A. \quad \square \end{aligned}$$



**Satz 1.14 (exists-cl: Kommutativgesetz).** Es gilt:

$$(\exists x: \exists y: P(x, y)) \iff (\exists y: \exists x: P(x, y)).$$

**Beweis (Natürliches Schließen).** Die Implikation von links nach rechts:

$$\frac{\frac{\frac{\exists x: \exists y: P(x, b) \quad \overline{P(a, b)}}{P(a, b)}^1}{\exists x: P(x, y)}}{\exists y: \exists x: P(x, y)}$$

Die Implikation von rechts nach links geht analog.  $\square$

**Satz 1.15 (all-cl: Kommutativgesetz).** Es gilt:

$$(\forall x: \forall y: P(x, y)) \iff (\forall y: \forall x: P(x, y)).$$

**Beweis (Natürliches Schließen).** Die Implikationen von links nach rechts:

$$\frac{\frac{\frac{\forall x: \forall y: P(x, y)}{P(x, y)}}{\forall x: P(x, y)}}{\forall y: \forall x: P(x, y)}$$

Die Implikation von rechts nach links geht analog.  $\square$

**Satz 1.16 (bounded-general-dl: allgemeine Distributivgesetze).** Es gilt:

$$A \wedge (\exists x \in M: P(x)) \iff (\exists x \in M: A \wedge P(x)), \quad (1.9)$$

$$A \vee (\forall x \in M: P(x)) \iff (\forall x \in M: A \vee P(x)). \quad (1.10)$$

**Beweis.** Nach Def. 1.1 (bounded) und Satz 1.8 (general-dl) gilt:

$$\begin{aligned} A \wedge \exists x \in M: P(x) &\equiv A \wedge \exists x: x \in M \wedge P(x) \equiv \exists x: A \wedge x \in M \wedge P(x) \\ &\equiv \exists x: x \in M \wedge A \wedge P(x) \equiv \exists x \in M: A \wedge P(x). \end{aligned}$$

Nach Def. 1.1 (bounded) und Satz 1.8 (general-dl) gilt:

$$\begin{aligned} A \vee \forall x \in M: P(x) &\equiv A \vee \forall x: (x \in M \Rightarrow P(x)) \equiv A \vee \forall x: x \notin M \vee P(x) \\ &\equiv \forall x: A \vee x \notin M \vee P(x) \equiv \forall x: (x \in M \Rightarrow A \vee P(x)) \\ &\equiv \forall x \in M: A \vee P(x). \quad \square \end{aligned}$$

**Satz 1.17.** Es gilt:

$$(\exists x \in A: \exists y \in B: P(x, y)) \iff (\exists y \in B: \exists x \in A: P(x, y)).$$

**Beweis.** Nach Def. 1.1 (bounded), Satz 1.8 (general-dl) und Satz 1.14 (exists-cl) gilt:

$$\begin{aligned} \exists x \in A: \exists y \in B: P(x, y) &\equiv \exists x: x \in A \wedge \exists y: y \in B \wedge P(x, y) \\ &\equiv \exists x: \exists y: x \in A \wedge y \in B \wedge P(x, y) \equiv \exists y: \exists x: y \in B \wedge x \in A \wedge P(x, y) \\ &\equiv \exists y: y \in B \wedge \exists x: x \in A \wedge P(x, y) \equiv \exists y \in B: \exists x \in A: P(x, y). \quad \square \end{aligned}$$

**Satz 1.18.** Es gilt:

$$(\forall x \in A: \forall y \in B: P(x, y)) \iff (\forall y \in B: \forall x \in A: P(x, y)).$$

**Beweis (LEM, boolesche Algebra).**

Nach Def. 1.1 (bounded), Satz 1.8 (general-dl) und Satz 1.15 (all-cl) gilt:

$$\begin{aligned} \forall x \in A: \forall y \in B: P(x, y) &\equiv \forall x: x \in A \Rightarrow \forall y: y \in B \Rightarrow P(x, y) \\ &\equiv \forall x: x \notin A \vee \forall y: y \notin B \vee P(x, y) \equiv \forall x: \forall y: x \notin A \vee y \notin B \vee P(x, y) \\ &\equiv \forall y: \forall x: y \notin B \vee x \notin A \vee P(x, y) \equiv \forall y: y \notin B \vee \forall x: x \notin A \vee P(x, y) \\ &\equiv \forall y: y \in B \Rightarrow \forall x: x \in A \Rightarrow P(x, y) \equiv \forall y \in B: \forall x \in A: P(x, y). \quad \square \end{aligned}$$

**Satz 1.19.** Für eine Aussage  $P$ , die nicht von  $x$  abhängt, und ein nichtleeres Diskursuniversum gilt:

$$(\exists x: P) \iff P.$$

**Beweis.** Nach 1.8 (general-dl) gilt:

$$\exists x: P \equiv \exists x: 1 \wedge P \equiv (\exists x: 1) \wedge P \equiv 1 \wedge P \equiv P.$$

Im vorletzten Schritt wurde dabei ausgenutzt, dass für ein nichtleeres Diskursuniversum immer  $(\exists x: 1) \equiv 1$  gelten muss.  $\square$

**Satz 1.20.** Es gilt

$$(\exists x \in M: P) \iff (M \neq \emptyset) \wedge P.$$

**Beweis.** Nach Def. 1.1 (bounded) und Satz 1.8 (general-dl) gilt:

$$\exists x \in M: P \equiv \exists x: x \in M \wedge P \equiv (\exists x: x \in M) \wedge P \equiv (M \neq \emptyset) \wedge P. \quad \square$$

## 1.3 Mengenlehre

### 1.3.1 Definitionen

**Definition 1.2 (seteq: Gleichheit von Mengen).**

$$A = B :\iff \forall x: (x \in A \iff x \in B).$$

**Definition 1.3 (subseq: Teilmenge).**

$$A \subseteq B :\iff \forall x: (x \in A \implies x \in B).$$

**Definition 1.4 (filter: beschreibende Angabe).**

$$a \in \{x \mid P(x)\} :\iff P(a).$$

**Definition 1.5 (cap: Schnitt).**

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

**Definition 1.6 (cup: Vereinigung).**

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

**Definition 1.7 (intersection: Schnitt).**

$$\bigcap_{i \in I} A_i := \{x \mid \forall i \in I: x \in A_i\} = \{x \mid \forall i: (i \in I \implies x \in A_i)\}.$$

**Definition 1.8 (union: Vereinigung).**

$$\bigcup_{i \in I} A_i := \{x \mid \exists i \in I: x \in A_i\} = \{x \mid \exists i: (i \in I \wedge x \in A_i)\}.$$

**Definition 1.9 (cart: kartesisches Produkt).**

$$A \times B := \{(a, b) \mid a \in A \wedge b \in B\} = \{t \mid \exists a: \exists b: (t = (a, b) \wedge a \in A \wedge b \in B)\}.$$

### 1.3.2 Rechenregeln

**Satz 1.21 (Kommutativgesetze).** Es gilt  $A \cap B = B \cap A$  und  $A \cup B = B \cup A$ .

**Beweis.** Nach Def. 1.2 (seteq) expandieren:

$$\forall x: (x \in A \cap B \iff x \in B \cap A).$$

Nach Def. 1.5 (cap) und Def. 1.4 (filter) gilt:

$$x \in A \cap B \iff x \in A \wedge x \in B \iff x \in B \wedge x \in A \iff x \in B \cap A.$$

Für die Vereinigung ist das analog.  $\square$

**Satz 1.22 (Assoziativgesetze).** Es gilt  $A \cap (B \cap C) = (A \cap B) \cap C$  und  $A \cup (B \cup C) = (A \cup B) \cup C$ .

**Beweis.** Nach Def. 1.2 (seteq) expandieren:

$$\forall x: [x \in A \cap (B \cap C) \iff x \in (A \cap B) \cap C].$$

Nach Def. 1.5 (cap) und Def. 1.4 (filter) gilt:

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

Für die Vereinigung ist das analog.  $\square$

**Satz 1.23.** Es gilt  $A \cap B \subseteq A$ .

**Beweis.** Expansion liefert die Formel  $x \in A \wedge x \in B \implies x \in A$ . Gemäß boolescher Algebra gilt allgemein

$$\varphi \wedge \psi \implies \varphi \equiv \neg(\varphi \wedge \psi) \vee \varphi \equiv \neg\varphi \vee \neg\psi \vee \varphi \equiv 1 \vee \neg\psi \equiv 1.$$

Setze  $\varphi := (x \in A)$  und  $\psi := (x \in B)$ .  $\square$

**Satz 1.24.** Es gilt  $A \subseteq B \iff A \cap B = A$ .

**Beweis.** Aufgrund von Satz 1.23 muss lediglich  $A \subseteq B \iff A \subseteq A \cap B$  gezeigt werden. Expansion führt zur Formel

$$x \in A \implies x \in B \iff x \in A \implies x \in A \wedge x \in B.$$

Die Formel  $\varphi \implies \psi \iff \varphi \implies \varphi \wedge \psi$  ist aber tautologisch, denn

$$\varphi \implies \varphi \wedge \psi \equiv \neg\varphi \vee (\varphi \wedge \psi) \equiv (\neg\varphi \vee \varphi) \wedge (\neg\varphi \wedge \psi) \equiv 1 \wedge (\varphi \implies \psi) \equiv \varphi \implies \psi.$$

Setze  $\varphi := (x \in A)$  und  $\psi := (x \in B)$ .  $\square$

**Satz 1.25.** Es gilt  $a = b \iff \forall x: (x = a \iff x = b)$ .

**Beweis.** Die Implikation  $a = b \implies \forall x: (x = a \iff x = b)$ . Wenn wir  $a = b$  voraussetzen, kann  $b$  gegen  $a$  ersetzt werden und es ergibt sich

$$(\forall x: (x = a \iff x = a)) \iff (\forall x: 1) \iff 1.$$

Die andere Implikation bringen wir zunächst in ihre Kontraposition:

$$a \neq b \implies \exists x: ((x = a) \oplus (x = b)).$$

Auf einer leeren Grundmenge wird der Allquantifizierung über  $a, b$  immer genügt. Besitzt die Grundmenge nur ein Element, dann muss  $a = b$  sein, womit  $a \neq b$  falsch ist und die Implikation somit erfüllt. Wir setzen nun  $a \neq b$  voraus. Wählt man nun  $x = a$ , dann ist  $x \neq b$ , womit die Kontravalenz erfüllt wird.  $\square$

**Satz 1.26.** Es gilt  $a = b \iff \{a\} = \{b\}$ .

**Beweis.** Es gilt:

$$\{a\} = \{b\} \iff \{x \mid x = a\} = \{x \mid x = b\} \iff \forall x: (x = a \iff x = b).$$

Nach Satz 1.25 ist das aber äquivalent zu  $a = b$ .  $\square$

**Satz 1.27.** Es gilt:

$$\forall x: \forall y: (x = y \wedge P(x) \iff P(y))$$

**Satz 1.28.** Es gilt:

$$(\forall t \in A \times B: P(t)) \iff \forall a \in A: \forall b \in B: P(a, b).$$

**Beweis.** Nach Def. 1.9 (cart) gilt:

$$\begin{aligned} (\forall t \in A \times B: P(t)) &\iff (\forall t: t \in A \times B \implies P(t)) \\ &\iff (\forall t: (\exists a: \exists b: t = (a, b) \wedge a \in A \wedge b \in B) \implies P(t)). \end{aligned}$$

Unter doppelter Anwendung von Satz 1.13 gilt weiter:

$$\iff \forall t: \forall a: \forall b: [t = (a, b) \wedge a \in A \wedge b \in B \implies P(t)].$$

Substituiert man  $t := (a, b)$ , dann ergibt sich:

$$\implies (\forall a: \forall b: a \in A \wedge b \in B \implies P(a, b)) \iff \forall a \in A: \forall b \in B: P(a, b),$$

wobei  $P(a, b)$  eine Kurzschreibweise für  $P((a, b))$  ist. Von der Gegenrichtung bilden wir die Kontraposition:

$$(\exists t: \exists a: \exists b: t = (a, b) \wedge a \in A \wedge b \in B \wedge \overline{P(t)}) \implies \exists a: \exists b: a \in A \wedge b \in B \wedge \overline{P(a, b)}.$$

Dem  $\exists t$  wird aber immer durch  $t := (a, b)$  genügt, so dass sich die äquivalente Formel

$$(\exists a: \exists b: a \in A \wedge b \in B \wedge \overline{P(a, b)}) \implies \exists a: \exists b: a \in A \wedge b \in B \wedge \overline{P(a, b)}.$$

ergibt.  $\square$

**Satz 1.29.** Es gilt:

$$(\exists t \in A \times B: P(t)) \iff (\exists a \in A: \exists b \in B: P(a, b)).$$

**Beweis.** Nach Def. 1.9 (cart) gilt:

$$\begin{aligned} (\exists t \in A \times B: P(t)) &\iff (\exists t: P(t) \wedge t \in A \times B) \\ &\iff (\exists t: P(t) \wedge \exists a: \exists b: t = (a, b) \wedge a \in A \wedge b \in B) \\ &\iff (\exists t: \exists a: \exists b: P(t) \wedge a \in A \wedge b \in B \wedge t = (a, b)) \\ &\iff \exists a \in A: \exists b \in B: \exists t: P(t) \wedge t = (a, b). \end{aligned}$$

Nun gilt aber ganz offensichtlich

$$(\exists t: P(t) \wedge t = (a, b)) \iff P(a, b).$$

Nimmt man  $P(a, b)$  an, dann lässt sich  $\exists t: P(t) \wedge t = (a, b)$  durch Wahl von  $t := (a, b)$  bestätigen. Nimmt man umgekehrt  $\exists t: P(t) \wedge t = (a, b)$  an, lässt sich  $P(a, b)$  daraus unter Anwendung von Satz 1.27 ableiten. Da  $\exists t: P(t) \wedge t = (a, b)$  gegen  $P(a, b)$  ersetzt werden darf, folgt die Behauptung.  $\square$

**Satz 1.30.** Es gilt:

$$\bigcup_{t \in I \times J} A_t = \bigcup_{i \in I} \bigcup_{j \in J} A_{ij}. \quad (t = (i, j))$$

**Beweis.** Nach Def. 1.8 (union) und Satz 1.29 gilt:

$$\begin{aligned} x \in \bigcup_{t \in I \times J} A_t &\iff (\exists t \in I \times J : x \in A_t) \iff (\exists i \in I : \exists j \in J : x \in A_{ij}) \\ &\iff (\exists i \in I : x \in \bigcup_{j \in J} A_{ij}) \iff x \in \bigcup_{i \in I} \bigcup_{j \in J} A_{ij}. \end{aligned}$$

Nach Def. 1.2 (seteq) folgt die Behauptung.  $\square$

**Satz 1.31.** Es gilt:

$$\bigcup_{i \in I} \bigcup_{j \in J} A_{ij} = \bigcup_{j \in J} \bigcup_{i \in I} A_{ij}.$$

**Beweis.** Nach Def. 1.8 (union) und Satz 1.17 gilt:

$$\begin{aligned} x \in \bigcup_{i \in I} \bigcup_{j \in J} A_{ij} &\iff (\exists i \in I : x \in \bigcup_{j \in J} A_{ij}) \iff (\exists i \in I : \exists j \in J : x \in A_{ij}) \\ &\iff (\exists j \in J : \exists i \in I : x \in A_{ij}) \iff (\exists j \in J : x \in \bigcup_{i \in I} A_{ij}) \iff x \in \bigcup_{j \in J} \bigcup_{i \in I} A_{ij}. \end{aligned}$$

Nach Def. 1.2 (seteq) folgt die Behauptung.  $\square$

## 1.4 Relationen

**Definition 1.10 (rel: Relation).**

Zu zwei Mengen  $X, Y$  bezeichnet man jede Menge  $R \subseteq X \times Y$  als Relation.

**Definition 1.11 (img: Bildmenge).** Zu einer Relation  $R$  wird die Menge

$$R(M) := \{y \mid \exists x \in M : (x, y) \in R\}$$

als Bildmenge von  $M$  unter  $R$  bezeichnet.

**Korollar 1.32.** Sei  $R$  eine Relation und seien  $A, B$  beliebige Mengen.

Es gilt  $R(A \cup B) = R(A) \cup R(B)$ .

**Beweis.** Expansion mit Def. 1.11 (img) führt zur Behauptung

$$(\exists x \in A \cup B : (x, y) \in R) \iff (\exists x \in A : (x, y) \in R) \vee (\exists x \in B : (x, y) \in R).$$

Die linke Seite lässt sich gemäß Def. 1.1 (bounded), Def. 1.6 (cup) und Satz 1.9 (exists-dl) äquivalent in die rechte umformen.  $\square$

**Korollar 1.33.** Sei  $R$  eine Relation und seien  $A, B$  beliebige Mengen.

Es gilt  $R(A) \setminus R(B) \subseteq R(A \setminus B)$ .

**Beweis.** Expansion mit Def. 1.11 (img) führt zur Behauptung

$$(\exists x \in A : (x, y) \in R) \wedge (\forall x \in B : (x, y) \notin R) \implies \exists x \in A \setminus B : (x, y) \in R.$$

Laut der ersten Prämisse existiert ein  $x \in A$  mit  $(x, y) \in R$ . Die zweite Prämisse ist äquivalent zur Kontraposition  $(x, y) \in R \implies x \notin B$ . Infolge ist  $x \in A \setminus B$ . Somit bezeugt  $x$  die Existenzaussage auf der rechten Seite.  $\square$

## 1.5 Abbildungen

### 1.5.1 Definitionen

**Definition 1.12 (app: Applikation).** Für eine Abbildung  $f$  ist

$$y = f(x) :\iff (x, y) \in G_f.$$

**Definition 1.13 (img: Bildmenge).**

Für eine Abbildung  $f: X \rightarrow Y$  und  $A \subseteq X$  wird die Menge

$$f(A) := \{y \mid \exists x \in A: y = f(x)\} = \{y \mid \exists x: (x \in A \wedge y = f(x))\}$$

als Bildmenge von  $A$  unter  $f$  bezeichnet.

**Definition 1.14 (preimg: Urbildmenge).** Für eine Abbildung  $f: X \rightarrow Y$  wird

$$f^{-1}(B) := \{x \mid f(x) \in B\} = \{x \mid \exists y \in B: y = f(x)\}$$

als Urbildmenge von  $B$  unter  $f$  bezeichnet.

**Definition 1.15 (inj: Injektion).**

Eine Abbildung  $f: X \rightarrow Y$  heißt genau dann injektiv, wenn gilt:

$$\forall x_1: \forall x_2: (f(x_1) = f(x_2) \implies x_1 = x_2)$$

bzw. äquivalent (Kontraposition)

$$\forall x_1: \forall x_2: (x_1 \neq x_2 \implies f(x_1) \neq f(x_2)).$$

**Definition 1.16 (sur: Surjektion).**

Eine Abbildung  $f: X \rightarrow Y$  heißt genau dann surjektiv, wenn gilt:

$$Y \subseteq f(X).$$

**Definition 1.17 (composition: Verkettung).**

Für Abbildungen  $f: X \rightarrow Y$  und  $g: Y \rightarrow Z$  heißt

$$(g \circ f): X \rightarrow Z, \quad (g \circ f)(x) := g(f(x))$$

Verkettung von  $f$  und  $g$ , sprich » $g$  nach  $f$ «.

### 1.5.2 Grundlagen

**Satz 1.34 (feq: Gleichheit von Abbildungen).** Zwei Abbildungen  $f: X \rightarrow Y$  und  $g: X' \rightarrow Y'$  sind genau dann gleich, kurz  $f = g$ , wenn  $X = X'$  und  $Y = Y'$  und

$$\forall x: f(x) = g(x).$$

**Beweis.** Nach Definition gilt  $f = g$  genau dann, wenn  $(G_f, X, Y) = (G_g, X', Y')$ , was äquivalent zu  $G_f = G_g \wedge X = X' \wedge Y = Y'$  ist. Nach Def. 1.2 (seteq) gilt

$$G_f = G_g \iff \forall t: (t \in G_f \iff t \in G_g).$$

Nach Satz 1.25 und Def. 1.12 (app) gilt

$$\begin{aligned} (\forall x: f(x) = g(x)) &\iff (\forall x: \forall y: (y = f(x) \iff y = g(x))) \\ &\iff (\forall x: \forall y: ((x, y) \in G_f \iff (x, y) \in G_g)) \iff \forall t: (t \in G_f \iff t \in G_g). \end{aligned}$$

Da die Quantifizierung auf  $x \in X$ ,  $y \in Y$  und  $t \in X \times Y$  beschränkt ist, konnte im letzten Schritt Satz 1.28 angewendet werden.  $\square$

**Satz 1.35 (preimg-dl: Distributivität der Urbildoperation).**

Für  $f: X \rightarrow Y$  und beliebige Mengen  $M_i$  gilt:

$$f^{-1}(M_1 \cap M_2) = f^{-1}(M_1) \cap f^{-1}(M_2), \quad (1.11)$$

$$f^{-1}(M_1 \cup M_2) = f^{-1}(M_1) \cup f^{-1}(M_2), \quad (1.12)$$

$$f^{-1}\left(\bigcap_{i \in I} M_i\right) = \bigcap_{i \in I} f^{-1}(M_i), \quad (1.13)$$

$$f^{-1}\left(\bigcup_{i \in I} M_i\right) = \bigcup_{i \in I} f^{-1}(M_i). \quad (1.14)$$

**Beweis.** Nach Def. 1.2 (seteq) expandieren:

$$\forall x: [x \in f^{-1}(M_1 \cap M_2) \iff x \in f^{-1}(M_1) \cap f^{-1}(M_2)].$$

Nach Def. 1.14 (preimg) und Def. 1.5 (cap) zusammen mit Def. 1.4 (filter) gilt:

$$\begin{aligned} x \in f^{-1}(M_1 \cap M_2) &\iff f(x) \in M_1 \cap M_2 \iff f(x) \in M_1 \wedge f(x) \in M_2 \\ &\iff x \in f^{-1}(M_1) \wedge x \in f^{-1}(M_2) \iff x \in f^{-1}(M_1) \cap f^{-1}(M_2). \end{aligned}$$

Für die Vereinigung ist das analog.

