Beweisarchiv

Mai 2022

Inhaltsverzeichnis

1 Grundlagen **5**
- 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 . . . 15
  - 1.4.1 Allgemeine Gesetzmäßigkeiten . . . 15
  - 1.4.2 Äquivalenzrelationen . . . 15
- 1.5 Abbildungen . . . 17
  - 1.5.1 Definitionen . . . 17
  - 1.5.2 Grundlagen . . . 17
  - 1.5.3 Kardinalzahlen . . . 24

2 Analysis **29**
- 2.1 Folgen . . . 29
  - 2.1.1 Konvergenz . . . 29
  - 2.1.2 Wachstum und Landau-Symbole . . . 32
- 2.2 Stetige Funktionen . . . 33
- 2.3 Differentialrechnung . . . 36
  - 2.3.1 Ableitungsregeln . . . 36
  - 2.3.2 Glatte Funktionen . . . 38
  - 2.3.3 Richtungsableitung . . . 38
- 2.4 Fixpunkt-Iterationen . . . 40

3 Topologie **41**
- 3.1 Grundbegriffe . . . 41
  - 3.1.1 Definitionen . . . 41
  - 3.1.2 Stetige Abbildungen . . . 42
  - 3.1.3 Elementares . . . 43
  - 3.1.4 Zusammenhang . . . 44
- 3.2 Metrische Räume . . . 46
  - 3.2.1 Metrische Räume . . . 46
  - 3.2.2 Normierte Räume . . . 46
  - 3.2.3 Homöomorphien . . . 47
- 3.3 Übungen . . . 47

4 Lineare Algebra **49**
- 4.1 Matrizen . . . 49
  - 4.1.1 Definitionen . . . 49
  - 4.1.2 Rechenregeln . . . 49
  - 4.1.3 Rechenregeln für komplexe Matrizen . . . 50
- 4.2 Eigenwerte . . . 50
  - 4.2.1 Quadratische Matrizen . . . 52
- 4.3 Lineare Abbildungen . . . 53

4.4 Bilinearformen 55
4.5 Euklidische Geometrie 56

5 Algebra **59**
5.1 Gruppentheorie 59
5.1.1 Grundlagen 59
5.2 Ringtheorie 61
5.2.1 Grundlagen 61
5.2.2 Ringhomomorphismen 63
5.3 Polynomringe 63
5.3.1 Einsetzungshomomorphismus 63
5.4 Körper 64
5.4.1 Geordnete Körper 64
5.5 Formale Potenzreihen 66
5.6 Zahlenbereiche 68
5.6.1 Die natürlichen Zahlen 68
5.6.2 Die ganzen Zahlen 70
5.6.3 Die rationalen Zahlen 72

6 Kombinatorik **75**
6.1 Endliche Mengen 75
6.1.1 Indikatorfunktion 75
6.1.2 Endliche Abbildungen 76
6.2 Endliche Summen 79
6.2.1 Allgemeine Regeln 79
6.2.2 Klassische Partialsummen 86
6.3 Funktionen 87
6.3.1 Floor und Ceil 87
6.3.2 Faktorielle 88
6.3.3 Binomialkoeffizient 89

7 Wahrscheinlichkeitsrechnung **91**
7.1 Diskrete Wahrscheinlichkeitsräume 91
7.2 Allgemeine Wahrscheinlichkeitsräume 96
7.3 Stochastische Prozesse 97
7.3.1 Markow-Prozesse mit endlichem Zustandsraum 97
7.4 Mathematische Statistik 99
7.4.1 Schätzfunktionen 99

8 Zahlentheorie **101**
8.1 Kongruenzen und Teilbarkeit 101
8.2 Primzahlen 103

1 Grundlagen

1.1 Aussagenlogik

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

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:

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

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 $\varphi :\equiv \neg A \vee B$. $\square$

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:

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

Programmterm:

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

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

$$\frac{A \vee B \quad \dfrac{\overline{A}^{\,1}}{B \vee A} \quad \dfrac{\overline{B}^{\,1}}{B \vee A}}{B \vee A}\,{}_{1}$$

Programmterm:

$$A + B \to B + A,\ s \mapsto \textbf{match}\ s \begin{cases} \mathrm{inl}(a) \mapsto \mathrm{inr}(a), \\ \mathrm{inr}(b) \mapsto \mathrm{inl}(b). \end{cases}$$

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

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

$$A \wedge (B \vee C) \iff A \wedge B \vee A \wedge C, \tag{1.3}$$
$$A \vee (B \wedge C) \iff (A \vee B) \wedge (A \vee C). \tag{1.4}$$

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

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

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

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

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

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

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

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

Sämtliche Teilaussagen sind bewiesen. □

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 $0 \to A$ 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 $\neg A \vee A \vdash (\neg\neg A \Rightarrow A)$. Gemäß Axiom 1.4 (PE) existiert ein Zeuge $\mathrm{ex}(A)$ für $0 \to A$. Damit lässt sich der Programmterm

$$(A \to 0) + A \to (((A \to 0) \to 0) \to A),\ s \mapsto \textbf{match}\ s \begin{cases} \mathrm{inl}\, f \mapsto g \mapsto \mathrm{ex}(A)(g(f)), \\ \mathrm{inr}\, a \mapsto g \mapsto a. \end{cases}$$

konstruieren. □

1.2 Prädikatenlogik

Definition 1.1 (bounded: beschränkte Quantifizierung).

$$(\forall x{\in}M\colon P(x)) \;:\Longleftrightarrow\; \forall x\colon (x \in M \implies P(x)), \tag{1.5}$$
$$(\exists x{\in}M\colon P(x)) \;:\Longleftrightarrow\; \exists x\colon (x \in M \land P(x)). \tag{1.6}$$

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

$$A \land (\exists x\colon P(x)) \iff \exists x\colon (A \land P(x)), \tag{1.7}$$
$$A \lor (\forall x\colon P(x)) \iff \forall x\colon (A \lor P(x)). \tag{1.8}$$

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

$$\dfrac{\exists x\colon P(x) \quad \dfrac{\dfrac{A \quad \overline{P(a)}^{\,1}}{A \land P(a)}}{\exists x\colon (A \land P(x))}}{\exists x\colon (A \land P(x))}\,{}_1 \qquad \dfrac{\exists x\colon (A \land P(x)) \quad \dfrac{\dfrac{\overline{A \land P(a)}^{\,1}}{A} \quad \dfrac{\overline{A \land P(a)}^{\,1}}{P(a)}}{\dfrac{}{\exists x\colon P(x)}}}{A \land \exists x\colon P(x)}\,{}_1$$

Zur zweiten Äquivalenz:

$$\dfrac{A \lor (\forall x\colon P(x)) \quad \dfrac{\dfrac{\overline{A}^{\,1}}{A \lor P(x)}}{\forall x\colon (A \lor P(x))} \quad \dfrac{\dfrac{\overline{\forall x\colon P(x)}^{\,1}}{P(x)}}{A \lor P(x)}}{\forall x\colon (A \lor P(x))}\,{}_1 \qquad \dfrac{\forall x\colon (A \lor P(x)) \quad \dfrac{\overline{A}^{\,1}}{A \lor \forall x\colon P(x)} \quad \dfrac{\dfrac{\overline{P(x)}^{\,1}}{\forall x\colon P(x)}}{A \lor \forall x\colon P(x)}}{A \lor \forall x\colon P(x)}\,{}_1$$

Die linken Schemata zeigen jeweils die Implikation von links nach rechts, die rechten jeweils die Implikation von rechts nach links. □

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

$$(\exists x\colon P(x) \lor Q(x)) \iff (\exists x\colon P(x)) \lor (\exists x\colon Q(x)).$$

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

$$\dfrac{\exists x\colon P(x) \lor Q(x) \quad \dfrac{\overline{P(a) \lor Q(a)}^{\,1} \quad \dfrac{\dfrac{\overline{P(a)}^{\,2}}{\exists x\colon P(x)}}{(\exists x\colon P(x)) \lor (\exists x\colon Q(x))} \quad \dfrac{\dfrac{\overline{Q(a)}^{\,2}}{\exists x\colon Q(x)}}{(\exists x\colon P(x)) \lor (\exists x\colon Q(x))}}{(\exists x\colon P(x)) \lor (\exists x\colon Q(x))}\,{}_1}{(\exists x\colon P(x)) \lor (\exists x\colon Q(x))}\,{}_2$$

Die Implikation von rechts nach links:

$$\dfrac{(\exists x\colon P(x)) \lor (\exists x\colon Q(x)) \quad \dfrac{\dfrac{\overline{\exists x\colon P(x)}^{\,1} \quad \dfrac{\overline{P(a)}^{\,2}}{P(a) \lor Q(a)}}{\exists x\colon P(x) \lor Q(x)}\,{}_2 \quad \dfrac{\overline{\exists x\colon Q(x)}^{\,1} \quad \dfrac{\overline{Q(b)}^{\,3}}{P(b) \lor Q(b)}}{\exists x\colon P(x) \lor Q(x)}\,{}_3}{\exists x\colon P(x) \lor Q(x)}\,{}_1$$

Man beachte, dass die Zeugen a, b unterschiedlich sein können. □

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

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

Beweis (Natürliches Schließen). Schematisch:

$$\dfrac{\exists x\colon P(x) \wedge Q(x) \qquad \dfrac{\dfrac{\dfrac{\overline{P(a) \wedge Q(a)}^{1}}{P(a)}}{\exists x\colon P(x)} \quad \dfrac{\dfrac{\overline{P(a) \wedge Q(a)}^{1}}{Q(a)}}{\exists x\colon Q(x)}}{(\exists x\colon P(x)) \wedge (\exists x\colon Q(x))}^{1}}{(\exists x\colon P(x)) \wedge (\exists x\colon Q(x))}$$

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

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

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

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

$$\neg \exists x\colon P(x) \equiv (\exists x\colon P(x)) \Rightarrow 0 \equiv \forall x\colon (P(x) \Rightarrow 0) \equiv \forall x\colon \neg P(x).$$ □

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

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

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

$$\neg \forall x\colon P(x) \equiv \neg \forall x\colon \neg\neg P(x) \equiv \neg\neg \exists x\colon \neg P(x) \equiv \exists x\colon \neg P(x).$$ □

Satz 1.13. Es gilt:

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

Beweis 1 (Natürliches Schließen). Schematisch:

$$\dfrac{\dfrac{\exists x\colon P(x) \qquad \dfrac{\dfrac{\forall x\colon (P(x) \implies A)}{P(a) \implies A} \quad \overline{P(a)}^{2}}{A}^{2}}{A}}{(\exists x\colon P(x)) \implies A}^{1} \qquad \dfrac{(\exists x\colon P(x)) \implies A \qquad \dfrac{\overline{P(x)}^{1}}{\exists x\colon P(x)}}{A}}{\dfrac{P(x) \implies A}{\forall x\colon (P(x) \implies A)}}^{1}$$

Die linke Seite zeigt die Implikation von links nach rechts, die rechte die Implikation von rechts nach links. □

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

$$\begin{aligned}\forall x\colon (P(x) \Rightarrow A) &\equiv \forall x\colon (\neg P(x) \vee A) \equiv (\forall x\colon \neg P(x)) \vee A \\ &\equiv \neg(\exists x\colon P(x)) \vee A \equiv (\exists x\colon P(x)) \Rightarrow A.\end{aligned}$$ □

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

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

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

$$\dfrac{\exists x\colon \exists y\colon P(x,y) \quad \dfrac{\overline{P(a,b)}^{\,1}}{\exists y\colon \exists x\colon P(x,y)}}{\exists y\colon \exists x\colon P(x,y)}\,{}_{1}$$

Die Implikation von rechts nach links geht analog. □

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

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

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

$$\dfrac{\dfrac{\dfrac{\forall x\colon \forall y\colon P(x,y)}{P(x,y)}}{\forall x\colon P(x,y)}}{\forall y\colon \forall x\colon P(x,y)}$$

Die Implikation von rechts nach links geht analog. □

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

$$A \wedge (\exists x{\in}M\colon P(x)) \iff (\exists x{\in}M\colon A \wedge P(x)), \tag{1.9}$$
$$A \vee (\forall x{\in}M\colon P(x)) \iff (\forall x{\in}M\colon A \vee P(x)). \tag{1.10}$$

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

$$A \wedge \exists x{\in}M\colon P(x) \equiv A \wedge \exists x\colon x \in M \wedge P(x) \equiv \exists x\colon A \wedge x \in M \wedge P(x)$$
$$\equiv \exists x\colon x \in M \wedge A \wedge P(x) \equiv \exists x{\in}M\colon A \wedge P(x).$$

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

$$A \vee \forall x{\in}M\colon P(x) \equiv A \vee \forall x\colon (x \in M \Rightarrow P(x)) \equiv A \vee \forall x\colon x \notin M \vee P(x)$$
$$\equiv \forall x\colon A \vee x \notin M \vee P(x) \equiv \forall x\colon (x \in M \Rightarrow A \vee P(x))$$
$$\equiv \forall x{\in}M\colon A \vee P(x).\ \square$$

Satz 1.17. Es gilt:

$$(\exists x{\in}A\colon \exists y{\in}B\colon P(x,y)) \iff (\exists y{\in}B\colon \exists x{\in}A\colon P(x,y)).$$

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

$$\exists x{\in}A\colon \exists y{\in}B\colon P(x,y) \equiv \exists x\colon x \in A \wedge \exists y\colon y \in B \wedge P(x,y)$$
$$\equiv \exists x\colon \exists y\colon x \in A \wedge y \in B \wedge P(x,y) \equiv \exists y\colon \exists x\colon y \in B \wedge x \in A \wedge P(x,y)$$
$$\equiv \exists y\colon y \in B \wedge \exists x\colon x \in A \wedge P(x,y) \equiv \exists y{\in}B\colon \exists x{\in}A\colon P(x,y).\ \square$$

Satz 1.18. Es gilt:

$$(\forall x{\in}A\colon \forall y{\in}B\colon P(x,y)) \iff (\forall y{\in}B\colon \forall x{\in}A\colon P(x,y)).$$

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

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

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

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

Beweis. Nach 1.8 (general-dl) gilt:

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

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

Satz 1.20. Es gilt

$$(\exists x{\in}M\colon P) \iff (M \neq \varnothing) \wedge P.$$

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

$$\exists x{\in}M\colon P \equiv \exists x\colon (x \in M \wedge P) \equiv (\exists x\colon x \in M) \wedge P \equiv (M \neq \varnothing) \wedge P. \qquad \square$$

1.3 Mengenlehre

1.3.1 Definitionen

Definition 1.2 (seteq: Gleichheit von Mengen).