Schnitt von beliebig vielen Mengen. Nach Def. 1.2 (seteq) expandieren:

$$\forall x: [x \in f^{-1}\left(\bigcap_{i \in I} M_i\right) \iff x \in \bigcap_{i \in I} f^{-1}(M_i)].$$

Nach Def. 1.14 (preimg) und Def. 1.7 (intersection) zusammen mit Def. 1.4 (filter) gilt:

$$\begin{aligned} x \in f^{-1}\left(\bigcap_{i \in I} M_i\right) &\iff f(x) \in \bigcap_{i \in I} M_i \iff \forall i(i \in I \implies f(x) \in M_i) \\ &\iff \forall i(i \in I \implies x \in f^{-1}(M_i)) \iff x \in \bigcap_{i \in I} f^{-1}(M_i). \end{aligned}$$

Für die Vereinigung ist das analog.  $\square$

**Satz 1.36 (img-cup-dl: Distributivität der Bildoperation über die Vereinigung).** Für  $f: X \rightarrow Y$  und Mengen  $M_i \subseteq X$  gilt:

$$f(M_1 \cup M_2) = f(M_1) \cup f(M_2), \quad (1.15)$$

$$f\left(\bigcup_{i \in I} M_i\right) = \bigcup_{i \in I} f(M_i). \quad (1.16)$$

**Beweis.** Nach Def. 1.2 (seteq) expandieren:

$$\forall y: (y \in f(M_1 \cup M_2) \iff y \in f(M_1) \cup f(M_2)).$$

Nach Def. 1.13 (img), Def. 1.6 (cup), Satz 1.3 (bool-dl) und Satz 1.9 (exists-dl) gilt:

$$\begin{aligned} y \in f(M_1 \cup M_2) &\iff (\exists x: x \in M_1 \cup M_2 \wedge y = f(x)) \\ &\iff (\exists x: (x \in M_1 \vee x \in M_2) \wedge y = f(x)) \\ &\iff (\exists x: x \in M_1 \wedge y = f(x) \vee x \in M_2 \wedge y = f(x)) \\ &\iff (\exists x: x \in M_1 \wedge y = f(x)) \vee (\exists x: x \in M_2 \wedge y = f(x)) \\ &\iff y \in f(M_1) \vee y \in f(M_2) \iff y \in f(M_1) \cup f(M_2). \end{aligned}$$



Nach Def. 1.2 (seteq) expandieren:

$$\forall y: [y \in f(\bigcup_{i \in I} M_i) \iff y \in \bigcup_{i \in I} f(M_i)].$$

Nach Def. 1.13 (img), Def. 1.8 (union), Satz 1.8 (general-dl) und Satz 1.14 (exists-cl) gilt:

$$\begin{aligned} y \in f(\bigcup_{i \in I} M_i) &\iff (\exists x: x \in \bigcup_{i \in I} M_i \wedge y = f(x)) \\ &\iff (\exists x: (\exists i: i \in I \wedge x \in M_i) \wedge y = f(x)) \iff (\exists x: \exists i: i \in I \wedge x \in M_i \wedge y = f(x)) \\ &\iff (\exists i \exists x: i \in I \wedge x \in M_i \wedge y = f(x)) \iff (\exists i: i \in I \wedge \exists x (x \in M_i \wedge y = f(x))) \\ &\iff (\exists i: i \in I \wedge y \in f(M_i)) \iff y \in \bigcup_{i \in I} f(M_i). \quad \square \end{aligned}$$

**Satz 1.37.** Es gilt:

$$f(M_1 \cap M_2) \subseteq f(M_1) \cap f(M_2), \quad (1.17)$$

$$f(\bigcap_{i \in I} M_i) \subseteq \bigcap_{i \in I} f(M_i). \quad (1.18)$$

**Beweis.** Nach Def. 1.3 (subsetq) expandieren:

$$\forall y: (y \in f(M_1 \cap M_2) \implies y \in f(M_1) \cap f(M_2)).$$

Nach Def. 1.13 (img), Def. 1.5 (cap) und Satz 1.10 (exists-asym-dl) gilt:

$$\begin{aligned} y \in f(M_1 \cap M_2) &\iff (\exists x: x \in M_1 \cap M_2 \wedge y = f(x)) \\ &\iff (\exists x: x \in M_1 \wedge x \in M_2 \wedge y = f(x)) \\ &\iff (\exists x: x \in M_1 \wedge y = f(x) \wedge x \in M_2 \wedge y = f(x)) \\ &\implies (\exists x: x \in M_1 \wedge y = f(x)) \wedge (\exists x: x \in M_2 \wedge y = f(x)) \\ &\iff y \in f(M_1) \wedge y \in f(M_2) \iff y \in f(M_1) \cap f(M_2). \end{aligned}$$

Nach Def. 1.3 (subsetq) expandieren:

$$\forall y: (y \in f(\bigcap_{i \in I} M_i) \implies y \in \bigcap_{i \in I} f(M_i))$$

Nach Def. 1.13 (img) und Def. 1.7 (intersection) gilt:

$$\begin{aligned} y \in f(\bigcap_{i \in I} M_i) &\iff (\exists x: x \in \bigcap_{i \in I} M_i \wedge y = f(x)) \\ &\iff (\exists x: (\forall i: i \in I \Rightarrow x \in M_i) \wedge y = f(x)) \\ &\iff (\exists x: \forall i: i \in I \Rightarrow x \in M_i \wedge y = f(x)) \\ &\implies (\forall i: \exists x: i \in I \Rightarrow x \in M_i \wedge y = f(x)) \\ &\iff (\forall i: i \in I \Rightarrow \exists x: x \in M_i \wedge y = f(x)) \\ &\iff (\forall i: i \in I \Rightarrow y \in f(M_i)) \iff y \in \bigcap_{i \in I} f(M_i). \quad \square \end{aligned}$$

**Korollar 1.38.** Zwei disjunkte Mengen haben disjunkte Urbilder.

**Beweis.** Sei  $A \cap B = \emptyset$ . Gemäß Satz 1.35 (preimg-dl) ist

$$f^{-1}(A) \cap f^{-1}(B) = f^{-1}(A \cap B) = f^{-1}(\emptyset) = \emptyset. \quad \square$$

**Satz 1.39.** Es gilt  $M \subseteq N \implies f^{-1}(M) \subseteq f^{-1}(N)$ .

**Beweis 1.** Gemäß Satz 1.24 ist  $M \subseteq N$  äquivalent zu  $M \cap N = M$ . Man wendet die Urbildoperation  $f^{-1}$  nun auf beide Seiten der Gleichung an und erhält mittels Satz 1.35 (preimg-dl) dann

$$f^{-1}(M \cap N) = f^{-1}(M) \cap f^{-1}(N) = f^{-1}(M).$$

Nochmalige Anwendung von Satz 1.24 liefert das gewünschte Resultat

$$f^{-1}(M) \subseteq f^{-1}(N). \quad \square$$

**Beweis 2.** Die Expansion der Aussage bringt

$$(y \in M \implies y \in N) \implies (f(x) \in M \implies f(x) \in N).$$

Trivialerweise kann die Prämisse mit  $y := f(x)$  spezialisiert werden werden.  $\square$

**Satz 1.40.** Es gilt  $M \subseteq N \implies f(M) \subseteq f(N)$ .

**Beweis.** Gemäß Satz 1.24 ist  $M \subseteq N$  äquivalent zu  $M \cap N = M$ . Man wendet die Bildoperation nun auf beide Seiten der Gleichung an und erhält mittels Satz 1.37 dann

$$f(M) = f(M \cap N) \subseteq f(M) \cap f(N).$$

Laut Satz 1.23 ist folglich  $f(M) = f(M) \cap f(N)$ . Nochmalige Anwendung von Satz 1.24 bringt das gewünschte Resultat  $f(M) \subseteq f(N)$ .  $\square$

**Satz 1.41.** Es gilt:

$$f(M) = \bigcup_{x \in M} \{f(x)\}.$$

**Beweis.** Nach Def. 1.13 (img) und Def. 1.8 (union) gilt:

$$y \in f(M) \iff (\exists x \in M: y = f(x)) \iff (\exists x \in M: y \in \{f(x)\}) \iff y \in \bigcup_{x \in M} \{f(x)\}.$$

Nach Def. 1.2 (seteq) folgt dann die Behauptung.  $\square$

**Satz 1.42.** Es gilt  $(g \circ f)^{-1}(M) = f^{-1}(g^{-1}(M))$ .

**Beweis.** Nach Def. 1.14 (preimg) und Def. 1.2 (seteq) expandieren und Def. 1.4 (filter) anwenden:

$$(g \circ f)(x) \in M \iff f(x) \in \{y \mid g(y) \in M\}.$$

Links Def. 1.17 (composition) anwenden und rechts nochmals Def. 1.4 (filter):

$$g(f(x)) \in M \iff g(f(x)) \in M. \quad \square$$

**Satz 1.43.** Es gilt  $(g \circ f)(M) = g(f(M))$ .

**Beweis.** Nach Def. 1.13 (img) und Def. 1.2 expandieren, dann 1.4 (filter) anwenden:

$$(\exists x: x \in M \wedge z = (g \circ f)(x)) \iff (\exists y: y \in f(M) \wedge z = g(y)).$$

Die rechte Seite mit Def. 1.13 (img) expandieren und Def. 1.4 (filter) anwenden. Unter Anwendung von Satz 1.8 (general-dl) und Satz 1.14 (exists-cl) ergibt sich

$$\begin{aligned} & (\exists y: (\exists x: x \in M \wedge y = f(x)) \wedge z = g(y)) \\ & \iff (\exists y: \exists x: x \in M \wedge y = f(x) \wedge z = g(y)) \\ & \iff (\exists x: x \in M \wedge \exists y: y = f(x) \wedge z = g(y)) \\ & \iff (\exists x: x \in M \wedge z = g(f(x))) \\ & \iff (\exists x: x \in M \wedge z = (g \circ f)(x)). \quad \square \end{aligned}$$

**Satz 1.44.** Sei  $f: A \rightarrow B$  eine Abbildung und  $A \neq \emptyset$ . Man nennt eine Funktion  $g: B \rightarrow A$  mit  $g \circ f = \text{id}_A$  Linksinverse zu  $f$ . Die Abbildung  $f$  ist genau dann injektiv, wenn eine Linksinverse zu  $f$  existiert.

**Beweis.** Sei  $f$  injektiv. Man wähle ein  $a \in A$ , das wegen  $A \neq \emptyset$  existieren muss. Man definiert nun  $g: B \rightarrow A$  mit

$$g(y) := \begin{cases} x \text{ wobei } y = f(x), & \text{wenn } y \in f(A), \\ a & \text{wenn } y \notin f(A). \end{cases}$$

Diese Funktion ist eindeutig definiert, weil  $f$  injektiv ist. Gemäß ihrer Definition gilt  $g(f(x)) = x$ , bzw.  $g \circ f = \text{id}$ .

Sei nun eine Linksinverse  $g$  mit  $g \circ f = \text{id}$  gegeben. Dann gilt

$$f(a) = f(b) \implies g(f(a)) = g(f(b))$$

und

$$g(f(a)) = g(f(b)) \iff (g \circ f)(a) = (g \circ f)(b) \iff \text{id}(a) = \text{id}(b) \iff a = b.$$

Es ergibt sich

$$f(a) = f(b) \implies a = b. \quad \square$$

**Korollar 1.45.** Es gilt  $f^{-1}(A^c) = f^{-1}(A)^c$  bzw.  $f(x) \in A^c \iff \neg f(x) \in A$ .

**Beweis.** Zufolge der Expansion von Def. 1.14 (preimg) ist

$$(\exists y \in A^c : y = f(x)) \iff \neg(\exists y \in A : y = f(x)) \iff (\forall y \in A : y \neq f(x))$$

zu zeigen. Weil  $x$  ein Element des Definitionsbereichs ist, muss der Funktionswert  $f(x)$  in irgendeiner Menge liegen. Die Implikation von rechts nach links. Weil  $f(x)$  nicht in  $A$  liegt, muss  $f(x)$  in  $A^c$  liegen. Die Implikation von links nach rechts. Gemäß Prämisse liegt  $f(x)$  in  $A^c$ . Weil  $x$  nur einen Funktionswert besitzt, kann  $f(x)$  nicht in  $A$  liegen.  $\square$

**Satz 1.46.** Für jede Abbildung  $f$  gilt  $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$ .

**Beweis 1.** Ergibt sich sofort gemäß Definition:

$$\begin{aligned} f^{-1}(A) \setminus f^{-1}(B) &= \{x \mid x \in f^{-1}(A) \wedge \neg x \in f^{-1}(B)\} \\ &= \{x \mid f(x) \in A \wedge f(x) \notin B\} = \{x \mid f(x) \in A \setminus B\} = f^{-1}(A \setminus B). \quad \square \end{aligned}$$

**Beweis 2.** Gemäß Korollar 1.45 und Satz 1.35 (preimg-dl) ist

$$\begin{aligned} f^{-1}(A) \setminus f^{-1}(B) &= \{x \mid x \in f^{-1}(A) \wedge \neg x \in f^{-1}(B)\} \\ &= \{x \mid x \in f^{-1}(A) \wedge x \in f^{-1}(B^c)\} = f^{-1}(A) \cap f^{-1}(B^c) \\ &= f^{-1}(A \cap B^c) = f^{-1}(A \setminus B). \quad \square \end{aligned}$$

**Satz 1.47.** Für jede Abbildung  $f$  gilt  $f(f^{-1}(N)) \subseteq N$ .

**Beweis.** Gemäß Definition bekommt man

$$y \in f(f^{-1}(N)) \iff (\exists x : x \in f^{-1}(N) \wedge y = f(x)) \iff (\exists x : f(x) \in N \wedge y = f(x)).$$

Leicht ersichtlich ist nun, dass

$$(\exists x : f(x) \in N \wedge y = f(x)) \implies y \in N. \quad \square$$

**Satz 1.48.** Für jede Abbildung  $f: A \rightarrow B$  gilt  $f(f^{-1}(N)) = N$ , sofern  $N \subseteq f(A)$  ist.

**Beweis.** Laut Satz 1.47 bleibt zu zeigen

$$y \in N \implies (\exists x \in A: f(x) \in N \wedge y = f(x)).$$

Setzt man nun  $N \subseteq f(A)$  voraus, dann ist  $f(x) \in N$  allgemeingültig. Man bekommt

$$(\exists x \in A: f(x) \in N \wedge y = f(x)) \iff (\exists x \in A: y = f(x)) \iff y \in f(A).$$

Die Implikation  $y \in N \implies y \in f(A)$  ist nun wiederum definitionsgemäß äquivalent zu  $N \subseteq f(A)$ , was Voraussetzung war.  $\square$

**Satz 1.49.** Für jede Abbildung  $f: A \rightarrow B$  gilt  $\exists M(f(M) = N) \iff N \subseteq f(A)$ .

**Beweis.** Hat man ein  $M$  mit  $f(M) = N$ , dann ist trivialerweise  $f(M) \subseteq f(A)$ , also  $N \subseteq f(A)$ . Liegt umgekehrt eine Menge  $N \subseteq f(A)$  vor, dann kann man  $M := f^{-1}(N)$  setzen, nach Satz 1.48 gilt dann  $f(M) = N$ .  $\square$

**Satz 1.50.** Ist  $f$  injektiv, dann gilt  $f(A \setminus B) = f(A) \setminus f(B)$ .

**Beweis.** Da  $f$  injektiv ist, gibt es nach Satz 1.44 eine Linksinverse  $f^{-1}$ . Nach Satz 1.43 ist für eine beliebige Menge  $M$  die Gleichung

$$f^{-1}(f(M)) = (f^{-1} \circ f)(M) = \text{id}(M) = M$$

erfüllt. Unter Heranziehung von Satz 1.46 bekommt man

$$f^{-1}(f(A) \setminus f(B)) = f^{-1}(f(A)) \setminus f^{-1}(f(B)) = \text{id}(A) \setminus \text{id}(B) = A \setminus B.$$

Wendet man nun auf beide Seiten der Gleichung  $f$  an, dann ergibt sich nach Satz 1.48 das gesuchte Resultat  $f(A) \setminus f(B) = f(A \setminus B)$ .  $\square$

**Satz 1.51.** Ist  $f$  eine bijektive Abbildung und  $f^{-1}$  die Umkehrabbildung von  $f$ , dann stimmt das Urbild  $f^{-1}(N)$  mit der Bildmenge von  $N$  unter der Umkehrabbildung – zur Unterscheidung  $(f^{-1})(N)$  geschrieben – überein.

**Beweis.** Expansion der Gleichung  $f^{-1}(N) = (f^{-1})(N)$  führt zur Bedingung

$$f(x) \in N \iff (\exists y: y \in N \wedge x = f^{-1}(y)).$$

Da  $f$  bijektiv ist, gilt  $x = f^{-1}(y) \iff f(x) = f(f^{-1}(y)) = y$ . Demnach ist

$$(\exists y: y \in N \wedge x = f^{-1}(y)) \iff (\exists y: f(x) \in N) \iff f(x) \in N.$$

Die Bedingung ist daher immer erfüllt.  $\square$

Es genügt nicht, wenn  $f$  injektiv ist. Als Gegenbeispiel setze

$$f: \{0\} \rightarrow \{0, 1\}, \quad f(x) := x.$$

Hier ist  $f^{-1}(\{1\}) = \emptyset$ . Jedoch ist  $(f^{-1})(\{1\}) = \{0\}$ .

**Satz 1.52 (Rechtskürzbarkeit von Surjektionen).**

Ist  $f: X \rightarrow Y$  eine surjektive Abbildung, dann gilt

$$g \circ f = h \circ f \implies g = h.$$

**Beweis.** Laut Prämisse und Satz 1.34 (feq) ist  $g(f(x)) = h(f(x))$  für jedes  $x \in X$ . Da  $f$  surjektiv ist, lässt sich zu jedem  $y \in Y$  ein  $x \in X$  finden, so dass  $y = f(x)$ . Demnach ist  $g(y) = h(y)$  für alle  $y \in Y$ , denn man kann immer mindestens ein  $x$  finden, so dass sich  $y := f(x)$  substituieren lässt. Laut Satz 1.34 (feq) ist daher  $g = h$ .  $\square$

### 1.5.3 Kardinalzahlen

**Axiom 1.53 (acc: abzählbares Auswahlaxiom).** Sei  $(A_n)_{n \in \mathbb{N}}$  eine Folge nichtleerer Mengen. Dann existiert eine Funktion  $f: \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} A_n$  mit  $f(n) \in A_n$ .

**Definition 1.18 (equipotent: Gleichmächtigkeit).** Zwei Mengen  $A, B$  heißen genau dann gleichmächtig, wenn eine Bijektion  $f: A \rightarrow B$  existiert.

**Satz 1.54.** Sei  $M$  eine beliebige Menge. Die Potenzmenge  $2^M$  ist zur Menge  $\{0, 1\}^M$  gleichmächtig.

**Beweis.** Für eine Aussage  $A$  sei

$$[A] := \begin{cases} 1 & \text{wenn } A \text{ gilt,} \\ 0 & \text{sonst.} \end{cases}$$

Für  $A \subseteq M$  betrachte man nun die Indikatorfunktion

$$\chi_A: M \rightarrow \{0, 1\}, \quad \chi_A(x) := [x \in A].$$

Die Abbildung

$$\varphi: 2^M \rightarrow \{0, 1\}^M, \quad \varphi(A) := \chi_A$$

ist eine kanonische Bijektion.

**Zur Injektivität.** Nach Def. 1.15 (inj) muss gelten:

$$\varphi(A) = \varphi(B) \implies A = B, \quad \text{d. h.} \quad \chi_A = \chi_B \implies A = B.$$

Nach Satz 1.34 (feq) und Def. 1.2 (seteq) wird die Aussage expandiert zu:

$$(\forall x: \chi_A(x) = \chi_B(x)) \implies (\forall x: x \in A \iff x \in B).$$

Es gilt aber nun:

$$\chi_A(x) = \chi_B(x) \iff [x \in A] = [x \in B] \iff (x \in A \iff x \in B).$$

**Zur Surjektivität.** Wir müssen nach Def. 1.16 (sur) prüfen, dass  $\{0, 1\}^M \subseteq \varphi(2^M)$  gilt. Expansion nach Def. 1.3 (subseq) und Def. 1.13 (img) ergibt:

$$\forall f: (f \in \{0, 1\}^M \implies \exists A \in 2^M: f = \varphi(A)).$$

Um dem Existenzquantor zu genügen, wähle

$$A := f^{-1}(\{1\}) = \{x \in M \mid f(x) \in \{1\}\} = \{x \in M \mid f(x) = 1\}.$$

Es gilt  $f = \chi_A$ , denn

$$\chi_A(x) = [x \in A] = [x \in \{x \mid f(x) = 1\}] = [f(x) = 1] = f(x).$$

Da  $\varphi$  eine Bijektion ist, müssen  $2^M$  und  $\{0, 1\}^M$  nach Def. 1.18 (equipotent) gleichmächtig sein.  $\square$

**Satz 1.55.** Man setze Axiom 1.53 (acc) voraus. Die Vereinigung von abzählbar vielen abzählbar unendlichen Mengen ist abzählbar unendlich. Kurz  $|\bigcup_{n \in \mathbb{N}} A_n| = |\mathbb{N}|$ , wenn  $|A_n| = |\mathbb{N}|$  für jedes  $n$ .

**Beweis.** Sei  $B_n$  die Menge der Bijektionen aus  $\text{Abb}(\mathbb{N}, A_n)$ . Nach Axiom 1.53 (acc) kann aus jeder Menge  $B_n$  eine Bijektion  $f_n: \mathbb{N} \rightarrow A_n$  ausgewählt werden. Man betrachte nun

$$\varphi: \mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} A_n, \quad \varphi(n, m) := f_n(m).$$

Die Abbildung  $\varphi$  ist surjektiv, denn nach Satz 1.41 und Satz 1.30 gilt

$$\begin{aligned} \varphi(\mathbb{N} \times \mathbb{N}) &= \bigcup_{(n,m) \in \mathbb{N} \times \mathbb{N}} \{f_n(m)\} = \bigcup_{n \in \mathbb{N}} \bigcup_{m \in \mathbb{N}} \{f_n(m)\} \\ &= \bigcup_{n \in \mathbb{N}} f_n\left(\bigcup_{m \in \mathbb{N}} \{m\}\right) = \bigcup_{n \in \mathbb{N}} f_n(\mathbb{N}) = \bigcup_{n \in \mathbb{N}} A_n. \end{aligned}$$

Daher gilt  $|\bigcup_{n \in \mathbb{N}} A_n| \leq |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$ . Für eine beliebige der Bijektionen  $f_n \in B_n$  lässt sich die Zielmenge erweitern, so dass man eine Injektion  $f: \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} A_n$  erhält. Daher ist auch  $|\mathbb{N}| \leq |\bigcup_{n \in \mathbb{N}} A_n|$ . Nach dem Satz von Cantor-Bernstein gilt also  $|\bigcup_{n \in \mathbb{N}} A_n| = |\mathbb{N}|$ .  $\square$

**Satz 1.56.** Wenn  $R$  abzählbar ist, dann ist auch der Polynomring  $R[X]$  abzählbar.

**Beweis.** Zu jedem Polynom vom Grad  $n \geq 1$  gehört auf kanonische Weise genau ein Tupel aus  $M_n := R^{n-1} \times R \setminus \{0\}$ . Da  $R$  abzählbar ist, sind auch  $R^{n-1}$  und  $R \setminus \{0\}$  abzählbar. Dann ist auch  $M_n$  abzählbar. Nach Satz 1.55 gilt

$$|R[X]| = 1 + \left| \bigcup_{n \in \mathbb{N}} M_n \right| = 1 + |\mathbb{N}| = |\mathbb{N}|. \quad \square$$

**Satz 1.57.** Es gibt nur abzählbar unendlich viele algebraische Zahlen.

**Beweis 1.** Zu zeigen ist  $|\mathbb{A}| = |\mathbb{N}|$  mit

$$\mathbb{A} := \{a \in \mathbb{C} \mid \exists p(p \in \mathbb{Q}[X] \setminus \{0\} \wedge p(a) = 0)\}.$$

Dass  $\mathbb{A}$  unendlich ist, ist leicht ersichtlich, denn schon jede rationale Zahl  $q$ , von denen es unendlich viele gibt, ist Nullstelle von  $p(X) := X - q$  und daher algebraisch.

Ein Polynom vom Grad  $n$  kann höchstens  $n$  Nullstellen besitzen. Nach Satz 1.56 gilt  $|\mathbb{Q}[X]| = |\mathbb{N}|$ . Für  $\mathbb{Q}[X]$  lässt sich also eine Abzählung angeben. Bei dieser Abzählung lässt sich für jedes Polynom  $p$  die Liste der Nullstellen von  $p$  einfügen. Streicht man alle Nullstellen, die schon einmal vorkamen, dann erhält man eine Abzählung der algebraischen Zahlen. Demnach gilt  $|\mathbb{A}| = |\mathbb{N}|$ .  $\square$

**Beweis 2.** Jedem  $p = \sum_{k=0}^n a_k X^k$  lässt sich eine Höhe  $h := n + \sum_{k=0}^n |a_k|$  zuordnen. Zu einer festen Höhe  $h$  kann es nur endlich viele Polynome  $p \in \mathbb{Z}[X]$  geben, wodurch man eine Abzählung der Polynome erhält, wenn für  $h = 1, h = 2, h = 3$  usw. jeweils die Liste der Polynome eingefügt wird. Für jedes Polynom  $p$  lässt sich die Liste der Nullstellen von  $p$  einfügen. Streicht man alle Nullstellen, die schon einmal vorkamen, dann erhält man eine Abzählung der algebraischen Zahlen.  $\square$

**Beweis 3.** Für  $n \in \mathbb{N}$  sei

$$A_n := \{x \in \mathbb{A} \mid x \text{ ist Nullstelle eines } p \in \mathbb{Z}[X] \setminus \{0\} \text{ mit } \deg(p) = n, \\ \text{dessen Koeffizienten } a_k \text{ alle } |a_k| \leq n \text{ erfüllen}\}.$$

Alle  $A_n$  sind endlich und es gilt  $\mathbb{A} = \bigcup_{n \in \mathbb{N}} A_n$ . Daher muss  $|\mathbb{A}| \leq |\mathbb{N}|$  sein.  $\square$

**Definition 1.19 (Satz und Def. Multiplikation von Kardinalzahlen).**

Die Operation  $|X| \cdot |Y| := |X \times Y|$  ist wohldefiniert.

**Beweis.** Zu zeigen ist, dass  $|X \times Y| = |X' \times Y'|$  aus  $|X| = |X'|$  und  $|Y| = |Y'|$  folgt. Nach Voraussetzung gibt es Bijektionen  $f_1: X \rightarrow X'$  und  $f_2: Y \rightarrow Y'$ . Gesucht ist mindestens eine Bijektion  $f: X \times Y \rightarrow X' \times Y'$ . Diese erhält man gemäß folgender Konstruktion:

$$f(x, y) := (f_1(x), f_2(y)).$$

Die Abbildung  $f$  ist injektiv, denn

$$\begin{aligned} f(x_1, y_1) = f(x_2, y_2) &\iff (f_1(x_1), f_2(y_1)) = (f_1(x_2), f_2(y_2)) \\ &\iff f_1(x_1) = f_1(x_2) \wedge f_2(y_1) = f_2(y_2) \iff x_1 = x_2 \wedge y_1 = y_2 \\ &\iff (x_1, y_1) = (x_2, y_2). \end{aligned}$$

Für die Surjektivität muss es für jedes  $(x', y')$  mindestens ein  $(x, y)$  mit  $(x', y') = f(x, y)$  geben. Die Konstruktion ergibt

$$(x', y') = (f_1(x), f_2(y)) \iff x' = f_1(x) \wedge y' = f_2(y).$$

Man findet  $x = f_1^{-1}(x')$  und  $y = f_2^{-1}(y')$ .

Die Umkehrabbildung ist gegeben gemäß

$$\begin{aligned} f^{-1}(x', y') &= f^{-1}((x', y')) := ((f_1^{-1} \circ \pi_1)(x', y'), (f_2^{-1} \circ \pi_2)(x', y')) \\ &= (f_1^{-1}(x'), f_2^{-1}(y')). \end{aligned}$$

Mit  $\pi_k$  ist die Projektion auf die  $k$ -te Komponente gemeint.  $\square$

**Definition 1.20 (Satz und Def. Addition von Kardinalzahlen).**

Für  $X \cap Y = \emptyset$  ist  $|X| + |Y| := |X \cup Y|$  wohldefiniert. Das schließt den Spezialfall  $|X| + |Y| := |X \sqcup Y|$  mit  $X \sqcup Y := (\{0\} \times X) \cup (\{1\} \times Y)$  ein.

**Beweis.** Zu zeigen ist, dass  $|X \cup Y| = |X' \cup Y'|$  aus  $|X| = |X'|$  und  $|Y| = |Y'|$  folgt. Nach Voraussetzung gibt es Bijektionen  $f_1: X \rightarrow X'$  und  $f_2: Y \rightarrow Y'$ , wobei  $X \cap Y = \emptyset$  und  $X' \cap Y' = \emptyset$  gilt. Gesucht ist mindestens eine Bijektion  $f: X \cup Y \rightarrow X' \cup Y'$ . Diese erhält man gemäß folgender Konstruktion:

$$f(x) := \begin{cases} f_1(x) & \text{für } x \in X, \\ f_2(x) & \text{für } x \in Y. \end{cases}$$

Die Abbildung  $f$  ist injektiv, denn entweder ist  $x' \in X'$  und somit

$$x' = f(a) = f(b) \iff x' = f_1(a) = f_1(b) \iff a = b$$

oder  $x' \in Y'$  und somit

$$x' = f(a) = f(b) \iff x' = f_2(a) = f_2(b) \iff a = b.$$

Zusammengefasst folgt  $f(a) = f(b) \iff a = b$  für alle  $a, b \in X \cup Y$ .

Für die Surjektivität muss es für jedes  $x'$  mindestens ein  $x$  mit  $x' = f(x)$  geben. Entweder ist  $x' \in X'$ , dann ist  $x' = f_1(x)$  und daher  $x = f_1^{-1}(x')$ . Oder es ist  $x' \in Y'$ , dann ist  $x' = f_2(x)$  und daher  $x = f_2^{-1}(x')$ .  $\square$

**Definition 1.21 (Satz und Def. Potenz von Kardinalzahlen).**

Die Operation  $|Y|^{|X|} := |Y^X|$  ist wohldefiniert.

**Beweis.** Zu zeigen ist, dass  $|\text{Abb}(X, Y)| = |\text{Abb}(X', Y')|$  aus  $|X| = |X'|$  und  $|Y| = |Y'|$  folgt. Nach Voraussetzung gibt es Bijektionen  $f_1: X \rightarrow X'$  und  $f_2: Y \rightarrow Y'$ . Gesucht ist eine Bijektion  $F: \text{Abb}(X, Y) \rightarrow \text{Abb}(X', Y')$ . Diese erhält man gemäß folgender Konstruktion:

$$F(f) := f_2 \circ f \circ f_1^{-1}.$$

Die Abbildung  $F$  ist injektiv, da

$$F(f) = F(g) \iff f_2 \circ f \circ f_1^{-1} = f_2 \circ g \circ f_1^{-1} \iff f_2 \circ f = f_2 \circ g \iff f = g,$$

denn Bijektionen sind kürzbar. Für die Surjektivität muss es für jedes  $f'$  mindestens ein  $f$  mit  $f' = F(f)$  geben. Das führt auf die Gleichung  $f' = f_2 \circ f \circ f_1^{-1}$ . Diese lässt sich umformen zu  $f_2^{-1} \circ f' = f \circ f_1^{-1}$ . Wendet man beide Seiten auf  $f_1$  an, ergibt sich  $f = f_2^{-1} \circ f' \circ f_1$ .  $\square$

**Definition 1.22 (Weniger mächtig).**

Eine Menge  $A$  heißt weniger mächtig als eine Menge  $B$ , kurz  $|A| < |B|$ , wenn es eine Injektion  $A \rightarrow B$  gibt, aber keine Bijektion  $A \rightarrow B$ .

**Satz 1.58 (Satz von Cantor).**

Eine Menge ist stets weniger mächtig als ihre Potenzmenge. Kurz  $|A| < |2^A|$ .

**Beweis.** Eine Injektion  $A \rightarrow 2^A$  können wir mit  $x \mapsto \{x\}$  trivial angeben. Nun wird die Annahme, es gäbe eine Surjektion  $f: A \rightarrow 2^A$  zum Widerspruch geführt, womit es erst recht keine Bijektion gibt. Dazu definiert man die Menge

$$D := \{x \in A \mid x \notin f(x)\}.$$

Weil  $D \subseteq A$  ist, hat man  $D \in 2^A$ . Weil  $f$  surjektiv sein soll, muss es ein  $x \in A$  geben, so dass  $f(x) = D$ . Gemäß Def. 1.2 (seteq) und Definition von  $D$  muss also

$$\exists x \in A: \forall u: (u \in f(x) \iff u \in A \wedge u \notin f(u))$$

gelten. Die allquantifizierte Aussage ist allerdings für  $u := x$  nicht erfüllt, denn unter Berücksichtigung  $x \in A$  gelangt man zur widersprüchlichen Aussage

$$x \in f(x) \iff x \notin f(x). \quad \square$$

**Satz 1.59.** Die Menge der endlichen Teilmengen der natürlichen Zahlen ist abzählbar.

**Beweis.** Zu jeder Teilmenge  $A \subseteq \mathbb{N}_0$  gehört genau eine Indikatorfunktion

$$1_A: \mathbb{N}_0 \rightarrow \{0, 1\}, \quad 1_A(n) := [n \in A].$$

Weil die Indikatorfunktion die natürlichen Zahlen als Definitionsbereich besitzt, handelt es sich um eine Folge, die wie jede Folge als formale Potenzreihe.

$$p_A(X) = \sum_{n=0}^{\infty} 1_A(n) X^n = \sum_{n \in A} X^n$$

dargestellt werden kann. Ist die Teilmenge  $A$  eine endliche, besitzt die Indikatorfunktion eine endliche Nichtnullstellenmenge, womit sich  $p_A(X)$  zu einem Polynom reduziert.



Wegen  $1_A(n) < 2$ , kann man  $1_A$  als Kodierung der Zahl  $p_A(2)$  auffassen, nämlich als Binärdarstellung der Zahl  $p_A(2)$  im Stellenwertsystem zur Basis  $X = 2$ . Deshalb fungiert die Abbildung

$$f(A) := p_A(2) = \sum_{n \in A} 2^n$$

als kanonische Bijektion zwischen der Menge der endlichen Teilmengen der natürlichen Zahlen und der natürlichen Zahlen. Sie ist injektiv aufgrund der Eigenheiten von Stellenwertsystemen. Sie ist surjektiv, weil  $1_A$  laut Prämisse jede beliebige Binärdarstellung sein darf, und man somit aufgrund der Eigenheiten von Stellenwertsystemen jede beliebige natürliche Zahl erhält.  $\square$



# 2 Analysis

## 2.1 Folgen

### 2.1.1 Konvergenz

**Definition 2.1 (open-ep-ball: offene Epsilon-Umgebung).** Sei  $(M, d)$  ein metrischer Raum. Unter der offenen Epsilon-Umgebung von  $a \in M$  versteht man:

$$U_\varepsilon(a) := \{x \mid d(x, a) < \varepsilon\}.$$

Man setze zunächst speziell  $d(x, a) := |x - a|$  bzw.  $d(x, a) := \|x - a\|$ .

**Definition 2.2 (lim: konvergente Folge, Grenzwert).** Eine Folge  $(a_n)$  heißt konvergent gegen einen Grenzwert  $a$ , wenn es zu jedem noch so kleinen  $\varepsilon$  einen Index  $n_0$  gibt, so dass ab diesem Index sämtliche ihrer Werte in der Umgebung  $U_\varepsilon(a)$  liegen. Formal:

$$\lim_{n \rightarrow \infty} a_n = a \iff \forall \varepsilon > 0: \exists n_0: \forall n \geq n_0: a_n \in U_\varepsilon(a)$$

bzw.

$$\lim_{n \rightarrow \infty} a_n = a \iff \forall \varepsilon > 0: \exists n_0: \forall n \geq n_0: \|a_n - a\| < \varepsilon.$$

**Definition 2.3 (bseq: beschränkte Folge).** Eine Folge  $(a_n)$  mit  $a_n \in \mathbb{R}$  heißt genau dann beschränkt, wenn es eine reelle Zahl  $S$  gibt mit  $|a_n| < S$  für alle  $n$ .

Eine Folge  $(a_n)$  von Punkten eines normierten Raums heißt genau dann beschränkt, wenn es eine reelle Zahl  $S$  gibt mit  $\|a_n\| < S$  für alle  $n$ .

**Satz 2.1 (Grenzwert bei Konvergenz eindeutig bestimmt).**

Eine konvergente Folge von Elementen eines metrischen Raumes besitzt genau einen Grenzwert.

**Beweis.** Sei  $(a_n)$  eine konvergente Folge mit  $a_n \rightarrow g_1$ . Sei weiterhin  $g_1 \neq g_2$ . Es wird nun gezeigt, dass  $g_2$  kein Grenzwert von  $a_n$  sein kann. Wir müssen also zeigen:

$$\neg \lim_{n \rightarrow \infty} a_n = g_2 \iff \exists \varepsilon > 0: \forall n_0: \exists n \geq n_0: a_n \notin U_\varepsilon(g_2)$$

mit  $a_n \notin U_\varepsilon(g_2) \iff d(a_n, g_2) \geq \varepsilon$ .

Um dem Existenzquantor zu genügen, wählt man nun  $\varepsilon = \frac{1}{2}d(g_1, g_2)$ . Nach Def. 3.8 (metric-space) gilt  $d(g_1, g_2) > 0$ , daher ist auch  $\varepsilon > 0$ . Nach Satz 3.5 sind die Umgebungen  $U_\varepsilon(g_1)$  und  $U_\varepsilon(g_2)$  disjunkt. Wegen  $a_n \rightarrow g_1$  gibt es ein  $n_0$  mit  $a_n \in U_\varepsilon(g_1)$  für alle  $n \geq n_0$ . Dann gibt es für jedes beliebig große  $n_0$  aber auch  $n \geq n_0$  mit  $a_n \notin U_\varepsilon(g_2)$ .  $\square$

**Satz 2.2 (lim-scaled-ep: skaliertes Epsilon).** Es gilt:

$$\lim_{n \rightarrow \infty} a_n = a \iff \forall \varepsilon > 0: \exists n_0: \forall n \geq n_0: \|a_n - a\| < R\varepsilon,$$

wobei  $R > 0$  ein fester aber beliebiger Skalierungsfaktor ist.

**Beweis.** Betrachte  $\varepsilon > 0$  und multipliziere auf beiden Seiten mit  $R$ . Dabei handelt es sich um eine Äquivalenzumformung. Setze  $\varepsilon' := R\varepsilon$ . Demnach gilt:

$$\varepsilon > 0 \iff \varepsilon' > 0.$$

Nach der Ersetzungsregel dürfen wir die Teilformel  $\varepsilon > 0$  nun ersetzen. Es ergibt sich die äquivalente Formel

$$\lim_{n \rightarrow \infty} a_n = a \iff \forall \varepsilon' > 0: \exists n_0: \forall n \geq n_0: \|a_n - a\| < \varepsilon'.$$

Das ist aber genau Def. 2.2 (lim).  $\square$

**Satz 2.3.** Es gilt:

$$\lim_{n \rightarrow \infty} a_n = a \implies \lim_{n \rightarrow \infty} \|a_n\| = \|a\|.$$

**Beweis.** Nach Satz 3.7 (umgekehrte Dreiecksungleichung) gilt:

$$|\|a_n\| - \|a\|| \leq \|a_n - a\| < \varepsilon.$$

Dann ist aber erst recht  $|\|a_n\| - \|a\|| < \varepsilon$ .  $\square$

**Satz 2.4.** Ist  $(a_n)$  eine Nullfolge und  $(b_n)$  eine beschränkte Folge, dann ist auch  $(a_n b_n)$  eine Nullfolge.

**Beweis.** Wenn  $(b_n)$  beschränkt ist, dann existiert nach Def. 2.3 (bseq) eine Schranke  $S$  mit  $|b_n| < S$  für alle  $n$ . Man multipliziert nun auf beiden Seiten mit  $|a_n|$  und erhält

$$|a_n b_n| = |a_n| |b_n| < |a_n| S.$$

Wenn  $a_n \rightarrow 0$ , dann muss für jedes  $\varepsilon$  ein  $n_0$  existieren mit  $|a_n| < \varepsilon$  für  $n \geq n_0$ . Multipliziert man auf beiden Seiten mit  $S$ , und ergibt sich

$$|a_n b_n - 0| = |a_n b_n| < |a_n| S < S\varepsilon.$$

Nach Satz 2.2 (lim-scaled-ep) gilt dann aber  $a_n b_n \rightarrow 0$ .  $\square$

**Satz 2.5.** Sind  $(a_n)$  und  $(b_n)$  Nullfolgen, dann ist auch  $(a_n b_n)$  eine Nullfolge.

**Beweis 1.** Wenn  $(b_n)$  eine Nullfolge ist, dann ist  $(b_n)$  auch beschränkt. Nach Satz 2.4 gilt dann die Behauptung.

**Beweis 2.** Sei  $\varepsilon > 0$  beliebig. Es gibt ein  $n_0$ , so dass  $|a_n| < \varepsilon$  und  $|b_n| < \varepsilon$  für  $n \geq n_0$ . Demnach ist

$$|a_n b_n| = |a_n| |b_n| < |a_n| \varepsilon < \varepsilon^2.$$

Wegen  $\varepsilon > 0 \iff \varepsilon' > 0$  mit  $\varepsilon' = \varepsilon^2$  gilt

$$\forall \varepsilon' > 0: \exists n_0: \forall n \geq n_0: |a_n b_n| < \varepsilon'.$$

Nach Def. 2.2 (lim) gilt somit die Behauptung.  $\square$

**Satz 2.6 (Grenzwertsatz zur Addition).** Seien  $(a_n), (b_n)$  Folgen von Vektoren eines normierten Raumes. Es gilt:

$$\lim_{n \rightarrow \infty} a_n = a \wedge \lim_{n \rightarrow \infty} b_n = b \implies \lim_{n \rightarrow \infty} a_n + b_n = a + b.$$

**Beweis.** Dann gibt es ein  $n_0$ , so dass für  $n \geq n_0$  sowohl  $\|a_n - a\| < \varepsilon$  als auch  $\|b_n - b\| < \varepsilon$ . Addition der beiden Ungleichungen führt zu

$$\|a_n - a\| + \|b_n - b\| < 2\varepsilon.$$

Laut der Dreiecksungleichung, das ist Axiom (N3) in Def. 3.10 (normed-space), gilt nun aber die Abschätzung

$$\|(a_n + b_n) - (a + b)\| = \|(a_n - a) + (b_n - b)\| \leq \|a_n - a\| + \|b_n - b\|.$$

Somit gilt erst recht

$$\|(a_n + b_n) - (a + b)\| < 2\varepsilon.$$

Nach Satz 2.2 (lim-scaled-ep) folgt die Behauptung.  $\square$

**Satz 2.7 (Grenzwertsatz zur Skalarmultiplikation).** Sei  $(a_n)$  eine Folge von Vektoren eines normierten Raumes und sei  $r \in \mathbb{R}$  oder  $r \in \mathbb{C}$ . Es gilt:

$$\lim_{n \rightarrow \infty} a_n = a \implies \lim_{n \rightarrow \infty} r a_n = r a.$$

**Beweis.** Sei  $\varepsilon > 0$  fest aber beliebig. Es gibt nun ein  $n_0$ , so dass  $\|a_n - a\| < \varepsilon$  für  $n \geq n_0$ . Multipliziert man auf beiden Seiten mit  $|r|$  und zieht Def. 3.10 (normed-space) Axiom (N2) heran, dann ergibt sich

$$\|r a_n - r a\| = |r| \|a_n - a\| < |r| \varepsilon.$$

Nach Satz 2.2 (lim-scaled-ep) folgt die Behauptung.  $\square$

**Satz 2.8 (Grenzwertsatz zum Produkt).**

Seien  $(a_n)$  und  $(b_n)$  Folgen reeller Zahlen. Es gilt:

$$\lim_{n \rightarrow \infty} a_n = a \wedge \lim_{n \rightarrow \infty} b_n = b \implies \lim_{n \rightarrow \infty} a_n b_n = ab.$$

**Beweis.** Nach Voraussetzung sind  $a_n - a$  und  $b_n - b$  Nullfolgen. Da das Produkt von Nullfolgen wieder eine Nullfolge ist, gilt

$$(a_n - a)(b_n - b) = a_n b_n - a_n b - a b_n + ab \rightarrow 0.$$

Da nach Satz 2.7 aber  $a_n b \rightarrow ab$  und  $a b_n \rightarrow ab$ , ergibt sich nach Satz 2.6 nun

$$(a_n - a)(b_n - b) + a_n b + a b_n = a_n b_n + ab \rightarrow 2ab.$$

Addiert man nun noch die konstante Folge  $-2ab$  und wendet nochmals Satz 2.6 an, dann ergibt sich die Behauptung

$$a_n b_n \rightarrow ab. \quad \square$$

**Satz 2.9.** Sei  $M$  ein metrischer Raum und  $X$  ein topologischer Raum. Eine Abbildung  $f: M \rightarrow X$  ist genau dann stetig, wenn sie folgenstetig ist.

**Satz 2.10 (Satz zur Fixpunktgleichung).** Sei  $M$  ein metrischer Raum und sei  $f: M \rightarrow M$ . Sei  $x_{n+1} := f(x_n)$  eine Fixpunktiteration. Wenn die Folge  $(x_n)$  zu einem Startwert  $x_0$  konvergiert mit  $x_n \rightarrow x$ , und wenn  $f$  eine stetige Abbildung ist, dann muss der Grenzwert  $x$  die Fixpunktgleichung  $x = f(x)$  erfüllen.

**Beweis.** Wenn  $x_n \rightarrow x$ , dann gilt trivialerweise auch  $x_{n+1} \rightarrow x$ . Weil  $f$  stetig ist, ist  $f$  nach Satz 2.9 auch folgenstetig. Daher gilt  $\lim f(a_n) = f(\lim a_n)$  für jede konvergente Folge  $(a_n)$ . Somit gilt:

$$x = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} f(x_n) = f(\lim_{n \rightarrow \infty} x_n) = f(x). \quad \square$$

### 2.1.2 Wachstum und Landau-Symbole

**Definition 2.4.** Seien  $f, g: D \rightarrow \mathbb{R}$  mit  $D = \mathbb{N}$  oder  $D = \mathbb{R}$ . Man sagt, die Funktion  $f$  wächst nicht wesentlich schneller als  $g$ , kurz  $f \in \mathcal{O}(g)$ , genau dann, wenn

$$\exists c > 0: \exists x_0 > 0: \forall x > x_0: |f(x)| \leq c|g(x)|.$$

**Korollar 2.11.** Ist  $r \in \mathbb{R}$  mit  $r \neq 0$  eine Konstante, dann gilt  $\mathcal{O}(rg) = \mathcal{O}(g)$ .

**Beweis.** Nach Def. 2.4 ist

$$f \in \mathcal{O}(rg) \iff \exists c > 0: \exists x_0 > 0: \forall x > x_0: |f(x)| \leq c|rg(x)|.$$

Man hat nun

$$|f(x)| \leq c|rg(x)| = c \cdot |r| \cdot |g(x)|.$$

Wegen  $r \neq 0$  ist  $|r| > 0$  und daher auch  $c > 0 \iff c|r| > 0$ . Sei  $c' := r|c|$ . Also gilt  $c > 0 \iff c' > 0$ . Nach der Ersetzungsregel darf  $c > 0$  gegen  $c' > 0$  ersetzt werden und man erhält die äquivalente Bedingung

$$\exists c' > 0: \exists x_0 > 0: \forall x > x_0: |f(x)| \leq c'|g(x)|.$$

Nach Def. 2.4 ist das gerade  $f \in \mathcal{O}(g)$ .  $\square$

## 2.2 Stetige Funktionen

**Definition 2.5 (Grenzwert einer Funktion).** Sei  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}$  und sei  $p$  ein Häufungspunkt von  $D$ . Die Funktion  $f$  heißt konvergent gegen  $L$  für  $x \rightarrow p$ , wenn

$$\forall \varepsilon > 0: \exists \delta > 0: \forall x \in D: (0 < |x - p| < \delta \implies |f(x) - L| < \varepsilon).$$

Bei Konvergenz schreibt man  $L = \lim_{x \rightarrow p} f(x)$  und nennt  $L$  den Grenzwert.