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

Definition 1.3 (subseteq: Teilmenge).

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

Definition 1.4 (filter: beschreibende Angabe).

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

Definition 1.5 (cap: Schnitt).

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

Definition 1.6 (cup: Vereinigung).

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

Definition 1.7 (intersection: Schnitt).

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

Definition 1.8 (union: Vereinigung).

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

Definition 1.9 (cart: kartesisches Produkt).

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

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:

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

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

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

Für die Vereinigung ist das analog. □

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:

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

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

$$x \in A \cap (B \cap C) \iff x \in A \land x \in B \cap C \iff x \in A \land (x \in B \land x \in C)$$
$$\iff (x \in A \land x \in B) \land x \in C \iff x \in A \cap B \land x \in C \iff x \in (A \cap B) \cap C.$$

Für die Vereinigung ist das analog. □

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

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

$$\varphi \land \psi \Rightarrow \varphi \equiv \neg(\varphi \land \psi) \lor \varphi \equiv \neg\varphi \lor \neg\psi \lor \varphi \equiv 1 \lor \neg\psi \equiv 1.$$

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

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

$$x \in A \Rightarrow x \in B \iff x \in A \Rightarrow x \in A \land x \in B.$$

Die Formel $\varphi \Rightarrow \psi \iff \varphi \Rightarrow \varphi \land \psi$ ist aber tautologisch, denn

$$\varphi \Rightarrow \varphi \land \psi \equiv \neg\varphi \lor (\varphi \land \psi) \equiv (\neg\varphi \lor \varphi) \land (\neg\varphi \land \psi) \equiv 1 \land (\varphi \Rightarrow \psi) \equiv \varphi \Rightarrow \psi.$$

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

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

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

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

Die andere Implikation bringen wir zunächst in ihre Kontraposition:

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

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. □

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

Beweis. Es gilt:

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

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

Satz 1.27. Es gilt:

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

Satz 1.28. Es gilt:

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

Beweis. Nach Def. 1.9 (cart) gilt:

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

Unter doppelter Anwendung von Satz 1.13 gilt weiter:

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

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

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

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

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

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

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

ergibt. □

Satz 1.29. Es gilt:

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

Beweis. Nach Def. 1.9 (cart) gilt:

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

Nun gilt aber ganz offensichtlich

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

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

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:

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

Nach Def. 1.2 (seteq) folgt die Behauptung. □

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:

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

Nach Def. 1.2 (seteq) folgt die Behauptung. □

Korollar 1.32. Die Relation *Teilmenge von* ist eine Partialordnung. Im Einzelnen gilt

(1)	$A\subseteq A$,	(Reflexivität)
(2)	$A\subseteq B \wedge B\subseteq A \implies A=B$,	(Antisymmetrie)
(3)	$A\subseteq B \wedge B\subseteq C \implies A\subseteq C$.	(Transitivität)

Beweis. Jeweils Def. 1.3 nutzen.

Zu (1). Die Aussage $A\subseteq A$ ist äquivalent zu $\forall x : (x\in A \Rightarrow x\in A)$. Eine Prämisse impliziert sich im Allgemeinen selbst.

Zu (2). Es findet sich die äquivalente Umformung

$$A\subseteq B \wedge B\subseteq A \iff (\forall x : x\in A \Rightarrow x\in B) \wedge (\forall x : x\in B \Rightarrow x\in A)$$
$$\iff (\forall x : (x\in A \Rightarrow x\in B) \wedge (x\in B \Rightarrow x\in A)) \iff (\forall x : (x\in A \Leftrightarrow x\in B))$$
$$\iff A=B.$$

Zu (3). Zu zeigen ist $\forall x : (x\in A \Rightarrow x\in C)$. Sei $x\in A$ fest, aber beliebig. Wegen $A\subseteq B$ muss $x\in B$ sein. Wegen $B\subseteq C$ muss infolge $x\in C$ sein. □

Korollar 1.33. Die Aussage $A\subseteq B$ ist äquivalent zu $1_A \leq 1_B$.

Beweis. Es gelte $A\subseteq B$. Um $\forall x : 1_A(x)\leq 1_B(x)$ zu zeigen, wird eine Fallunterscheidung in drei Fälle vorgenommen. Sei $x\notin B$. Dann ist $x\notin A$ und daher $1_A(x)=0$ und $1_B(x)=0$, womit $1_A(x)\leq 1_B(x)$ gilt. Sei $x\in A$. Dann ist $x\in B$, und daher $1_A(x)=1$ und $1_B(x)=1$, womit $1_A(x)\leq 1_B(x)$ gilt. Sei $x\notin A$, aber $x\in B$. Dann ist $1_A(x)=0$ und $1_B(x)=1$, womit $1_A(x)\leq 1_B(x)$ gilt.

Es gelte $\forall x : 1_A(x)\leq 1_B(x)$. Sei $x\in A$ fest, aber beliebig. Wegen $1_A(x)=1$ ist $1\leq 1_B(x)$. Weil somit $1_B(x)\neq 0$ ist, verbleibt nur noch $1_B(x)=1$, was gleichbedeutend mit $x\in B$ ist. □

1.4 Relationen

1.4.1 Allgemeine Gesetzmäßigkeiten

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

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

als Bildmenge von M unter R bezeichnet.

Korollar 1.34. 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

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

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. □

Korollar 1.35. 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

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

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

1.4.2 Äquivalenzrelationen

Definition 1.12 (Äquivalenzrelation).
Sei M eine Menge. Man nennt $R \subseteq M \times M$, notiert als $R(x, y) = (x \sim y)$, eine Äquivalenzrelation auf M, wenn für alle $x, y, z \in M$ erfüllt ist:

$$\begin{aligned} &x \sim x, && \text{(Reflexivität)} \\ &x \sim y \implies y \sim x, && \text{(Symmetrie)} \\ &x \sim y \wedge y \sim z \implies x \sim y. && \text{(Transitivität)} \end{aligned}$$

Definition 1.13 (Äquivalenzklasse).
Sei M eine Menge und $x \sim y$ eine Äquivalenzrelation für $x, y \in M$. Die Menge

$$[a] := \{x \in M \mid x \sim a\}$$

nennt man Äquivalenzklasse zum Repräsentanten $a \in M$.

Definition 1.14 (Quotientenmenge).
Die Menge $M/\sim := \{A \mid \exists a \in M : A = [a]\}$ aller Äquivalenzklassen heißt Quotientenmenge von M bezüglich der Äquivalenzlation $\sim$.

Definition 1.15 (Quotientenabbildung).
Die Abbildung $\pi : M \to M/\sim$ mit $\pi(x) := [x]$ heißt Quotientenabbildung.

Korollar 1.36 (Äquivalenzrelation induziert disjunkte Zerlegung).
Eine Menge wird durch eine auf ihr definierte Äquivalenzrelation in paarweise disjunkte Äquivalenzklassen zerlegt.

Beweis. Es ist zu zeigen, dass zwei unterschiedliche Äquivalenzklassen kein Element gemeinsam haben. Wir zeigen die Kontraposition, dass die Existenz eines x mit $x \in [a]$ und $x \in [b]$ bereits $[a] = [b]$ impliziert. Laut Prämisse ist $x \sim a$ und $x \sim b$, und wegen der Transitivtät infolge $a \sim b$, was äquivalent zu $[a] = [b]$ ist.

Zu bestätigen verbleibt noch, dass die Quotientenabbildung eine surjektive ist. Dies ist wahr, weil $M/\sim$ gerade so definiert ist, dass direkt $M/\sim = \pi(M)$ gilt. □

Korollar 1.37 (Disjunkte Zerlegung induziert Äquivalenzrelation).
Sei M eine Menge. Die Familie (A_k) der $A_k \subseteq M$ sei eine Zerlegung von M in paarweise disjunkte Mengen. Dann definiert

$$x \sim y \ :\Longleftrightarrow\ \exists k : x \in A_k \wedge y \in A_k$$

eine Äquivalenzrelation.

Beweis. Da die A_k die Menge M überdecken, muss für jedes $x \in M$ ein k mit $x \in A_k$ existieren, womit die Reflexivität $x \sim x$ erfüllt ist.

Die Symmetrie folgt unmittelbar aus der Kommutativität der Konjunktion.

Zur Transitivität. Seien x, y, z fest, aber beliebig. Zudem seien die Prämissen $x \sim y$ und $y \sim z$ erfüllt. Wir haben daher einen Zeugen i mit $x \in A_i$ und $y \in A_i$ und einen Zeugen j mit $y \in A_j$ und $z \in A_j$. Infolge ist $y \in A_i \cap A_j$. Wegen $A_i \cap A_j = \emptyset$ für $i \neq j$ muss $i = j$ sein. Deshalb ist i ein Zeuge für $\exists i : x \in A_i \wedge z \in A_i$, womit $x \sim z$ gilt. □

Satz 1.38 (Charakterisierung von Äquivalenzklassen).
Sei $\sim$ eine Äquivalenzrelation auf der Menge M. Eine Teilmenge $A \subseteq M$ ist genau dann eine Äquivalenzklasse, wenn

(1) $A \neq \emptyset$,
(2) $x, y \in A \implies x \sim y$,
(3) $x \in A \wedge y \in M \wedge x \sim y \implies y \in A$.

Beweis. Sei A eine Äquivalenzklasse. Dann existiert definitionsgemäß ein a, so dass $A = [a]$ gilt. Ergo ist $a \in A$, womit $A \neq \emptyset$ sein muss. Mit $x, y \in A$ ergibt sich $[x] = [y]$, was äquivalent zu $x \sim y$ ist. Sei nun $x \in A$ und y irgendein Element in M mit $x \sim y$. Dies bedeutet $A = [x] = [y]$, womit $y \in A$ sein muss.

Umgekehrt seien die drei Eigenschaften erfüllt. Zu zeigen ist, dass ein Zeuge a mit $A = [a]$ existiert. Weil A gemäß (1) nichtleer ist, muss ein Element $a \in A$ existieren. Für jedes weitere Element $x \in A$ ergibt sich $x \sim a$ aufgrund (2), also $x \in [a]$, womit wir $A \subseteq [a]$ haben. Es verbleibt $[a] \subseteq A$ zu zeigen. Sei also $x \in [a]$. Wir haben damit die Situation $a \in A$ und $x \sim a$, womit laut (3) ebenso $x \in A$ sein muss. □

1.5 Abbildungen

1.5.1 Definitionen

Definition 1.16 (app: Applikation). Für eine Abbildung f ist

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

Definition 1.17 (img: Bildmenge).
Für eine Abbildung $f : X \to Y$ 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 1.18 (preimg: Urbildmenge). Für eine Abbildung $f : X \to Y$ 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 1.19 (inj: Injektion).
Eine Abbildung $f : X \to Y$ 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 1.20 (sur: Surjektion).
Eine Abbildung $f : X \to Y$ heißt genau dann surjektiv, wenn gilt:

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

Definition 1.21 (composition: Verkettung).
Für Abbildungen $f : X \to Y$ und $g : Y \to Z$ heißt

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

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

1.5.2 Grundlagen

Satz 1.39 (feq: Gleichheit von Abbildungen). Zwei Abbildungen $f : X \to Y$ und $g : X' \to Y'$ 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

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

Nach Satz 1.25 und Def. 1.16 (app) gilt

$$(\forall x : f(x) = g(x)) \iff (\forall x : \forall y : (y = f(x) \Leftrightarrow y = g(x)))$$
$$\iff (\forall x : \forall y : ((x,y) \in G_f \Leftrightarrow (x,y) \in G_g)) \iff \forall t : (t \in G_f \Leftrightarrow t \in G_g).$$

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

Korollar 1.40. Für eine Abbildung f gilt

$$f(x) \in A \cap B \iff f(x) \in A \wedge f(x) \in B.$$

Beweis. Es gelte $f(x) \in A \cap B$. Dann existiert laut Definition ein $y \in A \cap B$ mit $y = f(x)$, womit $y \in A$ und $y \in B$ gilt. Folglich gilt $f(x) \in A$ und $f(x) \in B$.

Es gelte $f(x) \in A$ und $f(x) \in B$. Dann existiert laut Definition ein $y \in A$ mit $y = f(x)$ und ein $y' \in B$ mit $y' = f(x)$. Weil f dem x genau ein Bild zuordnet, muss $y = y'$ gelten. Folglich gilt $y \in A \cap B$, und somit $f(x) \in A \cap B$. □

Korollar 1.41. Für eine Abbildung f gilt

$$f(x) \in A \cup B \iff f(x) \in A \vee f(x) \in B.$$

Beweis. Es findet sich die äquivalente Umformung

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

Satz 1.42 (preimg-dl: Distributivität der Urbildoperation).
Für $f : X \to Y$ und beliebige Mengen M_i gilt:

$$f^{-1}(M_1 \cap M_2) = f^{-1}(M_1) \cap f^{-1}(M_2), \tag{1.11}$$
$$f^{-1}(M_1 \cup M_2) = f^{-1}(M_1) \cup f^{-1}(M_2), \tag{1.12}$$
$$f^{-1}(\bigcap_{i \in I} M_i) = \bigcap_{i \in I} f^{-1}(M_i), \tag{1.13}$$
$$f^{-1}(\bigcup_{i \in I} M_i) = \bigcup_{i \in I} f^{-1}(M_i). \tag{1.14}$$

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

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

Für die Vereinigung ist das analog.