**Definition 2.6 (cont: stetig).** Eine Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}$  heißt stetig an der Stelle  $x_0 \in D$ , wenn

$$\forall \varepsilon > 0: \exists \delta > 0: \forall x \in D: (|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon).$$

**Definition 2.7 (Lipschitz-stetig).**

Eine Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}$  heißt Lipschitz-stetig, wenn eine Konstante  $L$  existiert, so dass

$$|f(b) - f(a)| \leq L|b - a|$$

für alle  $a, b \in D$ .

**Definition 2.8 (Lipschitz-stetig an einer Stelle).**

Eine Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}$  heißt Lipschitz-stetig an der Stelle  $x_0 \in D$ , wenn eine Konstante  $L$  existiert, so dass

$$|f(x_0) - f(a)| \leq L|x_0 - a|$$

für alle  $a \in D$ .

**Korollar 2.12.** Eine Funktion ist genau dann Lipschitz-stetig, wenn sie an jeder Stelle Lipschitz-stetig ist und die Menge der optimalen Lipschitz-Konstanten dabei beschränkt.

**Beweis.** Eine Lipschitz-stetige Funktion ist trivialerweise an jeder Stelle Lipschitz-stetig. Ist  $f: D \rightarrow \mathbb{R}$  an der Stelle  $b$  Lipschitz-stetig, dann existiert eine Lipschitz-Konstante  $L_b$  mit

$$\forall a \in D: |f(b) - f(a)| \leq L_b|b - a|.$$

Nach Voraussetzung ist  $L = \sup_{b \in D} L_b$  endlich. Alle  $L_b$  können nun zu  $L$  abgeschwächt werden und es ergibt sich

$$\forall b \in D: \forall a \in D: |f(b) - f(a)| \leq L|b - a|. \quad \square$$

**Definition 2.9 (lokal Lipschitz-stetig).**

Eine Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}$  heißt lokal Lipschitz-stetig in der Nähe einer Stelle  $x_0 \in D$ , wenn es eine Epsilon-Umgebung  $U_\varepsilon(x_0)$  gibt, so dass die Einschränkung von  $f$  auf diese Umgebung Lipschitz-stetig ist. Die Funktion heißt lokal Lipschitz-stetig, wenn sie in der Nähe jeder Stelle Lipschitz-stetig ist.

**Satz 2.13.** Ist die Funktion  $f: D \rightarrow \mathbb{R}$  an der Stelle  $x_0$  differenzierbar, dann gibt es ein  $\delta > 0$ , so dass die Einschränkung von  $f$  auf  $U_\delta(x_0)$  an der Stelle  $x_0$  Lipschitz-stetig ist.

**Beweis.** Def. 2.5 wird in Def. 2.10 (diff) eingesetzt. Es ergibt sich:

$$0 < |x - x_0| < \delta \implies \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon.$$

Nach der umgekehrten Dreiecksungleichung 3.7 gilt

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| \leq \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon.$$

Daraus ergibt sich

$$|f(x) - f(x_0)| < (|f'(x_0)| + \varepsilon) \cdot |x - x_0|$$

und somit erst recht

$$|f(x) - f(x_0)| \leq (|f'(x_0)| + \varepsilon) \cdot |x - x_0|,$$

wobei jetzt auch  $x = x_0$  erlaubt ist. Demnach wird Def. 2.8 erfüllt:

$$\exists \delta > 0: \forall x \in U_\delta(x_0): |f(x) - f(x_0)| \leq (|f'(x_0)| + \varepsilon) \cdot |x - x_0|. \quad \square$$

**Satz 2.14.** Eine differenzierbare Funktion ist genau dann Lipschitz-stetig, wenn ihre Ableitung beschränkt ist.

**Beweis.** Wenn  $f: I \rightarrow \mathbb{R}$  Lipschitz-stetig ist, dann gibt es  $L$  mit

$$\left| \frac{f(b) - f(a)}{b - a} \right| \leq L$$

für alle  $a, b \in D$  mit  $a \neq b$ . Daraus folgt

$$|f'(a)| = \left| \lim_{b \rightarrow a} \frac{f(b) - f(a)}{b - a} \right| = \lim_{b \rightarrow a} \left| \frac{f(b) - f(a)}{b - a} \right| \leq L.$$

Demnach ist die Ableitung beschränkt.

Sei nun umgekehrt die Ableitung beschränkt. Für  $a, b \in I$  mit  $a \neq b$  gibt es nach dem Mittelwertsatz ein  $x_0 \in (a, b)$ , so dass

$$|f'(x_0)| = \left| \frac{f(b) - f(a)}{b - a} \right|.$$

Da die Ableitung beschränkt ist gibt es ein Supremum  $L = \sup_{x \in I} |f'(x)|$ . Demnach ist  $|f'(x)| \leq L$  für alle  $x$ . Es ergibt sich

$$\left| \frac{f(b) - f(a)}{b - a} \right| \leq L |b - a| \implies |f(b) - f(a)| \leq L |b - a|.$$

Nun darf auch  $a = b$  gewählt werden.  $\square$

**Satz 2.15.** Eine auf einem kompakten Intervall  $[a, b]$  definierte stetig differenzierbare Funktion ist Lipschitz-stetig.

**Beweis.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar. Dann ist  $f'(x)$  stetig. Nach dem Satz vom Minimum und Maximum ist  $|f'(x)|$  beschränkt. Nach Satz 2.14 muss  $f$  Lipschitz-stetig sein.  $\square$

**Korollar 2.16.** Eine stetig differenzierbare Funktion ist lokal Lipschitz-stetig.

**Beweis.** Sei  $f: D \rightarrow \mathbb{R}$  stetig differenzierbar. Sei  $[a, b] \in D$ . Sei  $x_0 \in [a, b]$ . Die Einschränkung von  $f$  auf  $[a, b]$  ist Lipschitz-stetig nach Satz 2.15. Dann ist auch die Einschränkung von  $f$  auf  $U_\varepsilon(x_0) \subseteq [a, b]$  Lipschitz-stetig.  $\square$

**Satz 2.17.** Es gibt differenzierbare Funktionen, die nicht überall lokal Lipschitz-stetig sind.

**Beweis.** Aus Satz 2.14 ergibt sich also Kontraposition, dass eine Funktion mit unbeschränkter Ableitung nicht Lipschitz-stetig sein kann.

Ist  $f: D \rightarrow \mathbb{R}$  an jeder Stelle differenzierbar und ist  $f'$  in jeder noch so kleinen Umgebung der Stelle  $x_0$  unbeschränkt, dann kann  $f$  also in der Nähe dieser Stelle auch nicht lokal Lipschitz-stetig sein.

Ein Beispiel für eine solche Funktion ist  $f: [0, \infty) \rightarrow \mathbb{R}$  mit

$$f(0) := 0 \quad \text{und} \quad f(x) := x^{3/2} \cos\left(\frac{1}{x}\right).$$

Einerseits gilt

$$f'(0) = \lim_{h \rightarrow 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \rightarrow 0} \frac{f(h)}{h} = \lim_{h \rightarrow 0} (h^{1/2} \cos(\frac{1}{h})) = 0.$$



Die Funktion ist also an der Stelle  $x = 0$  differenzierbar. Andererseits gilt nach den Ableitungsregeln

$$f'(x) = \frac{3}{2} \sqrt{x} \cos\left(\frac{1}{x}\right) + \frac{1}{\sqrt{x}} \sin\left(\frac{1}{x}\right).$$

für  $x > 0$ . Der Term  $\frac{1}{\sqrt{x}}$  erwirkt für  $x \rightarrow 0$  immer größere Maxima von  $|f'(x)|$ . Daher kann  $f$  in der Nähe von  $x = 0$  nicht lokal Lipschitz-stetig sein.  $\square$

**Satz 2.18.** Sei  $f: \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar und  $f(x)$  konvergent für  $x \rightarrow \infty$ . Ist außerdem  $f'$  Lipschitz-stetig, zieht dies  $f'(x) \rightarrow 0$  für  $x \rightarrow \infty$  nach sich.

**Beweis.** Gemäß dem Cauchyschen Konvergenzkriterium gibt es zu jedem  $\varepsilon > 0$  eine Stelle  $x_0$ , so dass

$$|f(b) - f(a)| < \varepsilon \quad (2.1)$$

für alle  $a, b$  mit  $x_0 < a \leq b$ . Nun ist  $f'$  aufgrund der Lipschitz-Stetigkeit erst recht stetig, womit

$$\left| \int_a^b f'(x) dx \right| = |f(b) - f(a)| \quad (2.2)$$

laut dem Fundamentalsatz gilt. Gezeigt wird nun, dass  $|f'(a)|$  beschränkt ist. Sei dazu  $L$  die Lipschitz-Konstante. Ohne Beschränkung der Allgemeinheit sei  $f'(a) > 0$ . Fallen darf  $f'$  maximal mit dem Anstieg  $-L$ . Geschieht dies linear bis zur Nullstelle  $b$ , ergibt sich ein rechtwinkliges Dreieck mit dem Flächeninhalt

$$\frac{1}{2L} f'(a)^2 = \int_a^b f'(x) dx < \varepsilon. \quad (2.3)$$

Demnach ist  $f'(a) < \sqrt{2L\varepsilon}$ . Weil dies für alle  $a > x_0$  gilt, muss  $f'$  jede Beschränkung unterbieten, womit der Beweis der Behauptung erbracht ist.  $\square$

Die Diskussion Gegenbeispiels  $f(0) := 0$ ,  $f(x) := \sin(x^2)/x$  macht ersichtlich, dass die Aussage ohne Lipschitz-Stetigkeit nicht einmal für glatte Funktionen gilt.

## 2.3 Differentialrechnung

### 2.3.1 Ableitungsregeln

**Definition 2.10 (diff: differenzierbar, Ableitung).** Eine Funktion  $f: D \rightarrow \mathbb{R}$  heißt differenzierbar an der Stelle  $x_0 \in D$ , wenn der Grenzwert

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

existiert. Man nennt  $f'(x_0)$  die Ableitung von  $f$  an der Stelle  $x_0$ .

**Satz 2.19.** Sei  $I$  ein Intervall und  $f, g: I \rightarrow \mathbb{R}$ . Sind  $f, g$  differenzierbar an der Stelle  $x \in I$ , dann ist auch

$$f + g \text{ dort differenzierbar mit } (f + g)'(x) = f'(x) + g'(x), \quad (2.4)$$

$$f - g \text{ dort differenzierbar mit } (f - g)'(x) = f'(x) - g'(x), \quad (2.5)$$

$$fg \text{ dort differenzierbar mit } (fg)'(x) = f'(x)g(x) + f(x)g'(x). \quad (2.6)$$

**Beweis.** Es gilt

$$(f + g)'(x) = \lim_{h \rightarrow 0} \frac{(f + g)(x + h) - (f + g)(x)}{h} \quad (2.7)$$

$$= \lim_{h \rightarrow 0} \frac{(f(x + h) + g(x + h)) - (f(x) + g(x))}{h} \quad (2.8)$$

$$= \lim_{h \rightarrow 0} \left( \frac{f(x + h) - f(x)}{h} + \frac{g(x + h) - g(x)}{h} \right) \quad (2.9)$$

$$= \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} + \lim_{h \rightarrow 0} \frac{g(x + h) - g(x)}{h} = f'(x) + g'(x). \quad (2.10)$$

Da die Grenzwerte auf der rechten Seite nach Voraussetzung existieren, muss auch der Grenzwert der Summe existieren. Die Rechnung für die Subtraktion ist analog.

Bei der Multiplikation wird ein Nullsummentrick angewendet:

$$g(x)f'(x) + f(x)g'(x) = g(x) \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} + f(x) \lim_{h \rightarrow 0} \frac{g(x + h) - g(x)}{h} \quad (2.11)$$

$$= \lim_{h \rightarrow 0} \left[ g(x + h) \frac{f(x + h) - f(x)}{h} \right] + \lim_{h \rightarrow 0} \left[ f(x) \frac{g(x + h) - g(x)}{h} \right] \quad (2.12)$$

$$= \lim_{h \rightarrow 0} \frac{f(x + h)g(x + h) - f(x)g(x + h)}{h} + \lim_{h \rightarrow 0} \frac{f(x)g(x + h) - f(x)g(x)}{h} \quad (2.13)$$

$$= \lim_{h \rightarrow 0} \frac{f(x + h)g(x + h) - f(x)g(x + h) + f(x)g(x + h) - f(x)g(x)}{h} \quad (2.14)$$

$$= \lim_{h \rightarrow 0} \frac{f(x + h)g(x + h) - f(x)g(x)}{h} = \lim_{h \rightarrow 0} \frac{(fg)(x + h) - (fg)(x)}{h} = (fg)'(x). \quad (2.15)$$

Hierbei wurde  $\lim_{h \rightarrow 0} g(x + h) = g(x)$  benutzt, was richtig ist, weil  $g$  an der Stelle  $x$  differenzierbar ist und dort somit ganz sicher stetig.  $\square$

**Satz 2.20.** Sei  $I$  ein Intervall. Sind  $f, g: I \rightarrow \mathbb{R}$  an der Stelle  $x$  differenzierbar und ist  $g(x) \neq 0$ , dann ist auch  $f/g$  differenzierbar und es gilt

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \quad (2.16)$$

**Beweis.** Nach der Produktregel (2.6) gilt

$$0 = 1' = \left(g \cdot \frac{1}{g}\right)' = g' \cdot \frac{1}{g} + g \cdot \left(\frac{1}{g}\right)'. \quad (2.17)$$

Umstellen bringt  $(1/g)'(x) = -g'(x)/g(x)^2$ . Nochmalige Anwendung der Produktregel (2.6) bringt

$$\left(\frac{f}{g}\right)'(x) = \left(f \cdot \frac{1}{g}\right)'(x) = f'(x) \cdot \frac{1}{g(x)} + f(x) \left(\frac{1}{g}\right)'(x) \quad (2.18)$$

$$= \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{g(x)^2} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \quad \square \quad (2.19)$$

**Satz 2.21.** Für  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) := x^n$  mit  $n \in \mathbb{N}$  gilt  $f'(x) = nx^{n-1}$ .

**Beweis 1.** Heranziehung des binomischen Lehrsatzes bringt

$$f'(x) = \lim_{h \rightarrow 0} \frac{(x+h)^n - x^n}{h} = \lim_{h \rightarrow 0} \frac{\sum_{k=0}^n \binom{n}{k} x^{n-k} h^k - x^n}{h} \quad (2.20)$$

$$= \lim_{h \rightarrow 0} \left( nx^{n-1} + \sum_{k=2}^n \binom{n}{k} x^{n-k} h^{k-1} \right) = nx^{n-1}. \quad \square \quad (2.21)$$

**Beweis 2.** Induktiv. Der Induktionsanfang  $\frac{d}{dx}x = 1$  ist klar. Induktionsschritt mittels Produktregel (2.6):

$$\frac{d}{dx}x^n = \frac{d}{dx}(x \cdot x^{n-1}) = x^{n-1} + x \frac{d}{dx}x^{n-1} = x^{n-1} + (n-1)x^{n-1} = nx^{n-1}. \quad \square \quad (2.22)$$

**Satz 2.22.** Für  $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ ,  $f(x) := x^n$  mit  $n \in \mathbb{Z}$  gilt  $f'(x) = nx^{n-1}$ .

**Beweis.** Der Fall  $n = 0$  ist trivial und  $n \geq 1$  wurde schon in Satz 2.21 gezeigt. Sei nun  $a \in \mathbb{N}$  und  $n = -a$ . Nach der Produktregel (2.6) und Satz 2.21 gilt

$$0 = \frac{d}{dx}1 = \frac{d}{dx}(x^a x^{-a}) = x^{-a} \frac{d}{dx}x^a + x^a \frac{d}{dx}x^{-a} = x^{-a} a x^{a-1} + x^a \frac{d}{dx}x^{-a}. \quad (2.23)$$

Dividiert man nun durch  $x^a$  und formt um, dann ergibt sich

$$\frac{d}{dx}x^{-a} = -a x^{-a-1} \implies \frac{d}{dx}x^n = nx^{n-1}. \quad \square \quad (2.24)$$

### 2.3.2 Glatte Funktionen

**Satz 2.23.** Sei  $f: \mathbb{R} \rightarrow \mathbb{R}$  eine Funktion mit der Eigenschaft  $f(x) = 0$  für  $x \leq 0$  und  $f(x) > 0$  für  $x > 0$ . Es gibt glatte Funktionen mit dieser Eigenschaft, jedoch keine analytischen.

**Beweis.** Wegen  $f(x) = 0$  für  $x \leq 0$  muss die linksseitige  $n$ -te Ableitung an der Stelle  $x = 0$  immer verschwinden. Wenn die  $n$ -te Ableitung stetig sein soll, muss auch die rechtsseitige Ableitung bei  $x = 0$  verschwinden. Da die Funktion glatt sein soll, muss das für jede Ableitung gelten. Daher verschwindet die Taylorreihe an der Stelle  $x = 0$ . Da aber  $f(x) > 0$  für  $x > 0$ , gibt es keine noch so kleine Umgebung mit Übereinstimmung von  $f$  und ihrer Taylorreihe. Daher kann  $f$  an der Stelle  $x = 0$  nicht analytisch sein.

Eine glatte Funktion lässt sich jedoch konstruieren:

$$f(x) := \begin{cases} e^{-1/x} & \text{wenn } x > 0, \\ 0 & \text{wenn } x \leq 0. \end{cases}$$

Ist nämlich  $g(x)$  an einer Stelle glatt, dann ist es nach Kettenregel, Produktregel und Summenregel auch  $e^{g(x)}$ . Die  $n$ -te Ableitung lässt sich immer in der Form

$$\sum_k e^{g(x)} r_k(x) = e^{g(x)} \sum_k r_k(x) = e^{g(x)} r(x)$$

darstellen, wobei die  $r_k(x)$  bzw.  $r(x)$  in diesem Fall rationale Funktionen mit Polstelle bei  $x = 0$  sind. Da aber  $e^{-1/x}$  für  $x \rightarrow 0$  schneller fällt als jede rationale Funktion steigen kann, muss die rechtsseitige Ableitung an der Stelle  $x = 0$  immer verschwinden.  $\square$

### 2.3.3 Richtungsableitung

**Definition 2.11 (Richtungsableitung).** Sei  $U \subseteq \mathbb{R}^n$  offen,  $x \in U$  eine Stelle und  $v \in \mathbb{R}^n$  ein Vektor. Man betrachte für ein kleines  $\varepsilon > 0$  die Parametergerade

$$\gamma: (-\varepsilon, \varepsilon) \rightarrow U, \quad \gamma(t) := x + tv.$$

Für eine Funktion  $f: U \rightarrow \mathbb{R}$  ist die Zahl

$$D_v f(x) := (f \circ \gamma)'(0) = \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h},$$

falls sie existiert, die Richtungsableitung von  $f$  an der Stelle  $x$  in Richtung  $v$ .

**Korollar 2.24.** Die Funktionen  $f, g$  seien an der Stelle  $x$  in Richtung  $v$  differenzierbar. Sei  $c$  eine reelle Zahl. Dann sind auch  $f + g$ ,  $f - g$ ,  $cf$ ,  $fg$  differenzierbar und es gelten die den üblichen Ableitungsregeln analogen Regeln

$$D_v(f + g)(x) = D_v f(x) + D_v g(x),$$

$$D_v(f - g)(x) = D_v f(x) - D_v g(x),$$

$$D_v(cf)(x) = c D_v f(x),$$

$$D_v(fg)(x) = g(x) D_v f(x) + f(x) D_v g(x).$$

**Beweis.** Die Ableitungsregeln werden über die Definition auf die Ableitungsregeln für gewöhnliche reelle Funktionen zurückgeführt. So ist

$$\begin{aligned} D_v(f + g)(x) &= ((f + g) \circ \gamma)'(0) = ((f \circ \gamma) + (g \circ \gamma))'(0) \\ &= (f \circ \gamma)'(0) + (g \circ \gamma)'(0) = D_v f(x) + D_v g(x). \end{aligned}$$

Der Beweis der restlichen Regeln ist analog.  $\square$

**Korollar 2.25 (Kettenregel).**

Sei  $g: \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar und  $f$  differenzierbar an der Stelle  $x$  in Richtung  $v$ . Dann ist auch  $g \circ f$  entsprechend differenzierbar, und es gilt

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

**Beweis.** Die Regel ist gemäß der Definition auf die gewöhnliche Kettenregel zurückführbar. Man bekommt

$$D_v(g \circ f)(x) = (g \circ f \circ \gamma)'(0) = g'(f(\gamma(0))) \cdot (f \circ \gamma)'(0) = g'(f(x)) \cdot D_v f(x). \quad \square$$

**Definition 2.12 (Partielle Ableitung).**

Sei  $(\mathbf{e}_1, \dots, \mathbf{e}_n)$  die Standardbasis. Die partielle Ableitung  $\partial_k f(x)$  ist definiert als die Richtungsableitung  $D_v f(x)$  bezüglich  $v = \mathbf{e}_k$ .

**Korollar 2.26.** Zur jeder gewöhnlichen Ableitungsregel besitzt die Richtungsableitung eine analoge Regel.

**Vorbereitung.** Sei  $f = (f_1, \dots, f_n)$  ein Tupel von Funktionen aus einem Funktionenraum und sei entsprechend  $f(x) := (f_1(x), \dots, f_n(x))$ . Sei  $p$  eine beliebige mehrstellige Operation. Sei  $\eta_p(f)(x)$  die punktweise Anwendung von  $p$ . Ein Beispiel ist die Addition  $p(y_1, y_2) := y_1 + y_2$ . Dann ist  $\eta_p(f_1, f_2)(x) = f_1(x) + f_2(x)$ . Sei

$$F(T)(f) := (Tf_1, \dots, Tf_n)$$

die komponentenweise Anwendung eines Operators  $T$ . Sei  $C_\gamma$  der durch  $C_\gamma f := f \circ \gamma$  definierte Kompositionsoperator. Allgemein gilt

$$C_\gamma \circ \eta_p = \eta_p \circ F(C_\gamma).$$

**Beweis.** Prämisse ist, dass der gewöhnliche Ableitungsoperator  $D$  die Regel

$$D(\eta_p(f))(x) = (D \circ \eta_p)(f)(x) = R(f(x), F(D)(f)(x))$$

erfüllt. Für die Richtungsableitung von  $\eta_p(f)$  gilt dann

$$\begin{aligned} D_v(\eta_p(f))(x) &= (\eta_p(f) \circ \gamma)'(0) = (D \circ C_\gamma \circ \eta_p)(f)(0) = (D \circ \eta_p \circ F(C_\gamma))(f)(0) \\ &= (D \circ \eta_p)(F(C_\gamma)(f))(0) = R(F(C_\gamma)(f)(0), F(D)(F(C_\gamma)(f))(0)) \\ &= R(f(x), F(D \circ C_\gamma)(f)(0)) = R(f(x), F(D_v)(f)(x)). \quad \square \end{aligned}$$

Beispiele sind

$$\begin{aligned} p(y_1, y_2) &= y_1 + y_2, & R((y_1, y_2), (y'_1, y'_2)) &= y'_1 + y'_2, \\ p(y_1, y_2) &= y_1 y_2, & R((y_1, y_2), (y'_1, y'_2)) &= y'_1 y_2 + y_1 y'_2, \\ p(y) &= cy, & R(y, y') &= cy', \\ p(y) &= g(y), & R(y, y') &= g'(y)y'. \end{aligned}$$

## 2.4 Fixpunkt-Iterationen

**Definition 2.13 (Kontraktion).** Sei  $(M, d)$  ein vollständiger metrischer Raum. Eine Abbildung  $\varphi: M \rightarrow M$  heißt Kontraktion, wenn sie Lipschitz-stetig mit Lipschitz-Konstante  $L < 1$  ist, d. h.

$$d(\varphi(x), \varphi(y)) < L d(x, y)$$

für alle  $x, y \in M$ .