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

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

Nach Def. 1.18 (preimg) und Def. 1.7 (intersection) zusammen mit Def. 1.4 (filter) gilt:

$$x \in f^{-1}(\bigcap_{i\in I} M_i) \iff f(x) \in \bigcap_{i\in I} M_i \iff \forall i(i \in I \implies f(x) \in M_i)$$
$$\iff \forall i(i \in I \implies x \in f^{-1}(M_i)) \iff x \in \bigcap_{i\in I} f^{-1}(M_i).$$

Für die Vereinigung ist das analog. □

Satz 1.43 (img-cup-dl: Distributivität der Bildoperation über die Vereinigung). Für $f : X \to Y$ und Mengen $M_i \subseteq X$ gilt:

$$f(M_1 \cup M_2) = f(M_1) \cup f(M_2), \tag{1.15}$$
$$f(\bigcup_{i\in I} M_i) = \bigcup_{i\in I} f(M_i). \tag{1.16}$$

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

$$y \in f(M_1 \cup M_2) \iff (\exists x : x \in M_1 \cup M_2 \land y = f(x))$$
$$\iff (\exists x : (x \in M_1 \lor x \in M_2) \land y = f(x))$$
$$\iff (\exists x : x \in M_1 \land y = f(x) \lor x \in M_2 \land y = f(x))$$
$$\iff (\exists x : x \in M_1 \land y = f(x)) \lor (\exists x : x \in M_2 \land y = f(x))$$
$$\iff y \in f(M_1) \lor y \in f(M_2) \iff y \in f(M_1) \cup f(M_2).$$

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

$$y \in f(\bigcup_{i\in I} M_i) \iff (\exists x : x \in \bigcup_{i\in I} M_i \land y = f(x))$$
$$\iff (\exists x : (\exists i : i \in I \land x \in M_i) \land y = f(x)) \iff (\exists x : \exists i : i \in I \land x \in M_i \land y = f(x))$$
$$\iff (\exists i \exists x : i \in I \land x \in M_i \land y = f(x)) \iff (\exists i : i \in I \land \exists x (x \in M_i \land y = f(x))$$
$$\iff (\exists i : i \in I \land y \in f(M_i)) \iff y \in \bigcup_{i\in I} f(M_i).$$ □

Satz 1.44. Es gilt:

$$f(M_1 \cap M_2) \subseteq f(M_1) \cap f(M_2), \tag{1.17}$$
$$f(\bigcap_{i\in I} M_i) \subseteq \bigcap_{i\in I} f(M_i). \tag{1.18}$$

Beweis. Nach Def. 1.3 (subseteq) expandieren:

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

Nach Def. 1.17 (img), Def. 1.5 (cap) und Satz. 1.10 (exists-asym-dl) gilt:

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

Nach Def. 1.3 (subseteq) expandieren:

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

Nach Def. 1.17 (img) und Def. 1.7 (intersection) gilt:

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

Korollar 1.45. Zwei disjunkte Mengen haben disjunkte Urbilder.

Beweis. Sei $A \cap B = \emptyset$. Gemäß Satz 1.42 (preimg-dl) ist

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

Satz 1.46. 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 mittles Satz 1.42 (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

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

Beweis 2. Die Expansion der Aussage bringt

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

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

Satz 1.47. 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 1.44 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 1.48. Es gilt:

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

Beweis. Nach Def. 1.17 (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. □

Satz 1.49. Es gilt $(g \circ f)^{-1}(M) = f^{-1}(g^{-1}(M))$.

Beweis. Nach Def. 1.18 (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. 1.21 (composition) anwenden und rechts nochmals Def. 1.4 (filter):

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

Satz 1.50. Es gilt $(g \circ f)(M) = g(f(M))$.

Beweis. Nach Def. 1.17 (img) und Def. 1.2 expandieren, dann 1.4 (filter) anwenden:

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

Die rechte Seite mit Def. 1.17 (img) expandieren und Def. 1.4 (filter) anwenden. Unter Anwendung von Satz 1.8 (general-dl) und Satz 1.14 (exists-cl) ergibt sich

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

Satz 1.51. Sei $f : A \to B$ eine Abbildung und $A \neq \emptyset$. Man nennt eine Funktion $g : B \to A$ mit $g \circ f = \mathrm{id}_A$ 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 $g : B \to A$ 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. $g \circ f = \mathrm{id}$.

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

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

und

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

Es ergibt sich

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

Korollar 1.52. Es gilt $f^{-1}(A^c) = f^{-1}(A)^c$ bzw. $f(x) \in A^c \Leftrightarrow \neg f(x) \in A$.

Beweis. Zufolge der Expansion von Def. 1.18 (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 Implikaton 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. □

Satz 1.53. Für jede Abbildung f gilt $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$.

Beweis 1. Ergibt sich sofort gemäß Definition:

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

Beweis 2. Gemäß Korollar 1.52 und Satz 1.42 (preimg-dl) ist

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

Satz 1.54. 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

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

Satz 1.55. Für jede Abbildung $f : A \to B$ gilt $f(f^{-1}(N)) = N$, sofern $N \subseteq f(A)$ ist.

Beweis. Laut Satz 1.54 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. □

Satz 1.56. Für jede Abbildung $f : A \to B$ 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 1.55 gilt dann $f(M) = N$. □

Satz 1.57. Ist f injektiv, dann gilt $f(A \setminus B) = f(A) \setminus f(B)$.

Beweis. Da f injektiv ist, gibt es nach Satz 1.51 eine Linksinverse f^{-1}. Nach Satz 1.50 ist für eine beliebige Menge M die Gleichung

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

erfüllt. Unter Heranziehung von Satz 1.53 bekommt man

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

Wendet man nun auf beide Seiten der Gleichung f an, dann ergibt sich nach Satz 1.55 das gesuchte Resultat $f(A) \setminus f(B) = f(A \setminus B)$. □

Satz 1.58. 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

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

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

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

Die Bedingung ist daher immer erfüllt. □

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

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

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

Satz 1.59 (Rechtskürzbarkeit von Surjektionen).
Ist $f \colon X \to Y$ eine surjektive Abbildung, dann gilt

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

Beweis. Laut Prämisse und Satz 1.39 (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 1.39 (feq) ist daher $g = h$. □

1.5.3 Kardinalzahlen

Axiom 1.60 (acc: abzählbares Auswahlaxiom). Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge nichtleerer Mengen. Dann existiert eine Funktion $f\colon \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n$ mit $f(n) \in A_n$.

Definition 1.22 (equipotent: Gleichmächtigkeit). Zwei Mengen A, B heißen genau dann gleichmächtig, wenn eine Bijektion $f\colon A \to B$ existiert.

Satz 1.61. 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 eine Menge $A \subseteq M$ betrachte man nun die Indikatorfunktion

$$1_A\colon M \to \{0,1\}, \quad 1_A(x) := [x \in A].$$

Die Abbildung

$$\varphi\colon 2^M \to \{0,1\}^M, \quad \varphi(A) := 1_A$$

ist eine kanonische Bijektion.

Zur Injektivität. Nach Def. 1.19 (inj) muss gelten:

$$\varphi(A) = \varphi(B) \implies A = B, \quad \text{d. h.} \quad 1_A = 1_B \implies A = B.$$

Nach Satz 1.39 (feq) und Def. 1.2 (seteq) wird die Aussage expandiert zu:

$$(\forall x\colon 1_A(x) = 1_B(x)) \implies (\forall x\colon x \in A \iff x \in B).$$

Es gilt aber nun:

$$1_A(x) = 1_B(x) \iff [x \in A] = [x \in B] \iff (x \in A \iff x \in B).$$

Zur Surjektivität. Wir müssen nach Def. 1.20 (sur) prüfen, dass $\{0,1\}^M \subseteq \varphi(2^M)$ gilt. Expansion nach Def. 1.3 (subseteq) und Def. 1.17 (img) ergibt:

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

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 $f = 1_A$, denn

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

Da φ eine Bijektion ist, müssen 2^M und $\{0,1\}^M$ nach Def. 1.22 (equipotent) gleichmächtig sein. □

Satz 1.62. Man setze Axiom 1.60 (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 $\mathrm{Abb}(\mathbb{N}, A_n)$. Nach Axiom 1.60 (acc) kann aus jeder Menge B_n eine Bijektion $f_n : \mathbb{N} \to A_n$ ausgewählt werden. Man betrachte nun

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

Die Abbildung φ ist surjektiv, denn nach Satz 1.48 und Satz 1.30 gilt

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

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

Satz 1.63. Wenn R abzählbar ist, dann ist auch der Polynomring $R[X]$ abzählbar.

Beweis. Zu jedem Polynom vom Grad $n \ge 1$ 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 1.62 gilt

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

Satz 1.64. 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 1.63 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 $|\mathbb{A}| = |\mathbb{N}|$. □

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 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. □

Beweis 3. Für $n \in \mathbb{N}$ sei

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

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

Definition 1.23 (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 $f_1 : X \to X'$ und $f_2 : Y \to Y'$. Gesucht ist mindestens eine Bijektion $f : X \times Y \to X' \times Y'$. 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 π_k ist die Projektion auf die k-te Komponente gemeint. □

Definition 1.24 (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 $f_1 : X \to X'$ und $f_2 : Y \to Y'$, wobei $X \cap Y = \emptyset$ und $X' \cap Y' = \emptyset$ gilt. Gesucht ist mindestens eine Bijektion $f : X \cup Y \to X' \cup Y'$. 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 $x = f_2^{-1}(x')$. □

Definition 1.25 (Satz und Def. Potenz von Kardinalzahlen).
Die Operation $|Y|^{|X|} := |Y^X|$ ist wohldefiniert.

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

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

Die Abbildung F ist injektiv, da

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

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 $f = f_2^{-1} \circ f' \circ f_1$. □

Definition 1.26 (Weniger mächtig).
Eine Menge A heißt weniger mächtig als eine Menge B, kurz $|A| < |B|$, wenn es eine Injektion $A \to B$ gibt, aber keine Bijektion $A \to B$.

Satz 1.65 (Satz von Cantor).
Eine Menge ist stets weniger mächtig als ihre Potenzmenge. Kurz $|A| < |2^A|$.

Beweis. Eine Injektion $A \to 2^A$ können wir mit $x \mapsto \{x\}$ trivial angeben. Nun wird die Annahme, es gäbe eine Surjektion $f : A \to 2^A$ 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 1.66. Die Menge der endlichen Teilmengen der natürlichen Zahlen ist abzählbar.

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

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

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 Indiktorfunktion 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. □

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 ε einen Index n_0 gibt, so dass ab diesem Index sämtliche ihrer Werte in der Umgebung $U_\varepsilon(a)$ liegen. Formal:

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

bzw.

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

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 $a_n \to g_1$. 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:

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

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. 3.13 (metric-space) gilt $d(g_1, g_2) > 0$, daher ist auch $\varepsilon > 0$. Nach Satz 3.12 sind die Umgebungen $U_\varepsilon(g_1)$ und $U_\varepsilon(g_2)$ disjunkt. Wegen $a_n \to g_1$ 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 $a_n \notin U_\varepsilon(g_2)$. □

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

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

wobei $R > 0$ ein fester aber beliebieger 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 düfen wir die Teilformel $\varepsilon > 0$ nun ersetzen. Es ergibt sich die äquivalente Formel

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

Das ist aber genau Def. 2.2 (lim). □

Satz 2.3. Es gilt:

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

Beweis. Nach Satz 3.14 (umgekehrte Dreiecksungleichung) gilt:

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

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

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 $a_n \to 0$, dann muss für jedes ε 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 $a_n b_n \to 0$. □

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. □

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

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

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. 3.15 (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. □

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:

$$\lim_{n\to\infty} a_n = a \implies \lim_{n\to\infty} ra_n \to ra.$$

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. 3.15 (normed-space) Axiom (N2) heran, dann ergibt sich

$$\|ra_n - ra\| = |r|\,\|a_n - a\| < |r|\varepsilon.$$

Nach Satz 2.2 (lim-scaled-ep) folgt die Behauptung. □

Satz 2.8 (Grenzwertsatz zum Produkt).
Seien (a_n) und (b_n) Folgen reeller Zahlen. Es gilt:

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

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

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

Da nach Satz 2.7 aber $a_n b \to ab$ und $ab_n \to ab$, ergibt sich nach Satz 2.6 nun

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

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

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

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

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

Beweis. Wenn $x_n \to x$, dann gilt trivialerweise auch $x_{n+1} \to x$. 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:

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

2.1.2 Wachstum und Landau-Symbole

Definition 2.4. Seien $f, g: D \to \mathbb{R}$ 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

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

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

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

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

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

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

Korollar 2.12. Sind $f_1, f_2 \in \mathcal{O}(g)$, ist auch $f_1 + f_2 \in \mathcal{O}(g)$.

Beweis. Als Prämissen liegen Zeugen $c' > 0, x_0'$ und $c'' > 0, x_0''$ für

$$\forall x > x_0': |f_1(x)| \leq c'|g(x)|,$$
$$\forall x > x_0'': |f_2(x)| \leq c''|g(x)|$$

vor. Mit der Dreiecksungleichung findet sich

$$|f_1(x) + f_2(x)| \leq |f_1(x)| + |f_2(x)| \leq c'|g(x)| + c''|g(x)| = (c' + c'')|g(x)|$$

für $x > \max(x_0', x_0'')$. Ergo sind $x_0 := \max(x_0', x_0'')$ und $c := c' + c''$ Zeugen für

$$\exists c > 0: \exists x_0: \forall x > x_0: |f_1(x) + f_2(x)| \leq c|g(x)|. \quad \square$$

2.2 Stetige Funktionen

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

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

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

Definition 2.6 (cont: stetig). Eine Funktion $f : D \to \mathbb{R}$ 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 $f : D \to \mathbb{R}$ 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 $f : D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ heißt Lipschitz-stetig an der Stelle $x_0 \in D$, wenn eine Konstante L existiert, so dass

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

für alle $a \in D$.

Korollar 2.13. 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 $f : D \to \mathbb{R}$ 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 $f : D \to \mathbb{R}$ 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 2.14. Ist die Funktion $f: D \to \mathbb{R}$ 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 3.14 gilt

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

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 2.15. Eine differenzierbare Funktion ist genau dann Lipschitz-stetig, wenn ihre Ableitung beschränkt ist.

Beweis. Wenn $f: I \to \mathbb{R}$ 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

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

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

Beweis. Sei $f: [a, b] \to \mathbb{R}$ stetig differenzierbar. Dann ist $f'(x)$ stetig. Nach dem Satz vom Minimum und Maximum ist $|f'(x)|$ beschränkt. Nach Satz 2.15 muss f Lipschitz-stetig sein. $\square$

Korollar 2.17. Eine stetig differenzierbare Funktion ist lokal Lipschitz-stetig.

Beweis. Sei $f: D \to \mathbb{R}$ 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 2.16. Dann ist auch die Einschränkung von f auf $U_\varepsilon(x_0) \subseteq [a,b]$ Lipschitz-stetig. □

Satz 2.18. Es gibt differenzierbare Funktionen, die nicht überall lokal Lipschitz-stetig sind.

Beweis. Aus Satz 2.15 ergibt sich also Kontraposition, dass eine Funktion mit unbeschränkter Ableitung nicht Lipschitz-stetig sein kann.

Ist $f: D \to \mathbb{R}$ 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 $f: [0,\infty) \to \mathbb{R}$ mit

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

Einerseits gilt

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

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

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

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

Satz 2.19. Sei $f: \mathbb{R} \to \mathbb{R}$ differenzierbar und $f(x)$ konvergent für $x \to \infty$. Ist außerdem f' Lipschitz-stetig, zieht dies $f'(x) \to 0$ für $x \to \infty$ nach sich.

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

$$|f(b)-f(a)| < \varepsilon \tag{2.1}$$

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

$$\left| \int_a^b f'(x)\,\mathrm{d}x \right| = |f(b)-f(a)| \tag{2.2}$$

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

$$\frac{1}{2L} f'(a)^2 = \int_a^b f'(x)\,\mathrm{d}x < \varepsilon. \tag{2.3}$$

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. □

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 $f: D \to \mathbb{R}$ heißt differenzieraber an der Stelle $x_0 \in D$, wenn der Grenzwert

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

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

Satz 2.20. Sei I ein Intervall und $f, g: I \to \mathbb{R}$. Sind f, g differenzierbar an der Stelle $x \in I$, dann ist auch

$$f + g \text{ dort differenzierbar mit } (f+g)'(x) = f'(x) + g'(x), \tag{2.4}$$
$$f - g \text{ dort differenzierbar mit } (f-g)'(x) = f'(x) - g'(x), \tag{2.5}$$
$$fg \text{ dort differenzierbar mit } (fg)'(x) = f'(x)g(x) + f(x)g'(x). \tag{2.6}$$

Beweis. Es gilt

$$(f+g)'(x) = \lim_{h \to 0} \frac{(f+g)(x+h) - (f+g)(x)}{h} \tag{2.7}$$
$$= \lim_{h \to 0} \frac{(f(x+h) + g(x+h)) - (f(x) + g(x))}{h} \tag{2.8}$$
$$= \lim_{h \to 0} \left(\frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h} \right) \tag{2.9}$$
$$= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} + \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} = f'(x) + g'(x). \tag{2.10}$$

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:

$$g(x)f'(x) + f(x)g'(x) = g(x) \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} + f(x) \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} \tag{2.11}$$
$$= \lim_{h \to 0} \left[g(x+h) \frac{f(x+h) - f(x)}{h} \right] + \lim_{h \to 0} \left[f(x) \frac{g(x+h) - g(x)}{h} \right] \tag{2.12}$$
$$= \lim_{h \to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h)}{h} + \lim_{h \to 0} \frac{f(x)g(x+h) - f(x)g(x)}{h} \tag{2.13}$$
$$= \lim_{h \to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \tag{2.14}$$
$$= \lim_{h \to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} = \lim_{h \to 0} \frac{(fg)(x+h) - (fg)(x)}{h} = (fg)'(x). \tag{2.15}$$

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

Satz 2.21. Sei I ein Intervall. Sind $f, g : I \to \mathbb{R}$ an der Stelle x differenzierbar und ist $g(x) \neq 0$, dann ist auch f/g differenzierbar und es gilt

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

Beweis. Nach der Produktregel (2.6) gilt

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

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

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

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

Satz 2.22. Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^n$ mit $n \in \mathbb{N}$ gilt $f'(x) = nx^{n-1}$.

Beweis 1. Heranziehung des binomischen Lehrsatzes bringt

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

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

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

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

Satz 2.23. Für $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) := x^n$ 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 2.22 gezeigt. Sei nun $a \in \mathbb{N}$ und $n = -a$. Nach der Produktregel (2.6) und Satz 2.22 gilt

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

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

$$\frac{\mathrm{d}}{\mathrm{d}x} x^{-a} = -a x^{-a-1} \implies \frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}. \square \tag{2.24}$$

2.3.2 Glatte Funktionen

Satz 2.24. Sei $f : \mathbb{R} \to \mathbb{R}$ 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 $x \to 0$ schneller fällt als jede rationale Funktion steigen kann, muss die rechtsseitige Ableitung an der Stelle $x = 0$ immer verschwinden. □

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

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

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

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

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

Korollar 2.25. 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

$$\begin{aligned} D_v(f + g)(x) &= D_v f(x) + D_v g(x), \\ D_v(f - g)(x) &= D_v f(x) + D_v g(x), \\ D_v(cf)(x) &= c D_v f(x), \\ D_v(fg)(x) &= g(x) D_v f(x) + f(x) D_v g(x). \end{aligned}$$

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. □

Korollar 2.26 (Kettenregel).
Sei $g\colon \mathbb{R} \to \mathbb{R}$ 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 $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ die Standardbasis. Die partielle Ableitung $\partial_k f(x)$ ist definiert als die Richtungsableitung $D_v f(x)$ bezüglich $v = \mathbf{e}_k$.

Korollar 2.27. Zur jeder gewöhnlichen Ableitungsregel besitzt die Richtungsableitung eine analoge Regel.

Vorbereitung. Sei $f = (f_1, \ldots, f_n)$ ein Tupel von Funktionen aus einem Funktionenraum und sei entsprechend $f(x) := (f_1(x), \ldots, f_n(x))$. 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

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

die komponentenweise Anwendung eines Operators T. Sei C_γ 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 $\varphi\colon M \to M$ 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 2.28 (Fixpunktsatz von Banach). Sei (M, d) ein nichtleerer vollständiger metrischer Raum und sei $\varphi\colon M \to M$ eine Kontraktion. Es gibt genau einen Fixpunkt $x \in M$ mit $x = \varphi(x)$ und die Folge $(x_n)\colon \mathbb{N} \to M$ mit $x_{n+1} = \varphi(x_n)$ konvergiert gegen den Fixpunkt, unabhängig vom Startwert x_0.

Satz 2.29 (Hinreichendes Konvergenzkriterium). Sei $M = [a, b]$. Ist $\varphi\colon M \to M$ differenzierbar und gibt es eine Zahl r mit $|\varphi'(x)| < r < 1$ für alle $x \in M$, dann hat φ 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 2.15 ist eine differenzierbare Funktion φ 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., φ ist eine Kontraktion. Die Konvergenz der Folge (x_n) ist gemäß Satz 2.28 gewährleistet. □

Satz 2.30 (Hinreichendes Konvergenzkriterium zum Newton-Verfahren). Sei $f\colon [a, b] \to \mathbb{R}$ zweimal stetig differenzierbar und $f'(x) \neq 0$ für alle x. Sei

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

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 φ 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 2.29 konvergiert die Iteration (x_n) gegen den einzigen Fixpunkt von φ. 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 φ ist also die einzige Nullstelle von f. □

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

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

Definition 3.5 (ext: Äußeres).
Sei X ein topologischer Raum und $M \subseteq X$. Das Äußere von M ist

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

Definition 3.6 (Abgeschlossene Hülle).
Sei X ein topologischer Raum und $M \subseteq X$. Die abgeschlossene Hülle von M ist

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

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

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

Definition 3.8 (Teilraumtopologie).
Sei (X, T) ein topologischer Raum und $M \subseteq X$. Man bezeichnet

$$T|M := \{A \cap M \mid A \in T\}.$$

als Teilraumtopologie und $(M, T|M)$ als Teilraum.

Definition 3.9 (Diskrete Topologie).
Man sagt, ein topologischer Raum (X, T) habe die diskrete Topologie T, wenn T die Potenzmenge von X ist – wenn also jede Teilmenge von X offen ist.

3.1.2 Stetige Abbildungen

Definition 3.10 (Stetige Abbildung).
Seien X und Y topologischen Räume. Eine Abbildung $f: X \to Y$ heißt stetig, wenn unter ihr das Urbild einer offenen Menge stets wieder offen ist.

Korollar 3.1. Sei $M \subseteq X$. Ist $f: X \to Y$ stetig, so ist die Einschränkung $f|_M$ stetig.

Beweis. Vorgelegt ist ein offenes $B \subseteq Y$. Laut Prämisse ist $A = f^{-1}(B)$ ebenfalls offen. Per Definition der Einschränkung gilt

$$(f|_M)^{-1}(B) = M \cap f^{-1}(B) = M \cap A.$$

Laut Definition ist $M \cap A$ ein Element der Teilraumtopologie $T(X)|M$. □

Korollar 3.2. Auf jedem topologischen Raum ist die identische Abbildung stetig.

Beweis. Ihre Urbildoperation ist ebenfalls eine identische Abbildung. Ergo verbleiben offene Mengen unverändert, also offen. □

Korollar 3.3. Die Verkettung stetiger Abbildungen ist stetig.

Beweis. Seien $f: X \to Y$ und $g: Y \to Z$ stetig. Gemäß Satz 1.49 gilt $(g \circ f)^{-1}(M) = f^{-1}(g^{-1}(M))$. Ist M offen, so auch $g^{-1}(M)$ und infolge $f^{-1}(g^{-1}(M))$. Ergo ist $g \circ f$ stetig. □

Definition 3.11 (Homöomorphismus).
Seien X und Y topologische Räume. Eine bijektive Abbildung $f: X \to Y$ heißt Homöomorphismus, wenn sowohl f als auch f^{-1} stetig sind.

3.1.3 Elementares

Korollar 3.4. Sei X ein topologischer Raum. Für jede Menge $M \subseteq X$ ist

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

eine disjunkte Zerlegung.

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

$$\operatorname{int}(M) \cup \partial M \cup \operatorname{ext}(M) = A \cup B^{\mathsf{C}} \cap A^{\mathsf{C}} \cup B = A \cup A^{\mathsf{C}} \cup B = X \cup B = X.$$

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

$$\begin{aligned} \operatorname{int}(M) \cap \partial M &= A \cap B^{\mathsf{C}} \cap A^{\mathsf{C}} = \emptyset, \\ \operatorname{ext}(M) \cap \partial M &= B \cap B^{\mathsf{C}} \cap A^{\mathsf{C}} = \emptyset. \end{aligned}$$

Wegen $M \cap M^{\mathsf{C}} = \emptyset$ ist erst recht $A \cap B = \emptyset$, denn $A \subseteq M$ und $B \subseteq M^{\mathsf{C}}$. □

Korollar 3.5. Für jede Menge $A \subseteq X$ gilt $\operatorname{int}(A)^{\mathsf{C}} \cup A = X$.

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

$$B^{\mathsf{C}} \cup A = (A \cap B)^{\mathsf{C}} \cup A = A^{\mathsf{C}} \cup B^{\mathsf{C}} \cup A = X \cup B^{\mathsf{C}} = X. \quad \square$$

Satz 3.6. Das Innere von M ist die Vereinigung der offenen Teilmengen von M, kurz

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

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. □

Satz 3.7. 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.

3.1.4 Zusammenhang

Definition 3.12 (Zusammenhängender Raum).
Ein topologischer Raum (X, T) heißt zusammenhängend, wenn er sich nicht in zwei disjunkte nichtleere offene Mengen zerlegen lässt. Gemeint ist

$$\forall A, B \in T : A \neq \emptyset \land B \neq \emptyset \land A \cap B = \emptyset \implies A \cup B \neq X.$$

Bemerkung. Ein Raum (X, T) ist demnach unzusammenhängend, wenn Zeugen A, B für die Aussage

$$\exists A, B \in T : A \neq \emptyset \land B \neq \emptyset \land A \cap B = \emptyset \land A \cup B = X$$

gefunden sind.

Satz 3.8. Sei X ein topologischer Raum und $\{0, 1\}$ der topologische Raum mit der diskreten Topologie. Es ist X genau dann zusammenhängend, wenn jede stetige Abbildung $f : X \to \{0, 1\}$ konstant sein muss.

Beweis. Sei f stetig und nicht-konstant. Man betrachte die Fasern $A := f^{-1}(\{0\})$ und $B := f^{-1}(\{1\})$. Weil f, wie gerade gefordert, beide Werte annehmen muss, gilt $A \neq \emptyset$ und $B \neq \emptyset$. Zudem sind A, B offen, weil sie die Urbilder offener Mengen sind unter stetigem f sind. Urbilder disjunkter Mengen sind immer disjunkt, womit $A \cap B = \emptyset$ gilt. Zudem gilt

$$A \cup B = f^{-1}(\{0\}) \cup f^{-1}(\{1\}) = f^{-1}(\{0\} \cup \{1\}) = X.$$

Somit ist ein Gegenbeispiel konstruiert, so dass X unzusammenhängend sein muss.

Sei X unzusammenhängend. Es existieren somit Zeugen $A, B \in T$ mit $A \neq \emptyset$, $B \neq \emptyset$, $A \cup B = \emptyset$ und $A \cup B = X$. Sei f definiert durch $f(x) := 0$ für alle $x \in A$ und $f(x) := 1$ für alle $x \in B$. Nun ist f keine konstante Abbildung, da sie auf nichtleeren A, B unterschiedliche Werte annimmt. Wohl aber ist f stetig, wie im Folgenden noch durchgerechnet wird. Die diskrete Topologie von $\{0, 1\}$ ist ihre Potenzmenge. Es bestätigt sich

$$f^{-1}(\emptyset) = \emptyset \in T,$$
$$f^{-1}(\{0\}) = A \in T,$$
$$f^{-1}(\{1\}) = B \in T,$$
$$f^{-1}(\{0, 1\}) = f^{-1}(\{0\} \cup \{1\}) = f^{-1}(\{0\}) \cup f^{-1}(\{1\}) = A \cup B = X \in T. \square$$

Korollar 3.9. Die Vereinigung zweier zusammenhängender nichtdisjunkter Räume ist ein zusammenhängender Raum.