**Satz 2.27 (Fixpunktsatz von Banach).** Sei  $(M, d)$  ein nichtleerer vollständiger metrischer Raum und sei  $\varphi: M \rightarrow M$  eine Kontraktion. Es gibt genau einen Fixpunkt  $x \in M$  mit  $x = \varphi(x)$  und die Folge  $(x_n): \mathbb{N} \rightarrow M$  mit  $x_{n+1} = \varphi(x_n)$  konvergiert gegen den Fixpunkt, unabhängig vom Startwert  $x_0$ .

**Satz 2.28 (Hinreichendes Konvergenzkriterium).** Sei  $M = [a, b]$ . Ist  $\varphi: M \rightarrow M$  differenzierbar und gibt es eine Zahl  $r$  mit  $|\varphi'(x)| < r < 1$  für alle  $x \in M$ , dann hat  $\varphi$  genau einen Fixpunkt und die Folge  $(x_n)$  mit  $x_{n+1} = \varphi(x_n)$  konvergiert für jeden Startwert  $x_0 \in M$  gegen diesen Fixpunkt.

**Beweis.** Nach Satz 2.14 ist eine differenzierbare Funktion  $\varphi$  mit beschränkter Ableitung auch Lipschitz-stetig, und  $L = \sup_{x \in M} |\varphi'(x)|$  eine Lipschitz-Konstante. Wegen  $|\varphi'(x)| < r$  muss  $L \leq r$  sein, und somit  $L < 1$ . D. h.,  $\varphi$  ist eine Kontraktion. Die Konvergenz der Folge  $(x_n)$  ist gemäß Satz 2.27 gewährleistet.  $\square$

**Satz 2.29 (Hinreichendes Konvergenzkriterium zum Newton-Verfahren).**

Sei  $f: [a, b] \rightarrow \mathbb{R}$  zweimal stetig differenzierbar und  $f'(x) \neq 0$  für alle  $x$ . Sei

$$\varphi: [a, b] \rightarrow [a, b], \quad \varphi(x) := x - \frac{f(x)}{f'(x)}.$$

Man beachte  $\varphi([a, b]) \subseteq [a, b]$ . Gilt für alle  $x$  die Ungleichung

$$|\varphi'(x)| = \left| \frac{f(x)f''(x)}{f'(x)^2} \right| < 1,$$

dann besitzt  $f$  genau eine Nullstelle und die Folge  $(x_n)$  mit  $x_{n+1} = \varphi(x_n)$  konvergiert gegen diese Nullstelle.

**Beweis.** Gemäß den Ableitungsregeln ist  $\varphi$  stetig differenzierbar und es gilt

$$\varphi'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}.$$

Da  $|\varphi'(x)|$  stetig ist, gibt es nach dem Satz vom Minimum und Maximum ein Maximum  $M$  und nach Voraussetzung ist  $M < 1$ . Man setze nun  $r := (M + 1)/2$ . Dann ist  $|\varphi'(x)| < r < 1$ . Gemäß Satz 2.28 konvergiert die Iteration  $(x_n)$  gegen den einzigen Fixpunkt von  $\varphi$ . Wegen  $f'(x) \neq 0$  gilt dabei

$$x = \varphi(x) = x - \frac{f(x)}{f'(x)} \iff \frac{f(x)}{f'(x)} = 0 \iff f(x) = 0.$$

Der Fixpunkt von  $\varphi$  ist also die einzige Nullstelle von  $f$ .  $\square$

# 3 Topologie

## 3.1 Grundbegriffe

### 3.1.1 Definitionen

**Definition 3.1 (Topologischer Raum).** Sei  $X$  eine Menge und  $T$  eine Menge von Teilmengen von  $X$ . Man nennt das System  $T$  eine Topologie und  $(X, T)$  einen topologischen Raum, falls die folgenden drei Axiome erfüllt sind:

1. Es gilt  $\emptyset \in T$  und  $X \in T$ .
2. Sind  $A, B \in T$ , dann ist auch  $A \cap B \in T$ .
3. Sind die  $A_i \in T$ , dann ist auch  $\bigcup_I A_i \in T$ , wobei  $I$  unendlich sein darf.

Die Elemente der Topologie nennt man offene Mengen.

**Definition 3.2 (Abgeschlossene Menge).** Sei  $X$  ein topologischer Raum. Eine Menge  $M \subseteq X$  nennt man abgeschlossen, wenn das Komplement  $X \setminus M$  offen ist.

**Definition 3.3 (nh-filter: Umgebungsfilter).** Zu einem Punkt  $x \in X$  ist

$$\underline{U}(x) := \{U \subseteq X \mid \exists O: O \in T \wedge x \in O \wedge O \subseteq U\}$$

der Umgebungsfilter. Eine Menge  $U \in \underline{U}(x)$  heißt Umgebung von  $x$ .

**Definition 3.4 (int: Inneres).**

Das Innere von  $M$ , auch offener Kern genannt, ist

$$\text{int}(M) := \{x \in M \mid M \in \underline{U}(x)\}.$$

**Definition 3.5 (ext: Äußeres).**

Sei  $X$  ein topologischer Raum und  $M \subseteq X$ . Das Äußere von  $M$  ist

$$\text{ext}(M) := \text{int}(X \setminus M) = \text{int}(M^c).$$

**Definition 3.6 (Abgeschlossene Hülle).**

Sei  $X$  ein topologischer Raum und  $M \subseteq X$ . Die abgeschlossene Hülle von  $M$  ist

$$\overline{M} := X \setminus \text{ext}(M) = \text{int}(M^c)^c.$$

**Definition 3.7 (Rand).** Der Rand einer Menge  $M$  ist

$$\partial M := \overline{M} \setminus \text{int}(M) = \overline{M} \cap \text{int}(M)^c.$$

### 3.1.2 Elementares

**Korollar 3.1.** Sei  $X$  ein topologischer Raum. Für jede Menge  $M \subseteq X$  ist

$$X = \text{int}(M) \cup \partial M \cup \text{ext}(M)$$

eine disjunkte Zerlegung.

**Beweis.** Sei  $A := \text{int}(M)$  und  $B := \text{ext}(M)$ . Dann ist

$$\text{int}(M) \cup \partial M \cup \text{ext}(M) = A \cup B^c \cap A^c \cup B = A \cup A^c \cup B = X \cup B = X.$$

Nun verbleibt zu prüfen, dass die Mengen paarweise disjunkt sind. Wir haben

$$\text{int}(M) \cap \partial M = A \cap B^c \cap A^c = \emptyset,$$

$$\text{ext}(M) \cap \partial M = B \cap B^c \cap A^c = \emptyset.$$

Wegen  $M \cap M^c = \emptyset$  ist erst recht  $A \cap B = \emptyset$ , denn  $A \subseteq M$  und  $B \subseteq M^c$ .  $\square$

**Korollar 3.2.** Für jede Menge  $A \subseteq X$  gilt  $\text{int}(A)^c \cup A = X$ .

**Beweis.** Setze  $B := \text{int}(A)$ . Gemäß Definition gilt  $B \subseteq A$ , was äquivalent zu  $A \cap B = B$  ist. Damit ergibt sich

$$B^c \cup A = (A \cap B)^c \cup A = A^c \cup B^c \cup A = X \cup B^c = X. \quad \square$$

**Satz 3.3.** Das Innere von  $M$  ist die Vereinigung der offenen Teilmengen von  $M$ , kurz

$$\text{int}(M) = \bigcup_{O \in 2^M \cap T} O.$$

**Beweis.** Nach Def. 1.2 (seteq) und Def. 3.4 (int) expandieren:

$$\forall x: [x \in M \wedge M \in \underline{U}(x) \iff x \in \bigcup_{O \in 2^M \cap T} O].$$

Den äußeren Allquantor brauchen wir nicht weiter mitschreiben, da alle freien Variablen automatisch allquantifiziert werden. Nach Def. 3.3 (nh-filter) weiter expandieren, wobei die Bedingung  $U \subseteq X$  als tautologisch entfallen kann, weil  $X$  die Grundmenge ist. Auf der rechten Seite wird nach Def. 1.8 (union) expandiert. Es ergibt sich:

$$x \in M \wedge (\exists O: O \in T \wedge x \in O \wedge O \subseteq M) \iff (\exists O: O \subseteq M \wedge O \in T \wedge x \in O).$$

Wegen  $A \wedge (\exists x: P(x)) \iff (\exists x: A \wedge P(x))$  ergibt sich auf der linken Seite:

$$\exists O: x \in M \wedge O \in T \wedge x \in O \wedge O \subseteq M.$$

Wenn aber  $O \subseteq M$  erfüllt sein muss, gilt  $x \in O \implies x \in M$ . Demnach kann  $x \in M$  entfallen. Auf beiden Seiten steht dann die gleiche Bedingung.  $\square$

**Satz 3.4.** Ein Punkt  $p$  liegt genau dann auf dem Rand einer Menge  $M$ , wenn jede Umgebung von  $p$  mindestens einen Punkt aus  $M$  und einen Punkt aus dem Komplement von  $M$  enthält.



## 3.2 Metrische Räume

### 3.2.1 Metrische Räume

**Definition 3.8 (metric-space: metrischer Raum).** Man bezeichnet  $(M, d)$  mit  $d: M^2 \rightarrow \mathbb{R}$  genau dann als metrischen Raum, wenn die folgenden Axiome erfüllt sind:

- |      |                                   |                                   |
|------|-----------------------------------|-----------------------------------|
| (M1) | $d(x, y) = 0 \iff x = y,$         | (Gleichheit abstandsloser Punkte) |
| (M2) | $d(x, y) = d(y, x),$              | (Symmetrie)                       |
| (M3) | $d(x, y) \leq d(x, z) + d(z, y).$ | (Dreiecksungleichung)             |

**Definition 3.9 (open-ep-ball: offene Epsilon-Umgebung).**

Für einen metrischen Raum  $(M, d)$  und  $p \in M$ :

$$U_\varepsilon(p) := \{x \mid d(p, x) < \varepsilon\}.$$

Bemerkung: Unter einer Epsilon-Umgebung ohne weitere Attribute versteht man immer eine offene Epsilon-Umgebung.

**Satz 3.5 (Konstruktion disjunkter Epsilon-Umgebungen).** Sei  $(M, d)$  ein metrischer Raum und  $p, q \in M$  mit  $p \neq q$ . Betrachte die Streckenzerlegung  $d(p, q) = A + B$ . Für  $a \leq A$  und  $b \leq B$  sind die Epsilon-Umgebungen  $U_a(p)$  und  $U_b(q)$  disjunkt.

**Beweis.** Angenommen  $U_a(p)$  und  $U_b(q)$  wären nicht disjunkt, dann gäbe es mindestens ein  $x$  mit  $x \in U_a(p)$  und  $x \in U_b(q)$ , d. h.  $d(p, x) < a$  und  $d(q, x) < b$ . Addition der beiden Ungleichungen bringt

$$d(p, x) + d(q, x) < a + b \leq d(p, q).$$

Gemäß der Dreiecksungleichung Def. 3.8 Axiom (M3) gilt nun aber

$$d(p, q) \leq d(p, x) + d(q, x)$$

für alle  $x$ . Sei  $c := d(p, x) + d(q, x)$ . Wir erhalten damit nun  $c < a + b \leq c$  und somit den Widerspruch  $c < c$ .  $\square$

**Korollar 3.6 (Unterschiedliche Punkte eines metrischen Raumes besitzen disjunkte Epsilon-Umgebungen).** Sei  $(M, d)$  ein metrischer Raum und  $p, q \in M$ . Wenn  $p \neq q$  ist, dann gibt es disjunkte offene Epsilon-Umgebungen  $U_a(p)$  und  $U_b(q)$ .

**Beweis.** Folgt trivial aus Satz 3.5. Wähle speziell z. B.  $a = b = d(p, q)/2$ .  $\square$

### 3.2.2 Normierte Räume

**Definition 3.10 (normed-space: normierter Raum).** Sei  $V$  ein Vektorraum über dem Körper der reellen oder komplexen Zahlen. Sei  $N(x) = \|x\|$  eine Abbildung, die jedem  $x \in V$  eine reelle Zahl zuordnet. Man nennt  $(V, N)$  genau dann einen normierten Raum, wenn die folgenden Axiome erfüllt sind:

- |      |                                    |                             |
|------|------------------------------------|-----------------------------|
| (N1) | $\ x\  = 0 \iff x = 0,$            | (Definitheit)               |
| (N2) | $\ \lambda x\  =  \lambda  \ x\ ,$ | (betragsmäßige Homogenität) |
| (N3) | $\ x + y\  \leq \ x\  + \ y\ .$    | (Dreiecksungleichung)       |

**Satz 3.7 (umgekehrte Dreiecksungleichung).** In jedem normierten Raum gilt

$$|||x|| - ||y||| \leq ||x - y||.$$

**Beweis.** Auf beiden Seiten von Def. 3.10 (normed-space) Axiom (N3) wird  $||y||$  subtrahiert. Es ergibt sich

$$||x + y|| - ||y|| \leq ||x||.$$

Substitution  $x := x - y$  bringt nun

$$||x|| - ||y|| \leq ||x - y||.$$

Vertauscht man nun  $x$  und  $y$ , dann ergibt sich

$$||y|| - ||x|| \leq ||y - x|| \iff -(||x|| - ||y||) \leq ||x - y||.$$

Wir haben nun  $a \leq b$  und  $-a \leq b$ , wobei  $a := ||x|| - ||y||$  und  $b := ||x - y||$  ist. Multipliziert man die letzte Ungleichung mit  $-1$ , dann ergibt sich  $a \geq -b$ . Somit ist  $-b \leq a \leq b$ , kurz  $|a| \leq b$ .  $\square$

### 3.2.3 Homöomorphismen

**Satz 3.8 (Verallgemeinerung des Zwischenwertsatzes).**

Ist  $f: X \rightarrow Y$  eine stetige Abbildung zwischen topologischen Räumen und  $A \subseteq X$  ein zusammenhängender Teilraum, dann ist auch  $f(A)$  zusammenhängend.

**Satz 3.9.** Eine injektive Abbildung  $f: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  kann nicht stetig sein.

**Beweis.** Da  $f$  injektiv ist, ist die Rechnung

$$f(\mathbb{R}_{>0}) = f(\mathbb{R}_{\geq 0} \setminus \{0\}) = f(\mathbb{R}_{\geq 0}) \setminus f(\{0\}) = \mathbb{R} \setminus \{f(0)\}$$

gültig gemäß Satz 1.50. Da  $\mathbb{R}_{>0}$  zusammenhängend ist,  $\mathbb{R} \setminus \{f(0)\}$  aber nicht, kann  $f$  laut Satz 3.8 nicht stetig sein.  $\square$

## 3.3 Übungen

**Satz 3.10.** Sei  $M \subseteq \mathbb{R}^n$ . Das Innere von  $M$  besitze den Punkt  $p$ , das Äußere den Punkt  $q$ . Dann schneidet das Bild jedes Weges von  $p$  nach  $q$  den Rand von  $M$ .

**Beweis 1.** Sei  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$  so ein Weg mit  $\gamma(0) = p$  und  $\gamma(1) = q$ . Angenommen, das Bild schneidet den Rand nicht. Das heißt,  $\gamma([0, 1]) \cap \partial M = \emptyset$ , oder äquivalent  $\gamma^{-1}(\partial M) = \emptyset$ . Allgemein ist

$$\mathbb{R}^n = \text{int}(M) \cup \partial M \cup \text{ext}(M)$$

laut Korollar 3.1 eine disjunkte Zerlegung. Gemäß Satz 1.35 (preimg-dl) gilt

$$[0, 1] = \gamma^{-1}(\mathbb{R}^n) = \gamma^{-1}(\text{int}(M)) \cup \gamma^{-1}(\partial M) \cup \gamma^{-1}(\text{ext}(M)),$$

was auch eine disjunkte Zerlegung ist, weil je zwei disjunkte Mengen gemäß Korollar 1.38 disjunkte Urbilder haben. Weil  $\gamma$  gemäß Definition stetig ist, sind die Urbilder  $\gamma^{-1}(\text{int}(M))$  und  $\gamma^{-1}(\text{ext}(M))$  offen im Raum  $[0, 1]$ . Sie sind nichtleer, weil sie jeweils laut Prämisse mindestens einen Punkt enthalten. Damit ist  $[0, 1]$  eine Zerlegung in disjunkte nichtleere offene Mengen, gemäß Definition also ein unzusammenhängender

Raum. Das steht im Widerspruch zur Erkenntnis, dass alle Intervalle zusammenhängend sind.  $\square$

**Beweis 2.** Sei  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$  ein solcher Weg mit  $\gamma(0) = p$  und  $\gamma(1) = q$ . Wir nehmen nun eine Bisektion vor. Sei  $a_0 := 0$  und  $b_0 := 1$ . Sei  $m := \frac{1}{2}(a_k + b_k)$ , also der Mittelwert. Liegt  $\gamma(m)$  im Inneren, dann ist  $[a_{k+1}, b_{k+1}] = [m, b_k]$  das nächste Intervall. Liegt  $\gamma(m)$  im Äußeren, dann  $[a_{k+1}, b_{k+1}] = [a_k, m]$ . Liegt  $\gamma(m)$  auf dem Rand, ist ein Schnittpunkt gefunden und das Verfahren bricht ab. Betrachten wir daher den Fall, dass das Verfahren nicht abbricht. Als Intervallschachtelung konvergieren die Folgen  $a_k, b_k$  gegen denselben Grenzwert  $a$ . Weil  $\gamma$  stetig ist, konvergiert  $\gamma(a_k) \rightarrow \gamma(a)$  für  $a_k \rightarrow a$  und  $\gamma(b_k) \rightarrow \gamma(a)$  für  $b_k \rightarrow a$ . Demnach sind in jeder Umgebung von  $\gamma(a)$  sowohl Punkte aus dem Inneren als auch Punkte aus dem Äußeren. Gemäß Satz 3.4 muss  $\gamma(a)$  infolge auf dem Rand liegen.  $\square$



# 4 Lineare Algebra

## 4.1 Matrizen

### 4.1.1 Definitionen

**Definition 4.1 (Transponierte Matrix).** Sei  $R$  ein Ring und  $A \in R^{m \times n}$  eine Matrix. Die Matrix  $A^T \in R^{n \times m}$  mit  $(A^T)_{ij} := A_{ji}$  heißt Transponierte von  $A$ .

**Definition 4.2 (Konjugierte Matrix).** Sei  $A \in \mathbb{C}^{m \times n}$ . Die Matrix  $\bar{A}$  mit  $(\bar{A})_{ij} := \overline{A_{ij}}$  heißt konjugierte Matrix zu  $A$ . Mit  $\overline{A_{ij}}$  ist die Konjugation der komplexen Zahl  $A_{ij}$  gemeint.

**Definition 4.3 (Adjungierte Matrix).** Sei  $A \in \mathbb{C}^{m \times n}$ . Die Adjungierte zu  $A$  ist definiert als  $A^H := (\bar{A})^T$ , d. h. die Transponierte der konjugierten Matrix zu  $A$ .

**Definition 4.4 (Inverse Matrix).** Sei  $K$  ein Körper und  $A \in K^{n \times n}$  eine quadratische Matrix. Man nennt  $A$  invertierbar, wenn es eine Matrix  $B$  gibt, mit  $AB = BA = E_n$ , wobei  $E_n$  die Einheitsmatrix ist. Die Matrix  $A^{-1} := B$  heißt dann inverse Matrix zu  $A$ .

### 4.1.2 Rechenregeln

**Korollar 4.1.** Sei  $R$  ein kommutativer Ring. Für Matrizen  $A \in R^{m \times n}$  und  $B \in R^{n \times p}$  gilt

$$(AB)^T = B^T A^T.$$

**Beweis.** Es gilt:

$$(AB)^T = \left( \sum_{k=1}^n A_{ik} B_{kj} \right)^T = \left( \sum_{k=1}^n A_{jk} B_{ki} \right) = \left( \sum_{k=1}^n B_{ki} A_{jk} \right) \quad (4.1)$$

$$= \left( \sum_{k=1}^n (B^T)_{ik} (A^T)_{kj} \right) = B^T A^T. \quad \square \quad (4.2)$$

**Korollar 4.2.** Sei  $A \in K^{n \times n}$  eine invertierbare Matrix. Dann ist auch  $A^T$  invertierbar und es gilt  $(A^{-1})^T = (A^T)^{-1}$ .

**Beweis.** Aus  $E = A^{-1}A = AA^{-1}$  und Korollar 4.1 folgt

$$E = E^T = (A^{-1}A)^T = A^T (A^{-1})^T = (AA^{-1})^T = (A^{-1})^T A^T. \quad (4.3)$$

Dann muss  $A^T$  nach Def. 4.4 die inverse Matrix zu  $(A^{-1})^T$  sein.  $\square$

**Korollar 4.3.** Sei  $v \in \mathbb{R}^n$  und  $w \in \mathbb{R}^m$ . Sei  $A \in \mathbb{R}^{m \times n}$ . Es gilt  $\langle Av, w \rangle = \langle v, A^T w \rangle$ , wobei links das Standardskalarprodukt auf dem  $\mathbb{R}^m$  und rechts das auf dem  $\mathbb{R}^n$  ausgewertet wird.

**Beweis.** Identifiziert man die Vektoren  $x, y \in \mathbb{R}^k$  mit den Matrizen  $x, y \in \mathbb{R}^{k \times 1}$ , dann ist  $\langle x, y \rangle = x^T y$ . Gemäß Korollar 4.1 darf man rechnen:

$$\langle Av, w \rangle = (Av)^T w = v^T A^T w = \langle v, A^T w \rangle. \quad \square$$

### 4.1.3 Rechenregeln für komplexe Matrizen

**Korollar 4.4.** Für Matrizen  $A \in \mathbb{C}^{m \times n}$  und  $B \in \mathbb{C}^{n \times p}$  gilt

$$\overline{AB} = \bar{A} \cdot \bar{B}$$

**Beweis.** Es gilt

$$\overline{AB} = \overline{\left( \sum_{k=1}^n A_{ik} B_{kj} \right)} = \left( \sum_{k=1}^n \overline{A_{ik} B_{kj}} \right) = \left( \sum_{k=1}^n \overline{A_{ik}} \cdot \overline{B_{kj}} \right) = \left( \sum_{k=1}^n (\bar{A})_{ik} (\bar{B})_{kj} \right) = \bar{A} \cdot \bar{B}. \quad \square$$

**Korollar 4.5.** Für Matrizen  $A \in \mathbb{C}^{m \times n}$  und  $B \in \mathbb{C}^{n \times p}$  gilt

$$(AB)^H = B^H A^H.$$

**Beweis.** Gemäß Korollar 4.4 und 4.1 gilt

$$(AB)^H = (\overline{AB})^T = (\bar{A} \cdot \bar{B})^T = (\bar{B})^T (\bar{A})^T = B^H A^H.$$

**Korollar 4.6.** Sei  $v \in \mathbb{C}^n$  und  $w \in \mathbb{C}^m$ . Sei  $A \in \mathbb{C}^{m \times n}$ . Es gilt  $\langle Av, w \rangle = \langle v, A^H w \rangle$ , wobei links das Standardskalarprodukt auf dem  $\mathbb{C}^m$  ausgewertet wird und rechts das auf dem  $\mathbb{C}^n$ .

**Beweis.** Identifiziert man die Vektoren  $x, y \in \mathbb{C}^k$  mit den Matrizen  $x, y \in \mathbb{C}^{k \times 1}$ , dann gilt  $\langle x, y \rangle = x^H y$ . Gemäß Korollar 4.5 darf man rechnen

$$\langle Av, w \rangle = (Av)^H w = v^H A^H w = \langle v, A^H w \rangle. \quad \square$$

## 4.2 Eigenwerte

**Satz 4.7.** Gegeben sei eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$ . Dann ist die Matrix  $M = A^T A$  symmetrisch und besitzt nur nichtnegative Eigenwerte, speziell bei  $\det(A) \neq 0$  nur positive.

**Beweis.** Gemäß Satz 4.1 gilt

$$M^T = (A^T A)^T = A^T (A^T)^T = A^T A = M. \quad (4.4)$$

Ist nun  $\lambda$  ein Eigenwert von  $M$  und  $v$  ein Eigenvektor dazu, dann gilt  $Mv = \lambda v$ . Unter Anwendung von Korollar 4.3 folgt daraus

$$\lambda |v|^2 = \langle \lambda v, v \rangle = \langle Mv, v \rangle = \langle A^T Av, v \rangle = \langle Av, Av \rangle = |Av|^2 \geq 0. \quad (4.5)$$

Ergo ist  $\lambda |v|^2 \geq 0$ . Unter der Voraussetzung  $v \neq 0$  ist  $|v| > 0$ . Dann muss auch  $\lambda \geq 0$  sein. Wenn nun  $\det(A) \neq 0$  ist, also  $A$  eine reguläre Matrix, dann hat  $A$  trivialen Kern, also  $Av = 0$  nur im Fall  $v = 0$ . Da  $v \neq 0$  vorausgesetzt wurde, muss auch  $Av \neq 0$ , und damit  $|Av| > 0$  sein. Dann ist auch  $\lambda > 0$ . Alternativ folgt  $\lambda > 0$  daraus, dass  $\det(A)$  das Produkt der Eigenwerte ist.  $\square$

**Satz 4.8.** Gegeben sei eine Matrix  $A \in \mathbb{C}^{m \times n}$ . Dann ist die Matrix  $M = A^H A$  hermitisch und besitzt nur nichtnegative reelle Eigenwerte.

**Beweis.** Gemäß Satz 4.5 gilt

$$M^H = (A^H A)^H = A^H (A^H)^H = A^H A = M. \quad (4.6)$$

Ist nun  $\lambda$  ein Eigenwert von  $M$  und  $v$  ein Eigenvektor dazu, dann gilt  $Mv = \lambda v$ . Unter Anwendung von Korollar 4.6 folgt daraus

$$\lambda |v|^2 = \langle \lambda v, v \rangle = \langle Mv, v \rangle = \langle A^H A v, v \rangle = \langle A v, A v \rangle = |A v|^2 \geq 0. \quad (4.7)$$

Ergo ist  $\lambda |v|^2 \geq 0$ . Unter der Voraussetzung  $v \neq 0$  ist  $|v| > 0$ . Dann muss auch  $\lambda \geq 0$  sein.  $\square$

**Definition 4.5 (Unitäre Matrix).**

Eine quadratische Matrix  $A$  heißt unitär, wenn  $A^H A = E$  gilt.

**Korollar 4.9.** Ist  $A$  unitär, dann gilt  $|Av| = |v|$  für jeden Vektor  $v$ .

**Beweis.** Laut Korollar 4.6 gilt

$$|Av|^2 = \langle Av, Av \rangle = \langle v, A^H A v \rangle = \langle v, E v \rangle = \langle v, v \rangle = |v|^2.$$

Radizieren ergibt  $|Av| = |v|$ .  $\square$

**Korollar 4.10.** Für jeden Eigenwert  $\lambda$  einer unitären Matrix gilt  $|\lambda| = 1$ .

**Beweis.** Sei  $v$  ein Eigenvektor zum Eigenwert  $\lambda$ . Laut Korollar 4.6 ist dann

$$|v|^2 = \langle v, v \rangle = \langle v, E v \rangle = \langle v, A^H A v \rangle = \langle A v, A v \rangle = |A v|^2 = |\lambda v|^2 = |\lambda|^2 |v|^2.$$

Daher ist  $|\lambda|^2 = 1$ , und wegen  $|\lambda| \geq 0$  folglich  $|\lambda| = 1$ .  $\square$

### 4.2.1 Quadratische Matrizen

**Satz 4.11.** Sei

$$I := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad aE + bI = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Die Menge  $M := \{aE + bI \mid a, b \in \mathbb{R}\}$  bildet bezüglich Matrizenaddition und Matrizenmultiplikation einen Körper  $(M, +, \cdot)$ . Die Abbildung

$$\Phi: \mathbb{C} \rightarrow M, \quad \Phi(a + bi) := aE + bI$$

ist ein Isomorphismus zwischen Körpern.

**Beweis.** Bei  $(M, +)$  handelt es sich um eine Untergruppe der kommutativen Gruppe  $(\mathbb{R}^{2 \times 2}, +)$ , denn gemäß

$$(aE + bI) + (cE + dI) = (a + b)E + (b + d)I \in M \quad (4.8)$$

und

$$-(aE + bI) = (-a)E + (-b)I \in M \quad (4.9)$$

ist das Untergruppenkriterium erfüllt. Die Abgeschlossenheit bezüglich Multiplikation:

$$\begin{aligned} (aE + bI)(cE + dI) &= aEcE + aEdI + bIcE + bIdI \\ &= acE + adI + bcI + bdI^2 = (ac - bd)E + (ad + bc)I \in M. \end{aligned} \quad (4.10)$$

Das Kommutativgesetz:

$$\begin{aligned} (aE + bI)(cE + dI) &= (ac - bd)E + (ad + bc)I \\ &= (ca - db)E + (cb + da)I = (cE + dI)(aE + bI). \end{aligned} \quad (4.11)$$

Das Assoziativgesetz ist für Matrizen allgemeingültig. Das multiplikativ neutrale Element ist die Einheitsmatrix  $E$ . Wird nun  $aE + bI \neq 0$  vorausgesetzt, dann ist  $a \neq 0 \vee b \neq 0$ . Daher ist  $\det(aE + bI) = a^2 + b^2 \neq 0$ . Demnach besitzt  $aE + bI$  eine Inverse. Somit muss  $(M, +, \cdot)$  ein Körper sein.

Die Abbildung  $\Phi$  ist invertierbar, denn jedes Bild  $A$  kann auf eindeutige Art in  $A = aE + bI$  zerlegt werden, wodurch  $a, b$  eindeutig bestimmt sind. Die Eigenschaften

$$\Phi((a + bi) + (c + di)) = \Phi(a + bi) + \Phi(c + di) \quad (4.12)$$

und

$$\Phi((a + bi)(c + di)) = \Phi(a + bi)\Phi(c + di) \quad (4.13)$$

ergeben sich aus den Rechnungen (4.8) und (4.10).  $\square$



## 4.3 Bilinearformen

### Definition 4.6 (Nicht-ausgeartete Bilinearform).

Sei  $B: V \times W \rightarrow K$  eine Bilinearform, sei

$$\begin{aligned} B_1: V &\rightarrow W^*, & B_1(v)(w) &:= B(v, w), \\ B_2: W &\rightarrow V^*, & B_2(w)(v) &:= B(v, w). \end{aligned}$$

Man nennt  $B$  nicht-ausgeartet, wenn  $B_1$  und  $B_2$  injektiv sind.

**Korollar 4.12.** Eine Bilinearform  $B: V \times W \rightarrow K$  ist genau dann nicht-ausgeartet, wenn  $B_1(v)$  für alle  $v \neq 0$  und  $B_2(w)$  für alle  $w \neq 0$  nicht die Nullabbildung ist. Die Abbildungen  $B_1, B_2$  aus Def. 4.6.

**Beweis.** Die lineare Abbildung  $B_1$  ist genau dann injektiv, wenn

$$\{0\} = \text{Kern}(B_1) := \{v \mid B_1(v) = 0\} \quad (4.14)$$

ist. Wegen  $B_1(0) = 0$  ist  $B_1$  schon dann injektiv, wenn

$$B_1(v) = 0 \implies v = 0, \quad (4.15)$$

was per Kontraposition äquivalent ist zu  $v \neq 0 \implies B_1(v) \neq 0$ . Für  $B_2$  gilt eine analoge Argumentation.  $\square$

**Korollar 4.13.** Eine symmetrische Bilinearform  $B: V \times V \rightarrow K$  ist genau dann nicht-ausgeartet, wenn es für alle  $v \neq 0$  ein  $w$  gibt, so dass  $B(v, w) \neq 0$ .

**Beweis.** Da  $B$  symmetrisch ist, ist  $B_1 = B_2$  in Def. 4.6. Es genügt also,  $B_1$  zu betrachten. Nun gilt

$$B_1(v) = 0 \iff (\forall w: B_1(v)(w) = 0(w)) \iff (\forall w: B(v, w) = 0). \quad (4.16)$$

Aus Korollar 4.12 ergibt sich dann die Behauptung, d. h. die Äquivalenz zu

$$v \neq 0 \implies \exists w: B(v, w) \neq 0. \quad \square \quad (4.17)$$

**Korollar 4.14.** Ein reelles Skalarprodukt  $\langle v, w \rangle$  ist nicht-ausgeartet.

**Beweis.** In Korollar 4.13 setze  $B(v, w) := \langle v, w \rangle$ . Wegen

$$\langle v, v \rangle = 0 \iff v = 0 \quad (4.18)$$

kann man für  $v \neq 0$  immer  $w := v$  setzen, dann ist  $B(v, w) = \langle v, v \rangle \neq 0$ .  $\square$

**Satz 4.15.** Sind  $V, W$  endlichdimensional, dann sind bei einer nicht-ausgearteten Bilinearform  $B: V \times W \rightarrow K$  die Abbildungen  $B_1, B_2$  aus Def. 4.6 Isomorphismen.

**Beweis.** Es gilt  $\dim B_1(V) \leq \dim W^*$  und  $\dim B_2(W) \leq \dim V^*$ . Gemäß Rangsatz erhält man  $\dim V = \dim B_1(V)$  und  $\dim W = \dim B_2(W)$ , da  $B_1, B_2$  nach Voraussetzung injektiv sind. Demnach ist

$$\dim V \leq \dim W^* = \dim W \leq \dim V^* = \dim V. \quad (4.19)$$

Folglich muss  $\dim V = \dim W = \dim V^* = \dim W^*$  sein. Somit haben  $B_1, B_2$  vollen Rang, sind also surjektiv.  $\square$

## 4.4 Euklidische Geometrie

### Satz 4.16 (Satz des Thales).

Gegeben seien zwei Punkte  $A, B$ , deren Strecke ein Durchmesser des Kreises ist. Sei  $C$  ein beliebiger weiterer Punkt auf dem Kreis. Dann ist das Dreieck  $\triangle ABC$  rechtwinklig.

**Beweis.** Wählt man den Mittelpunkt des Kreises als Ursprung aus, wird die Ebene zu einem euklidischen Vektorraum. Jeder Punkt kann nun mit seinem Ortsvektor identifiziert werden, setze  $\mathbf{a} := A$ ,  $\mathbf{b} := B$ ,  $\mathbf{c} := C$ . Zu zeigen ist, dass  $\mathbf{v} := \mathbf{c} - \mathbf{a}$  rechtwinklig auf  $\mathbf{w} := \mathbf{c} - \mathbf{b}$  steht. Das ist genau dann der Fall, wenn  $\langle \mathbf{v}, \mathbf{w} \rangle = 0$  ist. Man beachte  $\mathbf{b} = -\mathbf{a}$ . Aufgrund der Bilinearität und Symmetrie des Skalarproduktes ergibt sich

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{c} - \mathbf{a}, \mathbf{c} + \mathbf{a} \rangle = \langle \mathbf{c}, \mathbf{c} \rangle + \langle \mathbf{c}, \mathbf{a} \rangle - \langle \mathbf{c}, \mathbf{a} \rangle - \langle \mathbf{a}, \mathbf{a} \rangle \quad (4.20)$$

$$= |\mathbf{c}|^2 - |\mathbf{a}|^2 = 0. \quad (4.21)$$

Die letzte Gleichung gilt wegen  $|\mathbf{a}| = |\mathbf{c}|$ .  $\square$

### Satz 4.17 (Kosinussatz).

Gegeben ist ein Dreieck  $\triangle ABC$ . Sei  $\gamma$  der Winkel  $\angle ACB$ . Dann gilt

$$c^2 = a^2 + b^2 - 2ab \cos \gamma.$$

**Beweis.** Sei  $\mathbf{a} := \overrightarrow{CB}$ ,  $\mathbf{b} := \overrightarrow{CA}$ , und  $\mathbf{c} := \mathbf{a} - \mathbf{b}$ . Dann gilt  $a = |\mathbf{a}|$ ,  $b = |\mathbf{b}|$  und  $c = |\mathbf{c}|$ . Die Rechenregeln des Skalarproduktes gestatten nun die folgende Rechnung:

$$c^2 = |\mathbf{a} - \mathbf{b}|^2 = \langle \mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle = \langle \mathbf{a}, \mathbf{a} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle - 2\langle \mathbf{a}, \mathbf{b} \rangle \quad (4.22)$$

$$= a^2 + b^2 - 2ab \cos \gamma. \quad (4.23)$$

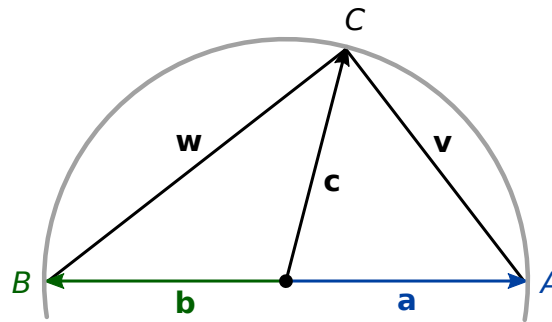


Abbildung 4.1: Zeichnung zum Satz des Thales

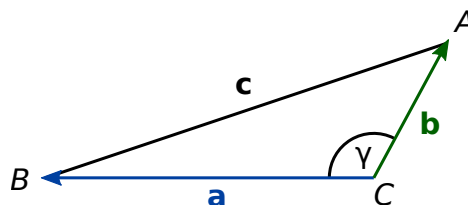


Abbildung 4.2: Zeichnung zum Kosinussatz

**Satz 4.18 (Sinussatz).** Für jedes Dreieck  $\triangle ABC$  gilt

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma} = \frac{abc}{2A},$$

wobei  $A$  der Flächeninhalt ist.

**Beweis.** Sei  $\mathbf{a} := \overrightarrow{CB}$ ,  $\mathbf{b} := \overrightarrow{CA}$  und  $\mathbf{c} := \mathbf{a} - \mathbf{b}$ . Dann gilt

$$ab \sin \gamma \mathbf{e}_1 \wedge \mathbf{e}_2 = \mathbf{b} \wedge \mathbf{a},$$

$$bc \sin \alpha \mathbf{e}_1 \wedge \mathbf{e}_2 = \mathbf{c} \wedge (-\mathbf{b}) = \mathbf{b} \wedge \mathbf{c} = \mathbf{b} \wedge (\mathbf{a} - \mathbf{b}) = \mathbf{b} \wedge \mathbf{a},$$

$$ac \sin \beta \mathbf{e}_1 \wedge \mathbf{e}_2 = (-\mathbf{a}) \wedge (-\mathbf{c}) = \mathbf{a} \wedge \mathbf{c} = \mathbf{a} \wedge (\mathbf{a} - \mathbf{b}) = -\mathbf{a} \wedge \mathbf{b} = \mathbf{b} \wedge \mathbf{a}$$

und  $\mathbf{b} \wedge \mathbf{a} = 2A \mathbf{e}_1 \wedge \mathbf{e}_2$ . Demnach gilt

$$ab \sin \gamma = bc \sin \alpha = ac \sin \beta = 2A.$$

Umformung der Gleichung führt zur Behauptung.  $\square$



# 5 Algebra

## 5.1 Gruppentheorie

### 5.1.1 Grundlagen

**Definition 5.1 (Gruppe).** Das Tupel  $(G, *)$  bestehend aus einer Menge  $G$  und Abbildung  $* : G \times G \rightarrow \Omega$  heißt Gruppe, wenn die folgenden Axiome erfüllt sind:

(G1) Für alle  $a, b \in G$  gilt  $a * b \in G$ . D. h., man darf  $G = \Omega$  setzen.

(G2) Es gilt das Assoziativgesetz: für alle  $a, b, c \in G$  gilt  $(a * b) * c = a * (b * c)$ .

(G3) Es gibt ein Element  $e \in G$ , so dass  $e * g = g = g * e$  für jedes  $g \in G$  gilt.

(G4) Zu jedem  $g \in G$  gibt es ein  $g^{-1} \in G$  so dass  $g * g^{-1} = e = g^{-1} * g$  gilt.

Das Element  $e$  wird neutrales Element der Gruppe genannt. Das Element  $g^{-1}$  wird inverses Element zu  $g$  genannt. Anstelle von  $a * b$  schreibt man auch kurz  $ab$ . Ist  $(G, +)$  eine Gruppe, dann schreibt man immer  $a + b$ , und  $-g$  anstelle von  $g^{-1}$ .

**Korollar 5.1.** Das neutrale Element einer Gruppe  $G$  ist eindeutig bestimmt. D. h., es gibt keine zwei unterschiedlichen neutralen Elemente.

**Beweis.** Seien  $e, e'$  zwei neutrale Elemente von  $G$ . Nach Axiom (G3) gilt dann  $e = e'e$ , und weiter  $e'e = e'$  bei nochmaliger Anwendung von (G3). Daher ist  $e = e'$ .  $\square$

**Korollar 5.2.** Sei  $G$  eine Gruppe. Zu jedem Element  $g \in G$  ist das inverse Element  $g^{-1}$  eindeutig bestimmt. D. h., es kann keine zwei unterschiedlichen inversen Elemente zu  $g$  geben.

**Beweis.** Seien  $a, b$  zwei inverse Elemente zu  $g$ . Nach Axiom (G3), Axiom (G2) und Axiom (G4) gilt

$$a \stackrel{(G3)}{=} ae \stackrel{(G4)}{=} a(gb) \stackrel{(G2)}{=} (ag)b \stackrel{(G4)}{=} eb \stackrel{(G3)}{=} b.$$

Daher ist  $a = b$ .  $\square$

**Definition 5.2 (Untergruppe).** Sei  $(G, *)$  eine Gruppe. Eine Teilmenge  $U \subseteq G$  heißt Untergruppe von  $G$ , kurz  $U \leq G$ , wenn  $U$  bezüglich derselben Verknüpfung  $*$  selbst eine Gruppe  $(U, *)$  bildet.

**Korollar 5.3.** Jede Gruppe  $G$  besitzt die Untergruppen  $\{e\} \leq G$  und  $G \leq G$ , wobei  $e \in G$  das neutrale Element ist. Man spricht von den trivialen Untergruppen.

**Beweis.** Die Aussage  $G \leq G$  ist trivial, denn  $G \subseteq G$  ist allgemeingültig und  $(G, *)$  bildet nach Voraussetzung eine Gruppe. Zu (G1): Es gilt  $ee = e$ . Da es nur diese eine Möglichkeit gibt, sind damit alle überprüft. Zu (G2): Das Assoziativgesetz wird auf Elemente der Teilmenge vererbt. Zu (G3): Das neutrale Element ist in  $\{e\}$  enthalten. Zu (G4): Das neutrale Element ist gemäß  $ee = e$  zu sich selbst invers. Da  $e$  das einzige Element von  $\{e\}$  ist, sind damit alle überprüft.  $\square$

**Satz 5.4 (Untergruppenkriterium).** Sei  $G$  eine Gruppe und  $H$  eine nichtleere Teilmenge von  $G$ . Sind die beiden Prämissen

1.  $a, b \in H \implies ab \in H$ ,
2.  $a \in H \implies a^{-1} \in H$ ,

erfüllt, ist  $H$  bereits eine Untergruppe von  $G$ .

**Beweis.** Die Abgeschlossenheit gilt wegen der ersten Prämisse. Das Assoziativgesetz vererbt sich auf die Elemente der Teilmenge. Nun existiert mindestens ein  $x \in H$ . Aufgrund der zweiten Prämisse muss somit  $x^{-1} \in H$  sein und infolge  $e = x^{-1}x$  ein Element von  $H$ . Die zweite Prämisse sichert schließlich ab, dass jedes weitere Element von  $H$  ein Inverses in  $H$  besitzt. Der Nachweis aller Gruppenaxiome ist erbracht.  $\square$

**Satz 5.5.** Sei  $X$  eine Menge. Die Menge der bijektiven Selbstabbildungen auf  $X$  bildet bezüglich der Verkettung eine Gruppe, die man als symmetrische Gruppe  $S(X)$  bezeichnet.

**Beweis.** Die Verkettung zweier Bijektionen ist ebenfalls bijektiv. Das Assoziativgesetz gilt für die Verkettung allgemein. Das neutrale Element ist offenbar die identische Abbildung. Zu einer Bijektion  $f$  nimmt die Umkehrabbildung  $f^{-1}$  die Rolle des inversen Elements ein, denn  $f \circ f^{-1} = \text{id}$  und  $f^{-1} \circ f = \text{id}$ . Demnach sind alle Gruppenaxiome erfüllt.  $\square$

**Korollar 5.6.** Sei  $X$  eine nichtleere algebraische Struktur, beispielsweise eine Gruppe, ein Ring oder Vektorraum. Die Menge der Automorphismen bildet bezüglich Verkettung eine Gruppe, die man Automorphismengruppe  $\text{Aut}(X)$  nennt. Sie ist eine Untergruppe der symmetrischen Gruppe  $S(X)$ .

**Beweis.** Offenbar ist  $\text{Aut}(X)$  eine nichtleere Teilmenge von  $S(X)$ . Die Verkettung zweier Automorphismen ist ebenfalls ein Automorphismus. Jeder Automorphismus besitzt einen inversen Automorphismus. Laut Untergruppenkriterium muss  $\text{Aut}(X)$  eine Untergruppe von  $S(X)$  sein.  $\square$

**Korollar 5.7.** Sei  $K$  ein Körper. Die invertierbaren Matrizen  $A \in K^{n \times n}$  bilden eine Gruppe, die man allgemeine lineare Gruppe  $\text{GL}(n, K)$  nennt. Sie ist kanonisch isomorph zur Automorphismengruppe  $\text{Aut}(K^n)$ .

**Beweis.** Die Gruppenaxiome sind erfüllt, wobei die Einheitsmatrix das neutrale Element ist und die jeweilige inverse Matrix das inverse Element zu einer Matrix. Weil die Abbildung

$$\varphi: K^{m \times n} \rightarrow \text{Hom}(K^n, K^m), \quad \varphi(A)(\mathbf{v}) := A\mathbf{v}$$

bereits einen kanonischen Isomorphismus darstellt, erhält man bei Setzung  $m = n$  und Einschränkung auf  $\text{GL}(n, K)$  einen Monomorphismus. Weil zu jeder bijektiven linearen Abbildung genau eine invertierbare Matrix gehört, muss

$$\varphi: \text{GL}(n, K) \rightarrow \text{Aut}(K^n)$$

außerdem surjektiv sein.  $\square$

## 5.2 Ringtheorie

### 5.2.1 Grundlagen

**Definition 5.3 (Ring).** Eine Struktur  $(R, +, \cdot)$  heißt genau dann Ring, wenn die folgenden Axiome erfüllt sind

1.  $(R, +)$  ist eine kommutative Gruppe.
2.  $(R, \cdot)$  ist eine Halbgruppe.
3. Für alle  $a, b, c \in R$  gilt  $a(b + c) = ab + ac$ . (Links distributivgesetz)
4. Für alle  $a, b, c \in R$  gilt  $(a + b)c = ac + bc$ . (Rechts distributivgesetz)

Bemerkung: Das neutrale Element von  $(R, +)$  wird als Nullelement bezeichnet und meist 0 geschrieben.