Beweis. Seien X, Y zusammenhängend und sei $X \cap Y \neq \emptyset$. Laut Prämisse existiert mindestens ein $p \in X \cap Y$. Sei $f : X \cup Y \to \{0, 1\}$ stetig. Nun ist $f|_X$ stetig und konstant mit $f(x) = f(p)$ für alle $x \in X$, da X zusammenhängend ist. Entsprechend ist $f|_Y$ stetig und konstant mit $f(y) = f(p)$ für alle $y \in Y$. Ergo ist f auf ganz $X \cup Y$ konstant $f(p)$. Laut Satz 3.8 muss $X \cup Y$ also zusammenhängend sein. $\square$

Korollar 3.10. Ein topologischer Raum X ist genau dann zusammenhängend, wenn mit Ausnahme von $\emptyset$ und X keine Teilmenge von X sowohl offen als auch abgeschlossen ist.

Beweis. Es existiere außer $\emptyset, X$ kein offenes A mit offenem $X \setminus A$. Sei $f : X \to \{0,1\}$ stetig. Angenommen, f wäre nicht konstant. Dann gäbe es die beiden nichtleeren offenen Mengen $A := f^{-1}(\{0\})$ und $B := f^{-1}(\{1\})$. Weil außerdem $X = A \cup B$ eine disjunkte Zerlegung wäre, gälte $B = X \setminus A$. Dies steht im Widerspurch zur Prämisse, womit f konstant sein muss. Ergo ist X laut Satz 3.8 ein zusammenhängender Raum.

Es existiere nun offnes, von $\emptyset, X$ verschiedenes A mit offenem $B := X \setminus A$, womit neben $A \neq \emptyset$ auch $B \neq \emptyset$ gilt. Wegen $A \cap B = \emptyset$ und $A \cup B = X$ gilt X als in zwei nichtleere disjunkte offene Mengen zerlegt. Ergo ist X unzusammenhängend. □

Korollar 3.11. Ist X zusammenhängend und $f : X \to Y$ stetig, dann ist der Teilraum $f(X)$ von Y ebenfalls zusammenhängend.

Beweis. Sei für die Kontraposition $f(X)$ unzusammenhängend. Hiermit existiert eine disjunkte Zerlegung in offene nichtleere Mengen A, B mit $A \cup B = f(X)$. Aufgrund der elementaren Eigenschaften der Urbildoperation hat man mit

$$X = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \qquad f^{-1}(A) \cap f^{-1}(B) = \emptyset$$

eine disjunkte Zerlegung von X. Wegen $A \subseteq f(X)$ und $B \subseteq f(X)$ gilt hierbei $f^{-1}(A) \neq \emptyset$ und $f^{-1}(B) \neq \emptyset$. Weil f stetig ist, sind $f^{-1}(A)$ und $f^{-1}(B)$ offen. Somit ist für X eine Zerlegung in zwei nichtleere disjunkte offene Mengen bezeugt. Ergo ist X ebenfalls unzusammenhängend. □

3.2 Metrische Räume

3.2.1 Metrische Räume

Definition 3.13 (metric-space: metrischer Raum). Man bezeichet (M, d) mit $d\colon M^2 \to \mathbb{R}$ 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 3.14 (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 3.12 (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. 3.13 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 $c < c$. □

Korollar 3.13 (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 3.12. Wähle speziell z. B. $a = b = d(p,q)/2$. □

3.2.2 Normierte Räume

Definition 3.15 (normed-space: normierter Raum). Sei V ein Vektorraum über dem Körper der rellen 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 3.14 (umgekehrte Dreiecksungleichung). In jedem normierten Raum gilt

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

Beweis. Auf beiden Seiten von Def. 3.15 (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 $|a| \leq b$. □

3.2.3 Homöomorphien

Satz 3.15 (Verallgemeinerung des Zwischenwertsatzes).
Ist $f : X \to Y$ eine stetige Abbildung zwischen topologischen Räumen und $A \subseteq X$ ein zusammenhängender Teilraum, dann ist auch $f(A)$ zusammenhängend.

Satz 3.16. Eine injektive Abbildung $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ 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 1.57. Da $\mathbb{R}_{>0}$ zusammenhängend ist, $\mathbb{R} \setminus \{f(0)\}$ aber nicht, kann f laut Satz 3.15 nicht stetig sein. □

3.3 Übungen

Satz 3.17. 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 $\gamma : [0, 1] \to \mathbb{R}^n$ 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

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

laut Korollar 3.4 eine disjunkte Zerlegung. Gemäß Satz 1.42 (preimg-dl) gilt

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

was auch eine disjunkte Zerlegung ist, weil je zwei disjunkte Mengen gemäß Korollar 1.45 disjunkte Urbilder haben. Weil γ gemäß Definition stetig ist, sind die Urbilder $\gamma^{-1}(\operatorname{int}(M))$ und $\gamma^{-1}(\operatorname{ext}(M))$ 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 unzusammhängender Raum. Das steht im Widerspruch zur Erkenntnis, dass alle Intervalle zusammenhängend sind. □

Beweis 2. Sei $\gamma\colon [0,1] \to \mathbb{R}^n$ 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 γ stetig ist, konvergiert $\gamma(a_k) \to \gamma(a)$ für $a_k \to a$ und $\gamma(b_k) \to \gamma(a)$ für $b_k \to a$. Demnach sind in jeder Umgebung von $\gamma(a)$ sowohl Punkte aus dem Inneren als auch Punkte aus dem Äußeren. Gemäß Satz 3.7 muss $\gamma(a)$ infolge auf dem Rand liegen. □

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 $\overline{A}$ mit $(\overline{A})_{ij} := \overline{A_{ij}}$ 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 $A^H := (\overline{A})^T$, 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:

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

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

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

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

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

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

Beweis. Es gilt

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

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

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

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

$$M^T = (A^T A)^T = A^T (A^T)^T = A^T A = M. \tag{4.4}$$

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

$$\lambda |v|^2 = \langle \lambda v, v\rangle = \langle Mv, v\rangle = \langle A^T Av, v\rangle = \langle Av, Av\rangle = |Av|^2 \geq 0. \tag{4.5}$$

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

$$M^H = (A^H A)^H = A^H (A^H)^H = A^H A = M. \tag{4.6}$$

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

$$\lambda |v|^2 = \langle \lambda v, v\rangle = \langle Mv, v\rangle = \langle A^H A v, v\rangle = \langle Av, Av\rangle = |Av|^2 \geq 0. \tag{4.7}$$

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

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

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

Radizieren ergibt $|Av| = |v|$. □

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

Beweis. Sei v ein Eigenvektor zum Eigenwert λ. Laut Korollar 4.6 ist dann

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

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

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

$$\Phi \colon \mathbb{C} \to M, \quad \Phi(a + b\mathrm{i}) := aE + bI$$

ist ein Isomorphismus zwischen Körpern.

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

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

und

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

ist das Untergruppenkriterium erfüllt. Die Abgeschlossenheit bezüglich Multiplikation:

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

Das Kommutativgesetz:

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

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

$$\Phi((a + b\mathrm{i}) + (c + d\mathrm{i})) = \Phi(a + b\mathrm{i}) + \Phi(c + d\mathrm{i}) \tag{4.12}$$

und

$$\Phi((a + b\mathrm{i})(c + d\mathrm{i})) = \Phi(a + b\mathrm{i})\Phi(c + d\mathrm{i}) \tag{4.13}$$

ergeben sich aus den Rechnungen (4.8) und (4.10). □

4.3 Lineare Abbildungen

Definition 4.6. Sei $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ ein System von Vektoren. Man nennt es linear unabhängig, wenn

$$\sum_{k=1}^{n} \lambda_k \mathbf{a}_k = \mathbf{0} \implies \forall k : \lambda_k = 0.$$

Korollar 4.12. Sei K einer Körper und $A \in K^{m \times n}$ eine Matrix. Es besitzt A genau dann den trivialen Kern, wenn ihre Spalten linear unabhängig sind.

Beweis. Die Spalten seien die $\mathbf{a}_k := A\mathbf{e}_k$. Für die Aussage, trivialen Kern zu haben, findet sich die äquivalente Umformung

$$\operatorname{Kern}(A) = \{\mathbf{v} \mid A\mathbf{v} = \mathbf{0}\} = \{\mathbf{0}\} \iff (A\mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v} = \mathbf{0}).$$

Dies stellt mit $\lambda_k = v_k$ nun lediglich eine Kurzschreibweise für die Eigenschaft

$$\sum_{k=1}^{n} \lambda_k \mathbf{a}_k = \mathbf{0} \implies \forall k : \lambda_k = 0.$$

dar, die die lineare Unabhängigkeit Def. 4.6 definiert. □

Korollar 4.13. Eine lineare Abbildung ist genau injektiv, wenn sie einen trivalen Kern besitzt.

Beweis. Sei L die lineare Abbildung. Für die Aussage, trivialen Kern zu haben, findet sich die äquivalente Umformung

$$\begin{aligned}
\operatorname{Kern}(L) = \{0\} &\iff L(v) = 0 \Rightarrow v = 0 \\
&\iff L(a-b) = 0 \Rightarrow a - b = 0 && (v = a - b) \\
&\iff L(a) - L(b) = 0 \Rightarrow a - b = 0 && (\text{weil } L \text{ linear ist}) \\
&\iff L(a) = L(b) \Rightarrow a = b.
\end{aligned}$$

Die letzte Aussage ist die definierende Eigenschaft der Injektivität Def. 1.19. □

Satz 4.14. Die Spaltenvektoren einer Matrix $A \in \mathbb{R}^{m \times n}$ sind genau dann linear unabhängig, wenn $\det(G) \neq 0$ gilt, wobei $G = A^T A$ die gramsche Matrix ist.

Beweis. Existiert ein $v \neq 0$ mit $Av = 0$, folgt sofort $Gv = A^T 0 = 0$, womit G keine reguläre Matrix sein kann, was gleichbedeutend mit $\det(G) = 0$ ist.

Sei umgekehrt $\det(G) = 0$. Da G hiermit keine reguläre Matrix ist, existiert ein $v \neq 0$ mit $Gv = 0$, und somit $\langle Gv, v \rangle = 0$. Laut Satz 4.3 ist $\langle Av, Av \rangle = \langle A^T Av, v \rangle$, was zur äquivalenten Umformung

$$Av = 0 \iff |Av| = 0 \iff 0 = |Av|^2 = \langle Av, Av \rangle = \langle Gv, v \rangle$$

führt. Demzufolge muss $Av = 0$ sein. □

Satz 4.15. Sei K ein geordneter Körper. Die Spalten einer Matrix $A \in K^{m \times n}$ sind genau dann linear unabhängig, wenn $\det(G) \neq 0$ gilt, wobei $G = A^T A$ die gramsche Matrix ist.

Beweis. Man betrachte zu $v, w \in K^n$ die Bilinearform $B(v, w) := v^T w$. Laut Korollar 5.23 gilt $B(v, v) > 0$ für $v \neq 0$. Analog zu Korollar 4.3 gilt $B(v, Aw) = B(A^T v, w)$, weshalb $B(Av, Av) = B(A^T Av, v)$ gilt. Mit den genannten Betrachtungen ist die zum Beweis von Satz 4.14 analoge Argumentation durchführbar. □

Satz 4.16. Die Spaltenvektoren einer Matrix $A \in \mathbb{C}^{m \times n}$ sind genau dann linear unabhängig, wenn $\det(G) \neq 0$ gilt, wobei $G = A^H A$ die gramsche Matrix ist.

Beweis. Die Argumentation verläuft analog zum Beweis von Satz 4.14, wobei Korollar 4.6 anstelle von Korollar 4.3 zur Anwendung kommt. □

Korollar 4.17. Zu einer linearen Abbildung $f : V \to W$ ist die Bildmenge $f(V)$ ein Untervektorraum von W.

Beweis. Wir ziehen das Untervektorraumkriterium heran. Es ist $f(V)$ nichtleer, weil der Definitionsbereich V nichtleer ist, da dieser mindestens den Nullvektor enthalten muss. Es gelte $w \in f(V)$ und $w' \in f(V)$. Dann existieren Zeugen v mit $w = f(v)$ und v' mit $w' = f(v')$. Infolge gilt $w + w' = f(v) + f(v') = f(v + v')$, womit $v + v'$ ein Zeuge für $w + w' \in f(V)$ ist. Gleichermaßen gilt $\lambda w = \lambda f(v) = f(\lambda v)$, womit λv ein Zeuge für $\lambda w \in f(V)$ ist. Die Forderungen des Kriteriums sind also erfüllt. □

Korollar 4.18. Zu einer linearen Abbildung $f : V \to W$ ist Graph von f ein Untervektorraum von $V \times W$.

Beweis. Der Graph von f ist die Menge

$$G = \{(v, w) \in V \times W \mid w = f(v)\}.$$

Der Raum $V \times W$ hat die Addition $(v, w) + (v', w') := (v + v', w + w')$ und die Skalarmultiplikation $\lambda(v, w) := (\lambda v, \lambda w)$.

Wir ziehen das Untervektorraumkriterium heran. Der Graph G ist nichtleer, weil der Definitionsbereich V nichtleer ist, da dieser mindestens den Nullvektor enthalten muss. Zu zeigen ist nun $(v + v', w + w') \in G$, sofern $(v, w) \in G$ und $(v', w') \in G$. Aufgrund der Prämissen ist $w = f(v)$ und $w' = f(v')$, womit

$$w + w' = f(v) + f(v') = f(v + v'), \iff (w + w', v + v') \in G.$$

Zu zeigen ist schließlich $(\lambda v, \lambda w) \in G$, sofern $(v, w) \in G$. Aufgrund der Prämisse ist $w = f(v)$, womit

$$\lambda w = \lambda f(v) = f(\lambda v), \iff (\lambda w, \lambda v) \in G. \square$$

4.4 Bilinearformen

Definition 4.7 (Nicht-ausgeartete Bilinearform).
Sei $B\colon V \times W \to K$ eine Bilinearform, sei

$$B_1\colon V \to W^*, \quad B_1(v)(w) := B(v,w),$$
$$B_2\colon W \to V^*, \quad B_2(w)(v) := B(v,w).$$

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

Korollar 4.19. Eine Bilinearform $B\colon V \times W \to K$ 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. 4.7.

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

$$\{0\} = \operatorname{Kern}(B_1) := \{v \mid B_1(v) = 0\} \tag{4.14}$$

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

$$B_1(v) = 0 \implies v = 0, \tag{4.15}$$

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

Korollar 4.20. Eine symmetrische Bilinearform $B\colon V \times V \to K$ 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. 4.7. Es genügt also, B_1 zu betrachten. Nun gilt

$$B_1(v) = 0 \iff (\forall w\colon B_1(v)(w) = 0(w)) \iff (\forall w\colon B(v,w) = 0). \tag{4.16}$$

Aus Korollar 4.19 ergibt sich dann die Behauptung, d. h. die Äquivalenz zu

$$v \neq 0 \implies \exists w\colon B(v,w) \neq 0. \ \square \tag{4.17}$$

Korollar 4.21. Ein reelles Skalarprodukt $\langle v, w\rangle$ ist nicht-ausgeartet.

Beweis. In Korollar 4.20 setze $B(v,w) := \langle v, w\rangle$. Wegen

$$\langle v, v\rangle = 0 \iff v = 0 \tag{4.18}$$

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

Satz 4.22. Sind V, W endlichdimensional, dann sind bei einer nicht-ausgearteten Bilinearform $B\colon V \times W \to K$ die Abbildungen B_1, B_2 aus Def. 4.7 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

$$\dim V \leq \dim W^* = \dim W \leq \dim V^* = \dim V. \tag{4.19}$$

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

4.5 Euklidische Geometrie

Satz 4.23 (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

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

$$= |\mathbf{c}|^2 - |\mathbf{a}|^2 = 0. \tag{4.21}$$

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

Satz 4.24 (Kosinussatz).
Gegeben ist ein Dreieck $\triangle ABC$. Sei γ 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:

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

$$= a^2 + b^2 - 2ab\cos\gamma. \ \square \tag{4.23}$$

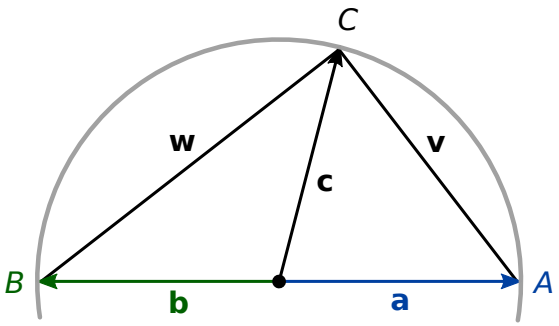

Abbildung 4.1: Zeichnung zum Satz des Thales

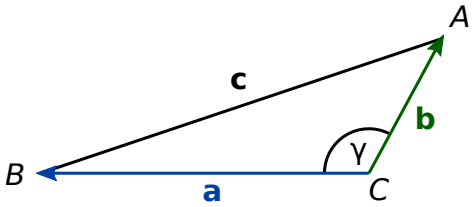

Abbildung 4.2: Zeichnung zum Kosinussatz

Satz 4.25 (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

$$\begin{aligned}
ab\sin\gamma\, \mathbf{e}_1 \wedge \mathbf{e}_2 &= \mathbf{b} \wedge \mathbf{a},\\
bc\sin\alpha\, \mathbf{e}_1 \wedge \mathbf{e}_2 &= \mathbf{c} \wedge (-\mathbf{b}) = \mathbf{b} \wedge \mathbf{c} = \mathbf{b} \wedge (\mathbf{a} - \mathbf{b}) = \mathbf{b} \wedge \mathbf{a},\\
ac\sin\beta\, \mathbf{e}_1 \wedge \mathbf{e}_2 &= (-\mathbf{a}) \wedge (-\mathbf{c}) = \mathbf{a} \wedge \mathbf{c} = \mathbf{a} \wedge (\mathbf{a} - \mathbf{b}) = -\mathbf{a} \wedge \mathbf{b} = \mathbf{b} \wedge \mathbf{a}
\end{aligned}$$

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. □

Lemma 4.26. Sei durch $\mathbf{p}(t) := \mathbf{a} + t\mathbf{v}$ mit $\mathbf{a}, \mathbf{v} \in \mathbb{R}^n$ eine Parametergerade gegeben und sei $\mathbf{b} \in \mathbb{R}^n$ nicht auf der Gerade. Wir betrachten die Abstandsvektoren $\mathbf{d}(t) = \mathbf{b} - \mathbf{p}(t)$. Man erhält den kürzesten Abstand, wenn $\mathbf{p}(t)$ der Lotfußpunkt ist.

Beweis. Man ermittelt die Ableitung

$$d'(t) = \tfrac{\partial}{\partial t}|\mathbf{d}(t)| = \frac{1}{2|\mathbf{d}(t)|} 2\langle \mathbf{d}(t), \mathbf{d}'(t)\rangle = -\frac{\langle \mathbf{d}(t), \mathbf{v}\rangle}{|\mathbf{d}(t)|}.$$

Aus dem notwendigen Kriterium $d'(t) = 0$ ergibt sich $\langle \mathbf{d}, \mathbf{v}\rangle = 0$, womit $\mathbf{d}$ rechtwinklig auf $\mathbf{v}$ stehen muss. Es verbleibt zu zeigen, dass es sich um ein Minimum handelt. Wir bestimmen dazu die zweite Ableitung. Mit $d'(t)d(t) = -\langle \mathbf{d}(t), \mathbf{v}\rangle$ findet sich

$$\tfrac{\partial}{\partial t}(d'(t)d(t)) = d''(t)d(t) + d'(t)^2 = -\tfrac{\partial}{\partial t}\langle \mathbf{d}(t), \mathbf{v}\rangle = -\langle \mathbf{d}'(t), \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle.$$

Man erhält an der kritischen Stelle also

$$d''(t) = \frac{|\mathbf{v}|^2}{d(t)} > 0,$$

womit dort ein Minimum befindlich sein muss. □

Satz 4.27. Sei durch $\mathbf{p}(t) := \mathbf{a} + t\mathbf{v}$ mit $\mathbf{a}, \mathbf{v} \in \mathbb{R}^n$ eine Parametergerade gegeben und sei $\mathbf{b} \in \mathbb{R}^n$ ein weiterer Punkt. Der Abstand von $\mathbf{b}$ zur Gerade beträgt

$$d = \left| \mathbf{b} - \mathbf{a} - \frac{\langle \mathbf{b} - \mathbf{a}, \mathbf{v}\rangle}{|\mathbf{v}|^2}\mathbf{v} \right|.$$

Beweis 1. Es ist der Parameter t gesucht, bei dem der Abstandsvektor $\mathbf{d} := \mathbf{b} - \mathbf{p}(t)$ rechtwinklig zum Richtungsvektor $\mathbf{v}$ steht. Allgemein stehen zwei Vektoren genau dann rechtwinklig, wenn ihr Skalarprodukt verschwindet. Es muss also gelten

$$0 = \langle \mathbf{d}, \mathbf{v}\rangle = \langle \mathbf{b} - \mathbf{a} - t\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{b} - \mathbf{a}, \mathbf{v}\rangle - t\langle \mathbf{v}, \mathbf{v}\rangle, \iff t = \frac{\langle \mathbf{b} - \mathbf{a}, \mathbf{v}\rangle}{|\mathbf{v}|^2}.$$

Es findet sich der Abstand

$$d = |\mathbf{d}| = |\mathbf{b} - \mathbf{p}| = \left| \mathbf{b} - \mathbf{a} - \tfrac{\langle \mathbf{b} - \mathbf{a}, \mathbf{v}\rangle}{|\mathbf{v}|^2}\mathbf{v} \right|. \quad \square$$

Beweis 2. Wir verschieben alle Ortsvektoren um $-\mathbf{a}$, damit die Gerade durch den Ursprung verläuft. Aus dem verschobenen Punkt $\mathbf{b}' = \mathbf{b} - \mathbf{a}$ lässt sich nun der Lotfußpunkt vermittels der orthogonalen Projektion $P_{\mathbf{v}}(\mathbf{b}')$ erhalten. Der gesuchte Abstand ist die zwischen $\mathbf{b}'$ und dem Lotfußpunkt befindliche Distanz. Man erhält

$$d = |\mathbf{b}' - P_{\mathbf{v}}(\mathbf{b}')| = \left|\mathbf{b} - \mathbf{a} - \tfrac{\langle \mathbf{b}-\mathbf{a}, \mathbf{v}\rangle}{\langle \mathbf{v}, \mathbf{v}\rangle}\mathbf{v}\right|. \square$$

Korollar 4.28. Seien $\mathbf{a}, \mathbf{v} \in \mathbb{R}^n$ fest. Die beiden Mengen

$$\{\mathbf{p} \mid \exists t \colon \mathbf{p} = \mathbf{a} + t\mathbf{v}\}, \quad \{\mathbf{p} \mid (\mathbf{p} - \mathbf{a}) \wedge \mathbf{v} = 0\}$$

beschreiben dieselbe Gerade.

Beweis. Zu zeigen ist die Äquivalenz

$$(\exists t \colon \mathbf{p} - \mathbf{a} = t\mathbf{v}) \iff (\mathbf{p} - \mathbf{a}) \wedge \mathbf{v} = 0.$$

Die linke Seite verlangt, dass $\mathbf{p} - \mathbf{a}$ kollinear zu $\mathbf{v}$ sein soll. Das äußere Produkt zweier Vektoren verschwindet im Allgemeinen genau dann, wenn diese kollinear sind. Mithin ist die rechte Seite äquivalent zur linken. $\square$

Satz 4.29 (Lagrange-Identität).
Für zwei Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ gilt

$$|\mathbf{a} \wedge \mathbf{b}|^2 = \sum_{i,j:i<j} (a_i b_j - a_j b_i)^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 - \langle \mathbf{a}, \mathbf{b}\rangle^2.$$

Beweis. Es findet sich die Umformung

$$\begin{aligned}
&\sum_{i,j:i<j} (a_i b_j - a_j b_i)^2 = \frac{1}{2}\sum_{i,j}(a_i b_j - a_j b_i)^2 = \frac{1}{2}\sum_{i,j}(a_i^2 b_j^2 - 2a_i b_i a_j b_j + a_j^2 b_i^2) \\
&= \frac{1}{2}\Big(2\sum_{i,j} a_i^2 b_j^2 - 2\sum_{i,j} a_i b_i a_j b_j\Big) = \Big(\sum_{i=1}^n a_i^2\Big)\Big(\sum_{j=1}^n b_j^2\Big) - \Big(\sum_{i=1}^n a_i b_i\Big)\Big(\sum_{j=1}^n a_j b_j\Big) \\
&= |\mathbf{a}|^2 |\mathbf{b}|^2 - \langle \mathbf{a}, \mathbf{b}\rangle^2. \square
\end{aligned}$$

5 Algebra

5.1 Gruppentheorie

5.1.1 Grundlagen

Definition 5.1 (Gruppe). Das Tupel $(G, *)$ bestehend aus einer Menge G und Abbildung $*: G \times G \to \Omega$ 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 $e = e'$. □

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

$$a \overset{(G3)}{=} ae \overset{(G4)}{=} a(gb) \overset{(G2)}{=} (ag)b \overset{(G4)}{=} eb \overset{(G3)}{=} b.$$

Daher ist $a = b$. □

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. □

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. $a \in H \implies a^{-1} \in H$,

erüllt, 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. □

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 $f \circ f^{-1} = \mathrm{id}$ und $f^{-1} \circ f = \mathrm{id}$. Demnach sind alle Gruppenaxiome erfüllt. □

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 $\mathrm{Aut}(X)$ nennt. Sie ist eine Untergruppe der symmetrischen Gruppe $S(X)$.

Beweis. Offenbar ist $\mathrm{Aut}(X)$ eine nichtleere Teilmenge von $S(X)$. Die Verkettung zweier Automorphismen ist ebenfalls ein Automorphismus. Jeder Automorphismus besitzt einen inversen Automorphismus. Laut Untergruppenkriterium muss $\mathrm{Aut}(X)$ eine Untergruppe von $S(X)$ sein. □

Korollar 5.7. Sei K ein Körper. Die invertierbaren Matrizen $A \in K^{n \times n}$ bilden eine Gruppe, die man allgemeine lineare Gruppe $\mathrm{GL}(n, K)$ nennt. Sie ist kanonisch isomorph zur Automorphismengruppe $\mathrm{Aut}(K^n)$.

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

$$\varphi \colon K^{m \times n} \to \mathrm{Hom}(K^n, K^m), \quad \varphi(A)(\mathbf{v}) := A\mathbf{v}$$

bereits einen kanonischen Isomorphismus darstellt, erhält man bei Setzung $m = n$ und Einschränkung auf $\mathrm{GL}(n, K)$ einen Monomorphismus. Weil zu jeder bijektiven linearen Abbildung genau eine invertierbare Matrix gehört, muss

$$\varphi \colon \mathrm{GL}(n, K) \to \mathrm{Aut}(K^n)$$

außerdem surjektiv sein. □

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 $a(b + c) = ab + ac$. (Linksdistributivgesetz)
4. Für alle $a, b, c \in R$ gilt $(a + b)c = ac + bc$. (Rechtsdistributivgesetz)

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. □

Korollar 5.9. Sei R ein Ring und $a, b \in R$, dann gilt $(-a)b = -(ab) = a(-b)$.

Beweis. Man rechnet

$$\begin{aligned}(-a)b &= (-a)b + 0 = (-a)b + ab - (ab) = ((-a) + a)b - (ab)\\ &= 0b - (ab) = 0 - (ab) = -(ab).\end{aligned}$$ □

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

$$(-a)(-b) = -((-a)b) = -(-(ab)) = ab.$$ □

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

$$(m+n+1)a \overset{\text{(Def)}}{=} (m+n)a + a \overset{\text{(IV)}}{=} ma+na+a \overset{\text{(Def)}}{=} ma+(n+1)a.$$

Induktionsanfang ist $(m\cdot 0)a = 0a = 0 = m\cdot 0 = m(0a)$. Der Schritt ist

$$\begin{aligned}(m(n+1))a &= (mn+m)a = (mn)a + ma \overset{\text{(IV)}}{=} m(na)+ma \\ &= m((na)+a) \overset{\text{(Def)}}{=} m((n+1)a). \qquad \square\end{aligned}$$

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

$$((n+1)e)a \overset{\text{(Def)}}{=} (ne+e)a = nea + ea \overset{\text{(IV)}}{=} na + a \overset{\text{(Def)}}{=} (n+1)a.$$

Für $n > 0$ gilt zudem

$$((-n)e)a \overset{\text{(Def)}}{=} (-(ne))a = -((ne)a) = -(na) \overset{\text{(Def)}}{=} (-n)a. \qquad \square$$

5.2.2 Ringhomomorphismen

Definition 5.6 (Ringhomomorphismus). Seien R, R' Ringe. Eine Abbildung $\varphi: R \to R'$ 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 φ gilt $\varphi(kx) = k\varphi(x)$ für $k \in \mathbb{Z}$.