**Definition 5.4 (Ring mit Eins).** Ein Ring  $R$  heißt genau dann Ring mit Eins, wenn  $(R, \cdot)$  ein Monoid ist. Monoid heißt, es gibt ein Element  $e \in R$ , so dass  $e \cdot a = a$  und  $a \cdot e = a$  für alle  $a \in R$ .

Bemerkung: Man bezeichnet  $e$  als Einselement des Rings.

**Korollar 5.8.** Sei  $R$  ein Ring und  $0 \in R$  das Nullelement. Für jedes  $a \in R$  gilt  $0 \cdot a = 0$  und  $a \cdot 0 = 0$ .

**Beweis.** Man rechnet

$$0a = 0a + 0 = 0a + 0a - 0a = (0 + 0)a - 0a = 0a - 0a = 0.$$

Die Rechnung für  $a \cdot 0$  ist analog.  $\square$

**Korollar 5.9.** Sei  $R$  ein Ring und  $a, b \in R$ , dann gilt  $(-a)b = -(ab) = a(-b)$ .

**Beweis.** Man rechnet

$$\begin{aligned} (-a)b &= (-a)b + 0 = (-a)b + ab - (ab) = ((-a) + a)b - (ab) \\ &= 0b - (ab) = 0 - (ab) = -(ab). \quad \square \end{aligned}$$

**Korollar 5.10 (»Minus mal minus macht plus«).**

Sei  $R$  ein Ring und  $a, b \in R$ , dann gilt  $(-a)(-b) = ab$ .

Beachtung von  $-(-x) = x$  nach zweifacher Anwendung von Korollar 5.9 bringt

$$(-a)(-b) = -((-a)b) = -(-(ab)) = ab. \quad \square$$

**Definition 5.5.** Sei  $(M, +, 0)$  ein Monoid. Für  $n \in \mathbb{Z}_{\geq 0}$  und  $a \in M$  definiert man  $na$  rekursiv als

$$0a := 0, \quad (n+1)a := na + a.$$

**Korollar 5.11.** Sei  $(M, +, 0)$  ein Monoid. Für  $m, n \in \mathbb{Z}_{\geq 0}$  und  $a \in M$  gilt

$$(m+n)a = ma + na, \tag{5.1}$$

$$(mn)a = m(na). \tag{5.2}$$

**Beweis.** Induktionsanfang ist  $(m + 0)a = ma = ma + 0 = ma + 0a$ . Der Schritt ist

$$(m + n + 1)a \stackrel{(\text{Def})}{=} (m + n)a + a \stackrel{(\text{IV})}{=} ma + na + a \stackrel{(\text{Def})}{=} ma + (n + 1)a.$$

Induktionsanfang ist  $(m \cdot 0)a = 0a = 0 = m \cdot 0 = m(0a)$ . Der Schritt ist

$$\begin{aligned} (m(n + 1))a &= (mn + m)a = (mn)a + ma \stackrel{(\text{IV})}{=} m(na) + ma \\ &= m((na) + a) \stackrel{(\text{Def})}{=} m((n + 1)a). \quad \square \end{aligned}$$

**Korollar 5.12.** Sei  $R$  ein Ring mit Eins  $e$ . Für  $n \in \mathbb{Z}$  und  $a \in R$  gilt  $(ne)a = na$ .

**Beweis.** Der Induktionsanfang ist  $(0e)a = 0_R a = 0_R = 0a$ . Der Schritt ist

$$((n + 1)e)a \stackrel{(\text{Def})}{=} (ne + e)a = nea + ea \stackrel{(\text{IV})}{=} na + a \stackrel{(\text{Def})}{=} (n + 1)a.$$

Für  $n > 0$  gilt zudem

$$((-n)e)a \stackrel{(\text{Def})}{=} -(ne)a = -((ne)a) = -(na) \stackrel{(\text{Def})}{=} (-n)a. \quad \square$$



### 5.2.2 Ringhomomorphismen

**Definition 5.6 (Ringhomomorphismus).** Seien  $R, R'$  Ringe. Eine Abbildung  $\varphi: R \rightarrow R'$  heißt Ringhomomorphismus, falls

$$\varphi(x + y) = \varphi(x) + \varphi(y),$$

$$\varphi(xy) = \varphi(x)\varphi(y)$$

für alle  $x, y \in R$  gilt. Liegen Ringe mit Eins vor, und gilt zusätzlich  $\varphi(1) = 1$ , dann spricht man von einem unitären Ringhomomorphismus.

**Korollar 5.13.** Bei jedem Ringhomomorphismus  $\varphi$  gilt  $\varphi(kx) = k\varphi(x)$  für  $k \in \mathbb{Z}$ .

**Beweis.** Für  $k > 0$  ist

$$\varphi(kx) = \varphi\left(\sum_{i=1}^k x\right) = \sum_{i=1}^k \varphi(x) = k\varphi(x).$$

Nun der Fall  $k = 0$ . Man rechnet  $f(0) = f(0 + 0) = f(0) + f(0)$ . Subtraktion von  $f(0)$  auf beiden Seiten ergibt  $f(0) = 0$ . Schließlich bleibt noch  $f(-kx) = -kf(x)$  für  $k > 0$  zu zeigen. Hier rechnet man zunächst

$$0 = f(0) = f(-x + x) = f(-x) + f(x).$$

Subtraktion von  $f(x)$  auf beiden Seiten ergibt  $f(-x) = -f(x)$ . Somit gilt

$$f(-kx) = -f(kx) = -kf(x). \quad \square$$

## 5.3 Polynomringe

### 5.3.1 Einsetzungshomomorphismus

**Definition 5.7 (Einsetzungshomomorphismus).**

Die Abbildung  $\Phi: \mathbb{R}[X] \rightarrow \text{Abb}(\mathbb{R}, \mathbb{R})$  mit  $\Phi(f)(x) := f(x)$  nennt man Einsetzungshomomorphismus.

**Satz 5.14.** Beim Einsetzungshomomorphismus  $\Phi$  handelt es sich um eine lineare Abbildung.

**Beweis.** Es gilt

$$\begin{aligned} \Phi(f + g)(x) &= \sum_k (a_k + b_k)x^k = \sum_k a_k x^k + \sum_k b_k x^k \\ &= \Phi(f)(x) + \Phi(g)(x) = (\Phi(f) + \Phi(g))(x) \end{aligned}$$

und

$$\Phi(\lambda f)(x) = \sum_k \lambda a_k x^k = \lambda \sum_k a_k x^k = \lambda \Phi(f)(x) = (\lambda \Phi(f))(x). \quad \square$$

**Satz 5.15.** Der Einsetzungshomomorphismus  $\Phi$  ist injektiv.

**Beweis.** Sei  $f = \sum_{k=0}^n a_k X^k$  und  $g = \sum_{k=0}^n b_k X^k$ , wobei  $n = \max(\deg f, \deg g)$ . Zu zeigen ist

$$(\forall x \in \mathbb{R}: \Phi(f)(x) = \Phi(g)(x)) \implies f = g,$$

das heißt

$$(\forall x \in \mathbb{R}: \sum_k a_k x^k = \sum_k b_k x^k) \implies (\forall k: a_k = b_k).$$

Die Umformung der Voraussetzung ergibt  $\sum_k (b_k - a_k) x^k = 0$ . D. h., jedes der  $(b_k - a_k)$  muss verschwinden. Zu zeigen ist also lediglich

$$(\forall x \in \mathbb{R}: \sum_{k=0}^n c_k x^k = 0) \implies (\forall k: c_k = 0).$$

Wenn  $f(x) = 0$  für alle  $x$  ist, muss auch die Ableitung  $D^m f(x) = 0$  sein. Es gilt  $D^k x^k = k!$ , und daher

$$D^n \sum_{k=0}^n c_k x^k = n! \cdot c_n = 0 \implies c_n = 0.$$

Demnach ergibt sich dann aber auch

$$D^{n-1} \sum_{k=0}^n c_k x^k = (n-1)! \cdot c_{n-1} = 0 \implies c_{n-1} = 0$$

usw. Man erhält  $c_k = 0$  für alle  $k$ .  $\square$

## 5.4 Körper

### 5.4.1 Geordnete Körper

**Definition 5.8 (Geordneter Körper).** Eine Körper  $K$  heißt geordnet bezüglich einer strengen Totalordnung, wenn die beiden Axiome

$$\begin{aligned} a < b &\implies a + c < b + c, \\ a > 0 \wedge b > 0 &\implies ab > 0 \end{aligned}$$

für alle  $a, b, c \in K$  erfüllt sind.

**Korollar 5.16.** Sei  $K$  ein geordneter Körper und  $a \in K$ . Es gilt  $a > 0 \iff -a < 0$ .

**Beweis.** Gemäß Def. 5.8 Axiom 1 gilt

$$0 < a \iff -a < a - a = 0. \square$$

**Korollar 5.17.** Sei  $K$  ein geordneter Körper. Für jedes  $a \in K$  mit  $a \neq 0$  gilt  $a^2 > 0$ .

**Beweis.** Im Fall  $a > 0$  ist  $a^2 > 0$  aufgrund von Def. 5.8 Axiom 2, setze  $b := a$ . Übrig bleibt der Fall  $a < 0$ . Gemäß Korollar 5.16 ist  $-a > 0$ . Laut Axiom 2 gilt andererseits

$$-a > 0 \implies 0 < (-a)^2 = a^2.$$

Somit ist auch in diesem Fall  $a^2 > 0$ .  $\square$

**Korollar 5.18.** Sei  $K$  ein geordneter Körper. Für  $1_K$  gilt  $1_K > 0$ .

**Beweis.** Wegen  $1_K = 1_K 1_K$  ist  $1_K$  ein Quadrat. Weil zudem  $1_K \neq 0$  ist, erhält man  $1_K > 0$  gemäß Korollar 5.17.  $\square$

**Korollar 5.19.** Sei  $K$  ein geordneter Körper. Für  $n \in \mathbb{Z}_{>0}$  gilt  $n1_K > 0$ .

**Beweis.** Der Induktionsanfang ist  $1 \cdot 1_K = 1_K > 0$ , wobei  $1_K > 0$  gemäß Korollar 5.18 gilt. Der Induktionsschritt ist

$$(n+1)1_K \stackrel{(\text{Def})}{=} n1_K + 1_K > 1_K > 0,$$

denn die Induktionsvoraussetzung  $n1_K > 0$  impliziert  $n1_K + 1_K > 1_K$  gemäß Def. 5.8 Axiom 1.  $\square$

**Korollar 5.20.** Sei  $K$  ein geordneter Körper. Für  $a, b, c \in K$  gilt

$$\begin{aligned} a < b \wedge c > 0 &\implies ca < cb, \\ a \leq b \wedge c \geq 0 &\implies ca \leq cb. \end{aligned}$$

**Beweis.** Gemäß Def. 5.8 Axiom 1 und Axiom 2 gilt

$$a < b \iff 0 < b - a \implies 0 < c(b - a) = cb - ca \iff ca < cb.$$

Im Fall  $c = 0$  ist  $ca \leq cb$  klar. Für  $c > 0$  gilt

$$a \leq b \iff a < b \vee a = b \implies ca < cb \vee ca = cb \iff ca \leq cb. \square$$

**Korollar 5.21.** Sei  $K$  ein geordneter Körper und  $n \in \mathbb{Z}_{\geq 0}$ . Für  $a \geq 0$  gilt  $na \geq 0$ .

**Beweis.** Der Induktionsanfang  $n = 0$  ist trivial. Der Schritt ist

$$(n+1)a \stackrel{(\text{Def})}{=} na + a \geq na \geq 0,$$

denn gemäß Induktionsvoraussetzung ist  $na \geq 0$ . Die Ungleichung  $na + a \geq na$  erhält man, indem auf beiden Seiten von  $a \geq 0$  der Wert  $an$  addiert wird.  $\square$

**Korollar 5.22.** Sei  $K$  ein geordneter Körper. Sei  $a \in K$  und  $m, n \in \mathbb{Z}$ . Es gilt

$$m \leq n \wedge a \geq 0 \implies ma \leq na.$$

**Beweis.** Wegen  $m \leq n$  ist  $n - m \geq 0$ . Laut Korollar 5.21 gilt somit

$$0 \leq a \implies 0 \leq (n - m)a = na - ma \iff ma \leq na. \square$$



# 6 Kombinatorik

## 6.1 Endliche Mengen

### 6.1.1 Indiktorfunktion

**Definition 6.1 (Iverson-Klammer).**

Für eine Aussage  $A$  der klassischen Aussagenlogik definiert man

$$[A] := \begin{cases} 1 & \text{wenn } A, \\ 0 & \text{sonst.} \end{cases}$$

**Korollar 6.1.** Es gilt

$$\begin{aligned} [A \wedge B] &= [A][B], \\ [A \vee B] &= [A] + [B] - [A][B], \\ [\neg A] &= 1 - [A], \\ [A \rightarrow B] &= 1 - [A](1 - [B]). \end{aligned}$$

**Beweis.** Trivial mittels Wertetabelle.  $\square$

**Korollar 6.2.** Für die Indikatorfunktion  $1_M(x) := [x \in M]$  gilt

$$\begin{aligned} 1_{A \cap B} &= 1_A 1_B, \\ 1_{A \cup B} &= 1_A + 1_B - 1_{A \cap B}. \end{aligned}$$

**Beweis.** Gemäß Korollar 6.1 darf gelten die Rechnungen

$$1_{A \cap B}(x) = [x \in A \cap B] = [x \in A \wedge x \in B] = [x \in A][x \in B] = 1_A(x) 1_B(x)$$

und

$$\begin{aligned} 1_{A \cup B}(x) &= [x \in A \cup B] = [x \in A \vee x \in B] = [x \in A] + [x \in B] - [x \in A][x \in B] \\ &= 1_A(x) + 1_B(x) - 1_{A \cap B}(x). \quad \square \end{aligned}$$

**Satz 6.3.** Für endliche Mengen  $A, B$  gilt  $|A \cup B| = |A| + |B| - |A \cap B|$ .

**Beweis.** Gemäß Korollar 6.2 darf man rechnen

$$\begin{aligned} |A \cup B| &= \sum_{x \in G} 1_{A \cup B}(x) = \sum_{x \in G} (1_A(x) + 1_B(x) - 1_{A \cap B}(x)) \\ &= \sum_{x \in G} 1_A(x) + \sum_{x \in G} 1_B(x) - \sum_{x \in G} 1_{A \cap B}(x) = |A| + |B| - |A \cap B|. \quad \square \end{aligned}$$

## 6.2 Summen

**Definition 6.2 (Summe).** Sei  $(M, +, 0)$  ein kommutatives Monoid und  $a_k \in M$ . Die Summe ist rekursiv definiert als

$$\sum_{k=m}^{m-1} a_k := 0, \quad \sum_{k=m}^n a_k := a_n + \sum_{k=m}^{n-1} a_k.$$

**Korollar 6.4.** Es gilt

$$\sum_{k=m}^n (a_k + b_k) = \sum_{k=m}^n a_k + \sum_{k=m}^n b_k.$$

**Beweis.** Induktion über  $n$ . Induktionsanfang:

$$\sum_{k=m}^{m-1} (a_k + b_k) = 0 = 0 + 0 = \sum_{k=m}^{m-1} a_k + \sum_{k=m}^{m-1} b_k.$$

Induktionsschritt:

$$\begin{aligned} \sum_{k=m}^n (a_k + b_k) &= a_n + b_n + \sum_{k=m}^{n-1} (a_k + b_k) \stackrel{\text{IV}}{=} a_n + b_n + \sum_{k=m}^{n-1} a_k + \sum_{k=m}^{n-1} b_k \\ &= \sum_{k=m}^n a_k + \sum_{k=m}^n b_k. \quad \square \end{aligned}$$

**Korollar 6.5.** Sei  $R$  ein Ring und  $c, a_k \in R$ . Sei  $c$  eine Konstante. Es gilt

$$\sum_{k=m}^n c a_k = c \sum_{k=m}^n a_k.$$

**Beweis.** Induktion über  $n$ . Induktionsanfang:

$$\sum_{k=m}^{m-1} c a_k = 0 = c \cdot 0 = c \sum_{k=m}^{m-1} a_k.$$

Induktionsschritt:

$$\sum_{k=m}^n c a_k = c a_n + \sum_{k=m}^{n-1} c a_k \stackrel{\text{IV}}{=} c a_n + c \sum_{k=m}^{n-1} a_k = c \left( a_n + \sum_{k=m}^{n-1} a_k \right) = c \sum_{k=m}^n a_k. \quad \square$$

**Korollar 6.6 (Aufteilung einer Summe).** Für  $m \leq p \leq n$  gilt

$$\sum_{k=m}^n a_k = \sum_{k=m}^{p-1} a_k + \sum_{k=p}^n a_k.$$

**Beweis.** Induktion über  $n$ . Im Induktionsanfang ist  $n = p$  und folglich:

$$\sum_{k=m}^p a_k = \sum_{k=m}^{p-1} a_k + a_p = \sum_{k=m}^{p-1} a_k + \sum_{k=p}^p a_k.$$

Induktionsschritt:

$$\sum_{k=m}^n a_k = a_n + \sum_{k=m}^{n-1} a_k \stackrel{\text{IV}}{=} a_n + \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{n-1} a_k = \sum_{k=m}^{p-1} a_k + \sum_{k=p}^n a_k. \quad \square$$

# 7 Wahrscheinlichkeitsrechnung

## 7.1 Diskrete Wahrscheinlichkeitsräume

### Definition 7.1 (Diskreter Wahrscheinlichkeitsraum).

Sei  $\Omega$  eine höchstens abzählbare Menge. Das Paar  $(\Omega, P)$  nennt man diskreten Wahrscheinlichkeitsraum, wenn

$$P: 2^\Omega \rightarrow [0, 1], \quad P(A) := \sum_{\omega \in A} P(\{\omega\})$$

die Eigenschaft  $P(\Omega) = 1$  besitzt.

Bemerkung: Man schreibt auch  $P(\omega) := P(\{\omega\})$ .

### Definition 7.2 (Reelle Zufallsgröße).

Sei  $(\Omega, P)$  ein diskreter Wahrscheinlichkeitsraum. Eine Funktion  $X: \Omega \rightarrow \mathbb{R}$  nennt man Zufallsgröße. Die Verteilung von  $X$  ist definiert gemäß  $P_X(A) := P(X^{-1}(A))$ .

### Definition 7.3 (Bedingte Wahrscheinlichkeit).

Seien  $A, B \subseteq \Omega$  und sei  $B \neq \emptyset$ . Dann nennt man

$$P(A | B) := \frac{P(A \cap B)}{P(B)}$$

die bedingte Wahrscheinlichkeit von  $A$ , gegeben  $B$ .

**Lemma 7.1.** Seien  $A, B$  disjunkt. Seien die  $A_i$  disjunkt. Dann gilt

$$\begin{aligned} 1_{A \cup B} &= 1_A + 1_B, \\ 1_{\left\{ \bigcup_{i \in I} A_i \right\}} &= \sum_{i \in I} 1_{\{A_i\}}. \end{aligned}$$

**Beweis.** Es gilt

$$\begin{aligned} 1_{A \cup B}(\omega) &= [\omega \in A \cup B] = [\omega \in A \vee \omega \in B] \stackrel{(P)}{=} [\omega \in A \oplus \omega \in B] \\ &= [\omega \in A] + [\omega \in B] = 1_A(\omega) + 1_B(\omega). \end{aligned}$$

Die allgemeine Rechnung ist

$$1_{\left\{ \bigcup_{i \in I} A_i \right\}}(\omega) = [\omega \in \bigcup_{i \in I} A_i] = [\exists i \in I: \omega \in A_i] \stackrel{(P)}{=} \sum_{i \in I} [\omega \in A_i] = \sum_{i \in I} 1_{\{A_i\}}(\omega).$$

Gleichung (P) gilt hierbei laut Prämisse.  $\square$

**Korollar 7.2 (Additivität des Wahrscheinlichkeitsmaßes).**

Seien die  $A_i$  paarweise disjunkte Mengen. Dann gilt

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i).$$

Für zwei disjunkte Menge  $A, B$  gilt speziell

$$P(A \cup B) = P(A) + P(B).$$

**Beweis 1.** Weil  $A \cap B = \emptyset$  gilt, ist  $[\omega \in A \cup B] = [\omega \in A] + [\omega \in B]$ . Deshalb gilt

$$\begin{aligned} P(A \cup B) &= \sum_{\omega \in A \cup B} P(\{\omega\}) = \sum_{\omega \in \Omega} [\omega \in A \cup B] P(\{\omega\}) \\ &= \sum_{\omega \in \Omega} ([\omega \in A] + [\omega \in B]) P(\{\omega\}) = \sum_{\omega \in \Omega} [\omega \in A] P(\{\omega\}) + \sum_{\omega \in \Omega} [\omega \in B] P(\{\omega\}) \\ &= \sum_{\omega \in A} P(\{\omega\}) + \sum_{\omega \in B} P(\{\omega\}) = P(A) + P(B). \end{aligned}$$

Nun allgemein. Weil die  $A_i$  disjunkt sind, gilt  $[\omega \in \bigcup_{i \in I} A_i] = \sum_{i \in I} [\omega \in A_i]$ . Deshalb gilt

$$\begin{aligned} P\left(\bigcup_{i \in I} A_i\right) &= \sum_{\omega \in \Omega} [\omega \in \bigcup_{i \in I} A_i] P(\{\omega\}) = \sum_{\omega \in \Omega} \sum_{i \in I} [\omega \in A_i] P(\{\omega\}) \\ &= \sum_{i \in I} \sum_{\omega \in \Omega} [\omega \in A_i] P(\{\omega\}) = \sum_{i \in I} \sum_{\omega \in A_i} P(\{\omega\}) = \sum_{i \in I} P(A_i). \quad \square \end{aligned}$$

**Beweis 2.** Laut Korollar 7.11, Lemma 7.1 und Korollar 7.7 gilt

$$P(A \cup B) = E(1_{A \cup B}) = E(1_A + 1_B) = E(1_A) + E(1_B) = P(A) + P(B).$$

Die allgemeine Rechnung ist

$$P\left(\bigcup_{i \in I} A_i\right) = E(1_{\{\bigcup_{i \in I} A_i\}}) = E\left(\sum_{i \in I} 1_{\{A_i\}}\right) = \sum_{i \in I} E(1_{\{A_i\}}) = \sum_{i \in I} P(A_i).$$

**Korollar 7.3.** Sei  $(\Omega, P)$  ein diskreter Wahrscheinlichkeitsraum in Form von Def. 7.1. Der Raum ist in Form des Tripels  $(\Omega, 2^\Omega, P)$  ein Wahrscheinlichkeitsraum, denn das Maß  $P$  erfüllt die drei kolmogorowschen Axiome.