Beweis. Für $k > 0$ ist

$$\varphi(kx) = \varphi(\sum_{i=1}^{k} x) = \sum_{i=1}^{k} \varphi(x) = k\varphi(x).$$

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 $\Phi: \mathbb{R}[X] \to \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ mit $\Phi(f)(x) := f(x)$ nennt man Einsetzungshomomorphismus.

Satz 5.14. Beim Einsetzungshomomorphismus Φ 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 Φ 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

$$(\forall x \in \mathbb{R}\colon \Phi(f)(x) = \Phi(g)(x)) \implies f = g,$$

das heißt

$$(\forall x \in \mathbb{R}\colon \sum_k a_k x^k = \sum_k b_k x^k) \implies (\forall k\colon a_k = b_k).$$

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

$$(\forall x \in \mathbb{R}\colon \sum_{k=0}^{n} c_k x^k = 0) \implies (\forall k\colon c_k = 0).$$

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. □

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

$$a < b \implies a + c < b + c,$$
$$a > 0 \wedge b > 0 \implies ab > 0$$

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 $a^2 > 0$. □

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. □

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

$$(n+1)1_K \overset{\text{(Def)}}{=} n1_K + 1_K > 1_K > 0,$$

denn die Induktionsvoraussetzung $n1_K > 0$ impliziert $n1_K + 1_K > 1_K$ gemäß Def. 5.8 Axiom 1. □

Korollar 5.20. Sei K ein geordneter Körper. Für $a, b, c \in K$ gilt

$$a < b \wedge c > 0 \implies ca < cb,$$
$$a \leq b \wedge c \geq 0 \implies ca \leq cb.$$

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

$$a \leq b \iff a < b \vee a = b \implies ca < cb \vee ca = cb \iff ca \leq cb. \quad \square$$

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

$$(n+1)a \overset{\text{(Def)}}{=} na + a \geq na \geq 0,$$

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. □

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

$$0 \leq a \implies 0 \leq (n-m)a = na - ma \iff ma \leq na. \quad \square$$

Korollar 5.23. Sei K ein geordneter Körper und $a_k \in K$. Es gilt

$$(\exists k : a_k \neq 0) \implies \textstyle\sum_{k=1}^{n} a_k^2 > 0.$$

Beweis. Aufgrund der Prämisse existiert die Permutation, die das erste nichtverschwindende a_k mit a_1 vertauscht, weshalb ohne Beschränkung der Allgemeinheit $a_1 \neq 0$ angenommen werden darf. Laut Korollar 5.17 ist $a_k^2 \geq 0$ für jedes k. Induktion über die k. Der Induktionsanfang $a_1 > 0$ wurde bereits erläutert. Nun der Induktionsschritt. Mit dem Axiom 1 von Def. 5.8 und der Induktionsvoraussetzung $0 < \sum_{k=1}^{n-1} a_k^2$ findet sich

$$0 \leq a_n^2 < a_n^2 + \sum_{k=1}^{n-1} a_k^2 = \sum_{k=1}^{n} a_k^2. \quad \square$$

5.5 Formale Potenzreihen

Definition 5.9 (Formale Potenzreihe).
Sei R ein kommutativer Ring mit Einselement. Eine Folge $a\colon \mathbb{N}_{\geq 0} \to R$ bezeichnet man auch als formale Potenzreihe und schreibt

$$\sum_{k=0}^{\infty} a_k X^k := a = (a_k) = (a_k)_{k=0}^{\infty} = (a_0, a_1, a_2, \ldots).$$

Addition und Multiplikation von formalen Potenzreihen ist hierbei definiert als

$$\sum_{k=0}^{\infty} a_k X^k + \sum_{k=0}^{\infty} b_k X^k := \sum_{k=0}^{\infty} (a_k + b_k) X^k, \qquad \text{bzw. } (a_k) + (b_k) := (a_k + b_k),$$

$$\Big(\sum_{i=0}^{\infty} a_i X^i\Big)\Big(\sum_{j=0}^{\infty} b_j X^j\Big) := \sum_{k=0}^{\infty} \Big(\sum_{i=0}^{k} a_i b_{k-i}\Big) X^k, \qquad \text{bzw. } (a_i)(b_j) := (\textstyle\sum_{i=0}^{k} a_i b_{k-i}).$$

Korollar 5.24. Es gilt

$$\sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{k} \sum_{j=0}^{k} a_i b_j [k = i + j]. \qquad ([k = i + j] = \delta_{k,i+j})$$

Beweis. Man findet

$$\sum_{i=0}^{k} \sum_{j=0}^{k} a_i b_j [k{=}i{+}j] = \sum_{i=0}^{k} \sum_{j=0}^{k} a_i b_{k-i} [k{-}i{=}j] = \sum_{i=0}^{k} a_i b_{k-i} [0 \leq k{-}i \leq k] = \sum_{k=0}^{k} a_i b_{k-i}.$$

Die letzte Gleichung wird ersichtlich durch die äquivalente Umformung

$$0 \leq k - i \wedge k - i \leq k \iff i \leq k \wedge 0 \leq i. \qquad \square$$

Korollar 5.25. Formale Potenzreihen mit Koeffizienten in R bilden bezüglich ihrer definierten Addition und Multiplikation einen kommutativen Ring mit Einselement, der durch $R[[X]]$ symbolisiert wird.

Beweis. Addition und Multiplikation sind per Definition abgeschlossen. Das neutrale Element der Addition ist, wie unschwer zur erkennen, die Nullfolge. Das additiv inverse Element zu (a_k) ist $(-a_k)$, denn

$$(a_k) + (-a_k) = (a_k - a_k) = (0, 0, 0, \ldots) = 0.$$

Die Addition ist kommutativ, denn

$$(a_k) + (b_k) = (a_k + b_k) = (b_k + a_k) = (b_k) + (a_k).$$

Die Addition ist assoziativ, denn

$$\begin{aligned}((a_k) + (b_k)) + (c_k) &= (a_k + b_k) + (c_k) = (a_k + b_k + c_k) \\ &= (a_k) + (b_k + c_k) = (a_k) + ((b_k) + (c_k)).\end{aligned}$$

Die Multiplikation ist kommutativ, denn mit Korollar 6.16 findet sich

$$(a_k)(b_k) = (\textstyle\sum_{i=0}^{k} a_i b_{k-i}) = (\sum_{i=0}^{k} a_{k-i} b_{k-(k-i)}) = (\sum_{i=0}^{k} b_i a_{k-i}) = (b_k)(a_k).$$

Die Multiplikation ist assoziativ, denn

$$\begin{aligned}&((a_k)(b_k))(c_k) = (\textstyle\sum_{i=0}^{k}\sum_{j=0}^{k} a_i b_j [k{=}i{+}j])(c_k)\\&= \textstyle\sum_{i'=0}^{k}\sum_{j'=0}^{k}\sum_{i=0}^{i'}\sum_{j=0}^{i'} a_i b_j c_{j'} [i'{=}i{+}j][k{=}i'{+}j']\\&= \textstyle\sum_{j'=0}^{k}\sum_{i=0}^{k-j'}\sum_{j=0}^{k-j'} a_i b_j c_{j'} [k{=}i{+}j{+}j']\\&= \textstyle\sum_{j'=0}^{k}\sum_{i=0}^{k}\sum_{j=0}^{k} a_i b_j c_{j'} [k{=}i{+}j{+}j'].\end{aligned}$$

Die letzte Gleichung wird klar durch die Überlegung

$$i > k - j' \iff i > i + j + j' - j' \iff i > i + j \iff 0 > j \iff j < 0.$$

Das Distributivgesetz ist erfüllt, denn

$$\begin{aligned}&(c_k)((a_k)+(b_k)) = (c_k)(a_k+b_k) = (\textstyle\sum_{i=0}^{k} c_i(a_{k-i}+b_{k-i}))\\&= (\textstyle\sum_{i=0}^{k} c_i a_{k-i} + \sum_{i=0}^{k} c_i b_{k-i}) = (\textstyle\sum_{i=0}^{k} c_i a_{k-i}) + (\sum_{i=0}^{k} c_i b_{k-i})\\&= (c_k)(a_k) + (c_k)(b_k).\end{aligned}$$

Das Einselement ist $a_k = [k = 0]$, also $(a_k) = (1, 0, 0, 0, \ldots)$. □

Definition 5.10 (Formale Potenzreihen in zwei Variablen).
Man definiert $R[[X, Y]] := R[[X]][[Y]]$.

Definition 5.11 (Exponentialreihe). Man definiert

$$\exp(X) := \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

Korollar 5.26. Es gilt $\exp(X + Y) = \exp(X)\exp(Y)$.

Beweis. Mit der kurzen Vorbetrachtung $\frac{1}{n!}\binom{n}{k} = \frac{1}{n!}\frac{n!}{k!(n-k)!} = \frac{1}{k!(n-k)!}$ findet sich

$$\begin{aligned}\exp(X+Y) &= \sum_{n=0}^{\infty}\frac{(X+Y)^n}{n!} = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}X^k Y^{n-k} = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{X^k}{k!}\frac{Y^{n-k}}{(n-k)!}\\&= \Big(\sum_{n=0}^{\infty}\frac{X^n}{n!}\Big)\Big(\sum_{n=0}^{\infty}\frac{Y^n}{n!}\Big) = \exp(X)\exp(Y).\ \square\end{aligned}$$

Definition 5.12 (Reihen der Hyperbelfunktionen). Man definiert

$$\cosh(X) := \sum_{k=0}^{\infty}\frac{X^k}{(2k)!}, \quad \sinh(X) := \sum_{k=0}^{\infty}\frac{X^{2k+1}}{(2k+1)!}.$$

Korollar 5.27 (Paritätszerlegung der Exponentialreihe).
Es gilt $\exp(X) = \cosh(X) + \sinh(X)$.

Beweis. Mit der Zerlegung $1 = [k \in 2\mathbb{Z}] + [k \in 2\mathbb{Z}+1]$ findet sich

$$\begin{aligned}\exp(X) &= \sum_{k=0}^{\infty}\frac{X^k}{k!} = \sum_{k=0}^{\infty}\frac{X^k}{k!}[k\in 2\mathbb{Z}] + \sum_{k=0}^{\infty}\frac{X^k}{k!}[k\in 2\mathbb{Z}+1]\\&= \sum_{k=0}^{\infty}\frac{X^{2k}}{(2k)!} + \sum_{k=0}^{\infty}\frac{X^{2k+1}}{(2k+1)!} = \cosh(X)+\sinh(X).\ \square\end{aligned}$$

Definition 5.13 (Reihen der Winkelfunktionen). Man definiert

$$\cos(X) := \sum_{k=0}^{\infty} (-1)^k \frac{X^k}{(2k)!}, \quad \sin(X) := \sum_{k=0}^{\infty} (-1)^k \frac{X^k}{(2k+1)!}.$$

Korollar 5.28. Es gilt $\cosh(\mathrm{i}X) = \cos(X)$ und $\sinh(\mathrm{i}X) = \mathrm{i}\sin(X)$, wobei i die imaginäre Einheit ist.

Beweis. Aufgrund $\mathrm{i}^2 = -1$ gilt $\mathrm{i}^{2k} = (-1)^k$ und $\mathrm{i}^{2k+1} = (-1)^k \mathrm{i}$. Hiermit findet sich

$$\cosh(\mathrm{i}X) = \sum_{k=0}^{\infty} \frac{(\mathrm{i}X)^{2k}}{(2k)!} = \sum_{k=0}^{\infty} \frac{\mathrm{i}^{2k} X^{2k}}{(2k)!} = \sum_{k=0}^{\infty} (-1)^k \frac{X^{2k}}{(2k)!} = \cos(X),$$

$$\sinh(\mathrm{i}X) = \sum_{k=0}^{\infty} \frac{(\mathrm{i}X)^{2k+1}}{(2k+1)!} = \sum_{k=0}^{\infty} \frac{\mathrm{i}^{2k+1} X^{2k+1}}{(2k+1)!} = \mathrm{i} \sum_{k=0}^{\infty} (-1)^k \frac{X^{2k+1}}{(2k+1)!} = \mathrm{i}\sin(X). \quad \square$$

Korollar 5.29 (Eulersche Formel).
Es gilt $\exp(\mathrm{i}X) = \cos(X) + \mathrm{i}\sin(X)$, wobei i die imaginäre Einheit ist.

Beweis. Man findet

$$\exp(\mathrm{i}X) \overset{(1)}{=} \cosh(\mathrm{i}X) + \sinh(\mathrm{i}X) \overset{(2)}{=} \cos(X) + \mathrm{i}\sin(X).$$

Die Überlegung zu (1) verläuft auf direkte Weise analog zu der von Korollar 5.27. Die Umformung (2) gilt gemäß Korollar 5.28. $\square$

5.6 Zahlenbereiche

5.6.1 Die natürlichen Zahlen

Definition 5.14 (Natürliche Zahlen).
Man bezeichnet $\mathbb{N}$ bezüglich einer Nachfolgerabbildung $s\colon \mathbb{N} \to \mathbb{N}$ als die natürlichen Zahlen, wenn die folgenden Axiome erfüllt sind:

(P1) $0 \in \mathbb{N}$,
(P2) s ist injektiv,
(P3) $\forall n \in \mathbb{N}\colon s(n) \neq 0$,
(P4) $\forall M\colon 0 \in M \wedge (\forall n \in \mathbb{N}\colon n \in M \Rightarrow s(n) \in M) \Rightarrow \mathbb{N} \subseteq M$.

Definition 5.15 (Addition natürlicher Zahlen).
Man definiert rekursiv

$$a + 0 := a, \quad a + s(b) := s(a + b).$$

Korollar 5.30. Mit $1 := s(0)$ gilt $s(a) = a + 1$.

Beweis. Es findet sich $a + 1 = a + s(0) = s(a + 0) = s(a)$. $\square$

Korollar 5.31 (Assoziativgesetz der Addition).
Für alle $a, b, c \in \mathbb{N}$ gilt $(a + b) + c = a + (b + c)$.