**Beweis.** Die ersten beiden Axiome,  $(\forall A: P(A) \geq 0)$  und  $P(\Omega) = 1$ , gelten per Definition. Das dritte Axiom, die Additivität, gilt laut Korollar 7.2. Dass es sich bei  $2^\Omega$  um eine sigma-Algebra handelt, ist kaum einer ausdrücklichen Erwähnung wert.  $\square$

**Satz 7.4 (Gesetz der totalen Wahrscheinlichkeit).** Sei  $Z$  eine Zerlegung der Ergebnismenge  $\Omega$  in paarweise disjunkte nichtleere Mengen  $B \in Z$ . Dann gilt

$$P(A) = \sum_{B \in Z} P(A | B)P(B).$$

**Beweis.** Es gilt

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P\left(A \cap \bigcup_{B \in Z} B\right) = P\left(\bigcup_{B \in Z} (A \cap B)\right) \\ &= \sum_{B \in Z} P(A \cap B) = \sum_{B \in Z} P(A | B)P(B). \quad \square \end{aligned}$$



**Korollar 7.5 (Gesetz der totalen Wahrscheinlichkeit für Zufallsgrößen).**

Seien  $X, Y: \Omega \rightarrow \Omega'$  Zufallsgrößen. Dann gilt

$$P(Y \in A) = \sum_{x \in X(\Omega)} P(Y \in A \mid X = x)P(X = x),$$

speziell

$$P(Y = y) = \sum_{x \in X(\Omega)} P(Y = y \mid X = x)P(X = x).$$

**Beweis.** Sei  $Z = X(\Omega)$ . Zunächst gilt

$$\Omega = X^{-1}(Z) = X^{-1}\left(\bigcup_{x \in Z} \{x\}\right) = \bigcup_{x \in Z} X^{-1}(x)$$

gemäß Satz 1.35 (preimg-dl), und gemäß Korollar 1.38 ist das eine Vereinigung nicht-leerer paarweise disjunkter Mengen. Laut dem Gesetz der totalen Wahrscheinlichkeit gilt daher

$$\begin{aligned} P(Y \in A) &= P(Y^{-1}(A)) = P(Y^{-1}(A) \cap \Omega) = \sum_{x \in Z} P(Y^{-1}(A) \mid X^{-1}(x))P(X^{-1}(x)) \\ &= \sum_{x \in Z} P(Y \in A \mid X = x)P(X = x). \quad \square \end{aligned}$$

**Definition 7.4 (Erwartungswert).**

Sei  $(\omega_k)$  eine beliebige Abzählung von  $\Omega$ . Ist die Reihe  $\sum_{k=0}^{|\Omega|} X(\omega_k)P(\{\omega_k\})$  absolut konvergent, dann nennt man

$$E(X) := \sum_{\omega \in \Omega} X(\omega)P(\{\omega\})$$

den Erwartungswert von  $X$ .

**Satz 7.6.** Es gilt

$$E(X) = \sum_{x \in X(\Omega)} xP(X^{-1}(x)) = \sum_{x \in X(\Omega)} xP(X = x).$$

**Beweis.** Zunächst gilt

$$\sum_{\substack{\omega \in \Omega \\ X(\omega)=x}} P(\omega) = P\left(\bigcup_{\substack{\omega \in \Omega \\ X(\omega)=x}} \{\omega\}\right) = P(\{\omega \in \Omega \mid X(\omega) = x\}) = P(X^{-1}(x)).$$

Da die Reihe zu  $E(X)$  nach Def. 7.4 absolut konvergent ist, darf sie beliebig umgeordnet werden und man bekommt

$$\begin{aligned} E(X) &= \sum_{\omega \in \Omega} X(\omega)P(\omega) = \sum_{x \in X(\Omega)} \sum_{\substack{\omega \in \Omega \\ X(\omega)=x}} xP(\omega) = \sum_{x \in X(\Omega)} x \sum_{\substack{\omega \in \Omega \\ X(\omega)=x}} P(\omega) \\ &= \sum_{x \in X(\Omega)} xP(X^{-1}(x)). \quad \square \end{aligned}$$

**Korollar 7.7.** Der Erwartungswertoperator ist ein lineares Funktional, das heißt, es gilt  $E(aX) = aE(X)$  und  $E(X + Y) = E(X) + E(Y)$ .

**Beweis.** Aufgrund der Konvergenz der Reihen gilt

$$E(aX) = \sum_{\omega \in \Omega} aX(\omega)P(\omega) = a \sum_{\omega \in \Omega} X(\omega)P(\omega) = aE(X)$$

und

$$\begin{aligned} E(X + Y) &= \sum_{\omega \in \Omega} (X(\omega) + Y(\omega))P(\omega) = \sum_{\omega \in \Omega} (X(\omega)P(\omega) + Y(\omega)P(\omega)) \\ &= \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega) = E(X) + E(Y). \quad \square \end{aligned}$$

**Korollar 7.8.** Ist  $X \leq Y$ , dann ist auch  $E(X) \leq E(Y)$ .

**Beweis.** Gemäß  $P(\omega) \geq 0$  ist

$$X \leq Y \iff X(\omega) \leq Y(\omega) \iff 0 \leq Y(\omega) - X(\omega) \iff 0 \leq (Y(\omega) - X(\omega))P(\omega).$$

Somit hat man

$$X \leq Y \implies 0 \leq E(Y - X) = \sum_{\omega \in \Omega} (Y(\omega) - X(\omega))P(\omega),$$

und gemäß Linearität daher

$$X \leq Y \implies 0 \leq E(Y - X) = E(Y) - E(X) \iff E(X) \leq E(Y). \quad \square$$

**Definition 7.5 (Unabhängige Ereignisse).**

Zwei Ereignisse  $A, B$  heißen unabhängig, falls  $P(A \cap B) = P(A)P(B)$ .

**Definition 7.6 (Unabhängige Zufallsgrößen).**

Zwei Zufallsgrößen  $X, Y: \Omega \rightarrow \mathbb{R}$  heißen unabhängig, wenn die Ereignisse  $\{X \in A\}$  und  $\{Y \in B\}$  für alle Mengen  $A, B \subseteq \mathbb{R}$  unabhängig sind.

**Satz 7.9.** Zwei Zufallsgrößen  $X, Y: \Omega \rightarrow \mathbb{R}$  sind genau dann unabhängig, wenn für alle  $x \in X(\Omega)$  und  $y \in Y(\Omega)$  gilt:

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

**Beweis.** Sind  $X, Y$  unabhängig, dann ist

$$\begin{aligned} P(X = x, Y = y) &= P(\{X \in \{x\}\} \cap \{Y \in \{y\}\}) = P(\{X \in \{x\}\})P(\{Y \in \{y\}\}) \\ &= P(X = x)P(Y = y). \end{aligned}$$

Umgekehrt gelte nun  $P(X = x, Y = y) = P(X = x)P(Y = y)$ , dann ist

$$\begin{aligned} P(\{X \in A\} \cap \{Y \in B\}) &= P\left(\bigcup_{x \in A} \{X = x\} \cap \bigcup_{y \in B} \{Y = y\}\right) \\ &= P\left(\bigcup_{x \in A} \bigcup_{y \in B} (\{X = x\} \cap \{Y = y\})\right) = \sum_{x \in A} \sum_{y \in B} P(\{X = x\} \cap \{Y = y\}) \\ &= \sum_{x \in A} \sum_{y \in B} P(X = x)P(Y = y) = \sum_{x \in A} P(X = x) \sum_{y \in B} P(Y = y) \\ &= P\left(\bigcup_{x \in A} \{X = x\}\right)P\left(\bigcup_{y \in B} \{Y = y\}\right) = P(X \in A)P(Y \in B). \quad \square \end{aligned}$$

**Definition 7.7 (Bedingter Erwartungswert).**

$$E(X | A) = \frac{E(1_A X)}{P(A)} = \frac{1}{P(A)} \sum_{\omega \in A} X(\omega) P(\{\omega\}).$$

**Satz 7.10.** Es gilt

$$E(X | A) = \frac{1}{P(A)} \sum_x x P(\{X = x\} \cap A) = \sum_x x P(X = x | A),$$

wobei sich die Summe über alle  $x \in X(\Omega)$  erstreckt.

**Beweis.** Man kann rechnen

$$\begin{aligned} E(1_A X) &= \sum_{\omega \in \Omega} 1_A(\omega) X(\omega) P(\{\omega\}) = \sum_x \sum_{\omega \in X^{-1}(x)} 1_A(\omega) X(\omega) P(\{\omega\}) \\ &= \sum_x x \sum_{\omega \in X^{-1}(x)} 1_A(\omega) P(\{\omega\}) = \sum_x x \sum_{\omega \in X^{-1}(x) \cap A} P(\{\omega\}) \\ &= \sum_x x P(X^{-1}(x) \cap A), \end{aligned}$$

wobei  $X^{-1}(x) = \{X = x\}$ .  $\square$

**Korollar 7.11.** Es gilt  $P(A) = E(1_A)$ , wobei  $1_A$  die Indikatorfunktion ist.

**Beweis.** Gemäß Definition des Erwartungswertes ist

$$E(1_A) = \sum_{\omega \in \Omega} 1_A(\omega) P(\{\omega\}) = \sum_{\omega \in A} P(\{\omega\}) = P(A). \quad \square$$

**Korollar 7.12.** Es gilt  $P(A | B) = E(1_A | B)$ , wobei  $1_A$  die Indikatorfunktion ist.

**Beweis.** Gemäß Definition 7.7 und Korollar 7.11 ist

$$E(1_A | B) = \frac{E(1_A 1_B)}{P(B)} = \frac{E(1_{A \cap B})}{P(B)} = \frac{P(A \cap B)}{P(B)} = P(A | B). \quad \square$$

**Korollar 7.13.** Der Erwartungswert ist der Schwerpunkt der Wahrscheinlichkeitsmassen.

**Beweis.** Der Schwerpunkt  $s$  ist charakterisiert durch die folgende Gleichgewichtsbedingung, die wie folgt äquivalent umgeformt werden darf:

$$\begin{aligned} \sum_{x \in X(\Omega)} (s - x) P(X = x) &= 0 \iff s \underbrace{\sum_{x \in X(\Omega)} P(X = x)}_{=1} - \sum_{x \in X(\Omega)} x P(X = x) = 0 \\ \iff s &= \sum_{x \in X(\Omega)} x P(X = x) \iff s = E(X). \quad \square \end{aligned}$$

## 7.2 Allgemeine Wahrscheinlichkeitsräume

**Korollar 7.14.** Der Erwartungswert ist der Schwerpunkt der Dichtefunktion.

**Beweis.** Der Schwerpunkt  $s$  ist charakterisiert durch die folgende Gleichgewichtsbedingung, die wie folgt äquivalent umgeformt werden darf:

$$\begin{aligned} \int_{\mathbb{R}} (s-x)f(x) dx &= 0 \iff s \underbrace{\int_{\mathbb{R}} f(x) dx}_{=1} - \int_{\mathbb{R}} xf(x) dx = 0 \\ \iff s &= \int_{\mathbb{R}} xf(x) dx = E(X). \quad \square \end{aligned}$$

**Satz 7.15.** Sei  $g: \mathbb{R} \rightarrow \mathbb{R}$  eine streng monotone Funktion. Seien  $X, Y$  Zufallsgrößen mit Dichten  $f_X, f_Y$ . Ist  $Y = g(X)$ , dann gilt

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}.$$

**Beweis.** Sei  $g$  streng monoton steigend. Dann kann man rechnen

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

Gemäß der Kettenregel findet man

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

Sei  $g$  nun streng monoton fallend. Dann kann man rechnen

$$F_Y(y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - P(X < g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

Entsprechend findet man

$$f_Y(y) = -\frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

Nun ist  $g'$  in beiden Fällen frei von Nullstellen. Demnach ist  $\text{sgn}(g'(x))$  konstant für alle  $x$  und wir haben allgemein

$$f_Y(y) = \text{sgn}(g'(x)) \frac{f_X(x)}{g'(x)} = \frac{f_X(x)}{|g'(x)|}$$

mit  $x = g^{-1}(y)$ .  $\square$

**Satz 7.16 (»LOTUS: Law of the unconscious statistician«).**

Ist  $g: \mathbb{R} \rightarrow \mathbb{R}$  streng monoton, dann muss gelten

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f_X(x) dx.$$

**Beweis.** Sei  $Y = g(X)$ . Mit Satz 7.15 und Substitution  $y = g(x)$  kann man rechnen

$$\begin{aligned} E(g(X)) &= E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{-\infty}^{\infty} y \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|} dy \\ &= \int_{g^{-1}(-\infty)}^{g^{-1}(\infty)} g(x) \frac{f_X(x)}{|g'(x)|} g'(x) dx \\ &= \text{sgn}(g') \int_{-\text{sgn}(g')\infty}^{\text{sgn}(g')\infty} g(x)f_X(x) dx = \int_{-\infty}^{\infty} g(x)f_X(x) dx. \quad \square \end{aligned}$$

## 7.3 Stochastische Prozesse

### 7.3.1 Markow-Prozesse mit endlichem Zustandsraum

**Definition 7.8 (Markow-Prozess).** Sei  $X(t)$  ein stochastischer Prozess mit  $t \geq 0$  und  $X(t) \in S$ , wobei  $S$  ein endlicher Zustandsraum ist. Der Prozess heißt Markow-Prozess, wenn die Markow-Eigenschaft

$$P(X(s+t) = j \mid \{X(s) = i\} \cap \bigcap_{0 \leq u < s} \{X(u) = x_u\}) = P(X(s+t) = j \mid X(s) = i)$$

erfüllt ist. Das heißt, für die bedingte Wahrscheinlichkeit, dass  $X$  in der Zukunft den Zustand  $j$  einnimmt, spielt nur der aktuelle Zustand  $i$  eine Rolle, nicht aber Zustände der Vergangenheit.

**Bemerkung.** Eine geläufige Notation ist  $P(A \mid B, C) := P(A \mid B \cap C)$ . Es gilt

$$P(A \mid B \cap C) = P_B(A \mid C) = P_C(A \mid B),$$

wobei  $P_B(M) := P(M \mid B)$  und  $P_C(M) := P(M \mid C)$ .

**Definition 7.9 (Homogener Markow-Prozess).** Ein Markow-Prozess  $X(t)$  heißt homogen, wenn die Gleichung

$$p_{ij}(t) := P(X(s+t) = j \mid X(s) = i) = P(X(t) = j \mid X(0) = i)$$

für jedes  $s$  erfüllt ist. Man nennt  $P = (p_{ij})$  die Übergangsmatrix.

**Satz 7.17 (Kolmogorow-Chapman-Gleichung).**

Für die Übergangsmatrix eines homogenen Markow-Prozesses gilt

$$P(s+t) = P(s)P(t).$$

**Beweis.** Wir schreiben kurz  $P_i(X(t) = j) := P(X(t) = j \mid X(0) = i)$ . Gemäß dem Gesetz der totalen Wahrscheinlichkeit 7.5 gilt

$$p_{ij}(s+t) = P_i(X(s+t) = j) = \sum_{k \in S} P_i(X(s) = k) P_i(X(s+t) = j \mid X(s) = k).$$

Aufgrund der Markow-Eigenschaft und der Homogenität gilt nun

$$\begin{aligned} P_i(X(s+t) = j \mid X(s) = k) &= P(X(s+t) = j \mid X(s) = k, X(0) = i) \\ &= P(X(s+t) = j \mid X(s) = k) = P(X(t) = j \mid X(0) = k) = P_k(X(t) = j) = p_{kj}(t). \end{aligned}$$

Man erhält die Matrizenmultiplikation

$$p_{ij}(s+t) = \sum_{k \in S} p_{ik}(s) p_{kj}(t), \text{ kurz } P(s+t) = P(s)P(t). \quad \square$$

**Definition 7.10 (Intensitätsmatrix).**

Ist  $P(t)$  die Übergangsmatrix eines homogenen Markow-Prozesses, nennt man

$$Q := P'(0) = \lim_{h \rightarrow 0} \frac{P(h) - E}{h}$$

die Intensitätsmatrix, wobei mit  $E$  die Einheitsmatrix gemeint ist.

**Bemerkung.** Weil  $P(t)$  für  $t \geq 0$  definiert ist, stimmt der gewöhnliche Grenzwert für  $h \rightarrow 0$  mit dem rechtsseitigen Grenzwert für  $h \searrow 0$  überein.

**Satz 7.18.** Der Grenzwert in Def. 7.10 existiert.

**Korollar 7.19.** Für die Übergangsmatrix gilt

$$P(h) = E + hQ + o(h), \text{ bzw. } p_{ij}(h) = \delta_{ij} + hq_{ij} + o(h).$$

**Beweis.** Weil  $P$  laut Satz 7.18 an der Stelle 0 differenzierbar ist, besteht die Linearisierung

$$P(h) = P(0) + hP'(0) + o(h) = E + hQ + o(h). \quad \square$$

**Satz 7.20 (Kolmogorowsche Vorwärtsgleichung).**

Die Übergangsmatrix erfüllt das Anfangswertproblem

$$P'(t) = P(t)Q, \quad P(0) = E.$$

**Beweis.** Laut der Chapman-Kolmogorow-Gleichung ist

$$P(t) = P(0 + t) = P(0)P(t).$$

Ergo muss  $P(0) = E$  sein. Laut der Chapman-Kolmogorow-Gleichung, Def. 7.10 und Satz 7.18 gilt zudem

$$P'(t) = \lim_{h \rightarrow 0} \frac{P(t+h) - P(t)}{h} = \lim_{h \rightarrow 0} \frac{P(t)P(h) - P(t)}{h} = P(t) \lim_{h \rightarrow 0} \frac{P(h) - E}{h} = P(t)Q. \quad \square$$

**Satz 7.21.** Die kolmogorowsche Vorwärtsgleichung besitzt die eindeutige Lösung  $P(t) = e^{tQ}$ , wobei mit  $e^A = \exp(A)$  das Matrixexponential gemeint ist.

**Beweis.** Analog zur gewöhnlichen Exponentialfunktion gilt

$$\begin{aligned} \frac{d}{dt} e^{tQ} &= \frac{d}{dt} \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!} = \sum_{k=1}^{\infty} \frac{d}{dt} \left( \frac{t^k Q^k}{k!} \right) = \sum_{k=1}^{\infty} k \frac{t^{k-1} Q^k}{k!} \\ &= \sum_{k=0}^{\infty} (k+1) \frac{t^k Q^{k+1}}{(k+1)!} = \sum_{k=0}^{\infty} \frac{t^k Q^{k+1}}{k!} = e^{tQ} Q = Q e^{tQ}. \end{aligned}$$

Daher ist  $e^{tQ}$  eine Lösung der Differentialgleichung und wegen  $\exp(0) = E$  des Anfangswertproblems. Angenommen, es gibt eine weitere Lösung  $P$ , dann kann man  $P(t) = R(t)e^{tQ}$  schreiben mit  $R(t) := P(t)e^{-tQ}$ . Da  $P$  die Differentialgleichung erfüllen soll, muss gelten

$$0 = P'(t) - P(t)Q = R'(t)e^{tQ} + R(t)e^{tQ}Q - R(t)e^{tQ}Q = R'(t)e^{tQ}.$$

Multipliziert man beide Seiten der Gleichung mit  $e^{-tQ}$ , ergibt sich  $R'(t) = 0$ . Damit ist  $R$  konstant mit  $R(t) = R(0) = E$  laut Anfangsbedingung. Es gibt also keine weitere Lösung außer  $P(t) = e^{tQ}$ .  $\square$

**Korollar 7.22.** Für die Intensitätsmatrix gilt  $\sum_{j \in S} q_{ij} = 0$ .

**Beweis.** Weil  $p_{ij}(t)$  eine Wahrscheinlichkeitsfunktion in  $j$  ist, muss  $\sum_{j \in S} p_{ij}(t) = 1$  gelten. Demzufolge ist

$$0 = \frac{d}{dt} 1 = \sum_{j \in S} p'_{ij}(t) = \sum_{j \in S} q_{ij}. \quad \square$$

**Korollar 7.23 (Mastergleichung).**

Die Übergangsmatrix erfüllt die Mastergleichung

$$p'_{ij}(t) = \sum_{k \neq j} (p_{ik}(t)q_{kj} - p_{ij}(t)q_{jk}).$$

**Beweis.** Da die Zeilensummen laut Korollar 7.22 verschwinden, gilt  $q_{jj} = -\sum_{k \neq j} q_{jk}$ . Damit erhält man die Umformung

$$\begin{aligned} p'_{ij}(t) &= \sum_k p_{ik}(t)q_{kj} = \sum_{k \neq j} p_{ik}(t)q_{kj} + p_{ij}(t)q_{jj} = \sum_{k \neq j} p_{ik}(t)q_{kj} - \sum_{k \neq j} p_{ij}(t)q_{jk} \\ &= \sum_{k \neq j} (p_{ik}(t)q_{kj} - p_{ij}(t)q_{jk}). \quad \square \end{aligned}$$

**Bemerkung.** Man kann die Mastergleichung alternativ in der Form

$$p'_j(t) = \sum_{k \neq j} (w_{jk}p_k(t) - w_{kj}p_j(t))$$

schreiben, wobei  $p_j(t) := p_{ij}(t)$  und  $w_{ij} := q_{ji}$  ist. Zudem kann man  $w_{kk} := 0$  für jedes  $k$  setzen, weil die Hauptdiagonale nicht in die Gleichung eingeht und redundante Information enthält.





# Index

- Abbildung, 15
- Ableitung, 34
- abzählbares Auswahlaxiom, 21
- adjungierte Matrix, 45
- algebraische Zahlen
  - Kardinalität, 22
- Assoziativgesetz
  - Mengen, boolesche Algebra, 12
- Aussagenlogik, 5
- Auswahlaxiom
  - abzählbares, 21
- Banach
  - Fixpunktsatz von, 38
- bedingte Wahrscheinlichkeit, 63
- bedingter Erwartungswert, 67
- beschränkte Folge, 27
- Bildmenge, 15
- differenzierbar, 34
- Distributivgesetz
  - boolesche Algebra, 6
  - Urbildoperation, 16
- Dreiecksungleichung, 41
  - umgekehrte, 42
- Epsilon-Umgebung, 27
- Erwartungswert, 65
  - bedingter, 67
- Fixpunkt-Iteration, 38
- Fixpunktgleichung, 30
- Fixpunktsatz von Banach, 38
- folgenstetig, 30
- Gesetz der totalen Wahrsch., 64
- Gleichheit
  - von Abbildungen, 15
  - von Mengen, 11
- gleichmächtig, 21
- Grenzwert, 27
- Grenzwertsätze, 29
- homogener Markow-Prozess, 69
- Indikatorfunktion, 21
- Injektion, 15
- inverse Matrix, 45
- kartesisches Produkt, 11
- Kolmogorow-Chapman-Gleichung, 69
- Kommutativgesetz
  - boolesche Algebra, 5
  - Mengen, boolesche Algebra, 11
- Komposition, 15
- konjugierte Matrix, 45
- Kontraktion, 38
- konvergente Folge, 27
- Markow-Prozess, 69
- Mengenlehre, 11
- metrischer Raum, 41
- Newton-Verfahren, 38
- normierter Raum, 41
- offene Epsilon-Umgebung, 27
- offener Kern, 39
- Prädikatenlogik, 7
- Produktregel, 34
- Relation, 14
- Schnittmenge, 11
  - stetig
    - folgenstetig, 30
  - stochastischer Prozess, 69
- Surjektion, 15
- Teilmenge, 11
- transponierte Matrix, 45
- Umgebungsfilter, 39
- umgekehrte Dreiecksungleichung, 42
- unabhängige Ereignisse, 66
- unabhängige Zufallsgrößen, 66
- unitäre Matrix, 47
- Urbildmenge, 15
- Vereinigungsmenge, 11
- Verkettung, 15
- Zufallsgröße, 63