Beweis. Induktion über c. Im Anfang $c = 0$ gilt

$$(a + b) + c = (a + b) + 0 = a + b = a + (b + 0) = a + (b + c).$$

Zum Schritt. Induktionsvoraussetzung sei $(a + b) + c = a + (b + c)$. Man findet

$$(a + b) + s(c) = s((a + b) + c) \overset{\text{IV}}{=} s(a + (b + c)) = a + s(b + c) = a + (b + s(c)). \quad \square$$

Korollar 5.32 (Neutrales Element der Addition).
Für alle $a \in \mathbb{N}$ gilt $0 + a = a + 0 = a$.

Beweis. Per Definition gilt $a + 0 = a$. Zu $0 + a$ per Induktion über a. Im Anfang $a = 0$ ist $0 + a = 0$ per Definition. Zum Schritt. Induktionsvorausetzung sei $0 + a = a$. Man findet

$$0 + s(a) = s(0 + a) \overset{\text{IV}}{=} s(a). \square$$

Korollar 5.33 (Kommutativgesetz der Addition).
Für alle $a, b \in \mathbb{N}$ gilt $a + b = b + a$.

Beweis. Zunächst $a + 1 = 1 + a$ per Induktion über a. Im Anfang $a = 0$ gilt die Aussage gemäß Korollar 5.32. Zum Schritt. Man findet

$$1 + s(a) = s(1 + a) \overset{\text{IV}}{=} s(a + 1) = a + s(1).$$

Nun $a + b = b + a$ per Induktion über b. Der Fall $b = 0$ gilt gemäß Korollar 5.32. Anfang sei $b = 1$. Dieser wurde zuvor gezeigt. Zum Schritt. Man findet

$$\begin{aligned} s(b) + a &= (b + 1) + a = b + (1 + a) = b + (a + 1) = b + s(a) \\ &= s(b + a) \overset{\text{IV}}{=} s(a + b) = a + s(b). \square \end{aligned}$$

Definition 5.16 (Multiplikation natürlicher Zahlen).
Man definiert rekursiv

$$a \cdot 0 := 0, \quad a \cdot s(b) := a \cdot b + a.$$

Korollar 5.34 (Distributivgesetz der Multiplikation).
Für alle $a, b, c \in \mathbb{N}$ gilt $(a + b)c = ac + bc$.

Beweis. Induktion über c. Im Anfang $c = 0$ resultieren beide Seiten der Gleichung im Wert 0. Zum Schritt. Man findet

$$(a + b)s(c) = (a + b)c + (a + b) \overset{\text{IV}}{=} ac + bc + a + b = ac + a + bc + b = as(c) + bs(c). \square$$

Lemma 5.35. Für alle $a \in \mathbb{N}$ gilt $0a = 0$.

Beweis. Induktion über a. Im Anfang $a = 0$ ist per Definition $0a = 0$. Der Schritt ist

$$0s(a) = 0a + 0 \overset{\text{IV}}{=} 0 + 0 = 0.$$

Korollar 5.36. Für alle $a \in \mathbb{N}$ gilt $a \cdot 1 = a$ und $1 \cdot a = a$.

Beweis. Die Formel $a \cdot 1 = a$ folgt unmittelbar aus der Definition und bereits bekannten Regeln.

Die Formel $1 \cdot a = a$ per Induktion über a. Im Anfang $a = 0$ folgt die Regel unmittelbar aus der Definition. Der Schritt ist

$$1s(a) = 1a + 1 \overset{\text{IV}}{=} a + 1 = s(a). \square$$

Korollar 5.37 (Kommutativgesetz der Multiplikation).
Für alle $a, b \in \mathbb{N}$ gilt $ab = ba$.

Beweis. Induktion über b. Im Anfang $b = 0$ gilt die Regel gemäß Lemma 5.35. Der Schritt ist

$$as(b) = ab + a \overset{\text{IV}}{=} ba + a = ba + 1a = (b+1)a = s(b)a. \quad \square$$

5.6.2 Die ganzen Zahlen

Definition 5.17 (Ganze Zahlen).
Man definiert die ganzen Zahlen $\mathbb{Z}$ als Quotientenmenge

$$\mathbb{Z} := (\mathbb{N} \times \mathbb{N})/\sim, \quad (x,y) \sim (x',y') :\iff x + y' = x' + y.$$

Satz 5.38. Die Menge $\mathbb{Z}$ bildet bezüglich den wohldefinierten Operationen

$$[(x,y)] + [(x',y')] := [(x + x', y + y')],$$
$$[(x,y)] \cdot [(x',y')] := [(xx' + yy', xy' + x'y)]$$

einen kommutativen unitären Ring.

Beweis. Wohldefiniert heißt, dass das Resultat der Operation nicht von den gewählten Repräsentanten abhängt. Sei dazu $(x,y) \sim (a,b)$ und $(x',y') \sim (a'b')$. Zu zeigen ist

$$(x + x', y + y') \sim (a + a', b + b'), \iff x + x' + b + b' = y + y' + a + a'.$$

Mit den Prämissen gilt $x + b = y + a$ und $x' + b' = y' + a'$, womit

$$(x + b) + (x' + b') = (y + a) + (y' + a'), \iff x + x' + b + b' = y + y' + a + a'.$$

Mit der Multiplikation verhält es sich ein wenig komplizierter. Zu Vereinfachung wird zunächst gezeigt:

$$\begin{aligned}
&[(x,y)] \cdot [(x',y')] = [(a,b)] \cdot [(x',y')] \\
&\iff (xx' + yy', xy' + yx') \sim (ax' + by', ay' + bx') \\
&\iff xx' + yy' + ay' + bx' = ax' + by' + xy' + yx' \\
&\iff (x + b)x' + (a + y)y' = (a + y)x' + (x + b)y'.
\end{aligned}$$

Diese Gleichung ist gemäß Voraussetzung $(x,y) \sim (a,b)$ bzw. $x + b = a + y$ erfüllt.
Analog bestätigt man

$$[(a,b)] \cdot [(x',y')] = [(a,b)] \cdot [(a',b')].$$

Gemäß Transitivität ergibt sich somit

$$[(x,y)] \cdot [(x',y')] = [(a,b)] \cdot [(a',b')].$$

Es ist nun zu bestätigen, dass $(\mathbb{Q}, +)$ eine kommutative Gruppe ist. Das Assoziativgesetz:

$$\begin{aligned}
&([(x,y)] + [(x',y')]) + [(x'',y'')] = [(x + x', y + y')] + [(x'',y'')] \\
&= [(x + x' + x'', y + y' + y'')] = [(x,y)] + [(x' + x'', y' + y'')] \\
&= [(x,y)] + ([(x',y')] + [(x'',y'')]).
\end{aligned}$$

Das neutrale Element ist $[(0,0)]$, denn

$$[(x,y)] + [(0,0)] = [(x+0, y+0)] = [(x,y)].$$

Das inverse Element zu $[(x,y)]$ ist $[(y,x)]$, denn es gilt

$$[(x,y)] + [(y,x)] = [(x+y, y+x)] = [(0,0)]$$
$$\iff (x+y, y+x) \sim (0,0) \iff x+y+0 = y+x+0.$$

Das Kommutativgesetz:

$$[(x,y)] + [(x',y')] = [(x+x', y+y')] = [(x'+x, y'+y)] = [(x',y')] + [(x,y)].$$

Es ist nun zu bestätigen, dass $(\mathbb{Q}, \cdot)$ ein kommutatives Monoid bildet. Das Assoziativgesetz:

$$([(x,y)] \cdot [(x',y')]) \cdot [(x'',y'')] = [(xx'+yy', xy'+x'y)] \cdot [(x'',y'')]$$
$$= [(xx'x''+x''yy'+xy'y''+x'yy'',\ xx'y''+yy'y''+xx''y'+x'x''y)]$$
$$= [(x,y)] \cdot [(x'x''+y'y'', x'y''+x''y')] = [(x,y)] \cdot ([(x',y')] \cdot [(x'',y'')]).$$

Das Kommutativgesetz:

$$[(x,y)] \cdot [(x',y')] = [(xx'+yy',\ xy'+yx')]$$
$$= [(x'x+y'y,\ x'y+xy')] = [(x',y')] \cdot [(x,y)].$$

Das neutrale Element ist $[(1,0)]$, denn es gilt

$$[(x,y)] \cdot [(1,0)] = [(x \cdot 1 + y \cdot 0,\ 1 \cdot y + x \cdot 0)] = [(x,y)].$$

Schließlich ist noch das Distributivgesetz zu bestätigen. Man findet

$$[(a,b)] \cdot ([(x,y)] + [(x',y')]) = [(a,b)] \cdot [(x+x', y+y')]$$
$$= [(ax+ax'+by+by',\ ay+ay'+bx+bx')]$$
$$= [(ax+by, ay+bx)] + [(ax'+by', ay'+bx')]$$
$$= [(a,b)] \cdot [(x,y)] + [(a,b)] \cdot [(x',y')].$$

Somit sind alle Axiome bestätigt. □

Definition 5.18 (Monoidhomomorphismus).
Seien $(M,+)$ und $(M',+')$ zwei Monoide. Eine Abbildung $\varphi\colon M \to M'$ heißt Monoidhomomorphismus, wenn für alle $a, b \in M$ gilt

$$\varphi(a+b) = \varphi(a) + \varphi(b)$$

und $\varphi(0) = 0'$ ist.

Satz 5.39 (Einbettung der natürlichen Zahlen in die ganzen).
Die Abbildung $\varphi\colon \mathbb{N}_0 \to \mathbb{Z}$ mit $\varphi(n) := [(n,0)]$ ist ein Monoidmonomorphismus.

Beweis. Es ergibt sich

$$\varphi(a+b) = [(a+b,0)] = [(a,0)] + [(b,0)] = \varphi(a) + \varphi(b).$$

Außerdem ist $\varphi(0) = [(0,0)]$, und $[(0,0)]$ ist das neutrale Element von $(\mathbb{Z},+)$.

Schließlich ist noch die Injektivität zu prüfen:

$$[(a,0)] = \varphi(a) = \varphi(b) = [(b,0)] \iff (a,0) \sim (b,0)$$
$$\iff a+0 = b+0 \iff a = b. \square$$

Bemerkung. Anstelle von $\varphi(n) = [(n,0)]$ darf man daher kurz $n = [(n,0)]$ schreiben. Außerdem definiert man $a-b := a+(-b)$. Daraus ergibt sich nun

$$[(x,y)] = [(x,0)] + [(0,y)] = [(x,0)] - [(y,0)] = x-y.$$

Die umständliche Schreibweise $[(x,y)]$ wird ab jetzt nicht mehr benötigt.

Definition 5.19 (Totalordnung der ganzen Zahlen).
Auf $\mathbb{Z}$ wird die folgende strenge Totalordnung definiert:

$$[(x,y)] < [(x',y')] :\iff x+y' < x'+y.$$

Satz 5.40 (Einbettung der Totalordnung).
Die Abbildung φ aus Satz 5.39 genügt der Forderung

$$a < b \implies \varphi(a) < \varphi(b).$$

Beweis. Nach den Definitionen ist

$$\varphi(a) < \varphi(b) \iff [(a,0)] < [(b,0)] \iff a+0 < 0+b \iff a < b. \square$$

5.6.3 Die rationalen Zahlen

Definition 5.20 (Rationale Zahlen).
Man definiert die rationalen Zahlen $\mathbb{Q}$ als Quotientenmenge

$$\mathbb{Q} := (\mathbb{Z} \times \mathbb{Z}_{\geq 1})/\sim, \quad (x,y) \sim (x',y') :\iff xy' = x'y.$$

Für die Äquivalenzklasse $[(x,y)] \in \mathbb{Q}$ schreiben wir im Folgenden $\frac{x}{y}$.

Satz 5.41. Die Menge $\mathbb{Q}$ bildet bezüglich den wohldefinierten Operationen

$$\frac{x}{y} + \frac{x'}{y'} := \frac{xy'+x'y}{yy'}, \quad \frac{x}{y} \cdot \frac{x'}{y'} := \frac{xx'}{yy'}$$

einen Körper.

Beweis. Wohldefiniert bedeutet, dass das Ergebnis der Operationen nicht von den gewählten Repräsentanten der Argumente abhängig ist. Sei dazu $(a,b) \sim (x,y)$ und $(a',b') \sim (x',y')$. Zu zeigen ist nun

$$(ab'+a'b, bb') \sim (xy'+x'y, yy')$$
$$\iff (ab'+a'b)(yy') = (xy'+x'y)(bb')$$
$$\iff ab'yy' + a'byy' = xy'bb' + x'ybb'.$$

Substituiert man $ay = xb$ und $a'y' = x'b'$ auf der linken Seite der Gleichung, ergibt sich die rechte Seite. Zu zeigen ist weiterhin

$$(aa', bb') \sim (xx', yy') \iff aa'yy' = xx'bb'.$$

Wieder wird linke Seite der Gleichung über $ay = xb$ und $a'y' = x'b'$ in die rechte Seite überführt. Die Wohldefiniertheit der Operationen ist damit bestätigt.

Es bleibt zu prüfen, dass $(\mathbb{Q}, +, \cdot)$ allen Körperaxiomen genügt. Das neutrale Element der Addition ist $0/1$, denn es gilt

$$\frac{x}{y} + \frac{0}{1} = \frac{x \cdot 1 + 0 \cdot y}{y \cdot 1} = \frac{x}{y}.$$

Das neutrale Element der Multiplikation ist $1/1$, denn es gilt

$$\frac{x}{y} \cdot \frac{1}{1} = \frac{x \cdot 1}{y \cdot 1} = \frac{x}{y}.$$

Die Assoziativität der Addition ergibt sich ohne größere Umstände:

$$\begin{aligned}\left(\frac{x}{y} + \frac{x'}{y'}\right) + \frac{x''}{y''} &= \frac{xy' + x'y}{yy'} + \frac{x''}{y''} = \frac{xy'y'' + x'yy'' + x''yy'}{yy'y''},\\ \frac{x}{y} + \left(\frac{x'}{y'} + \frac{x''}{y''}\right) &= \frac{x}{y} + \frac{x'y'' + x''y'}{y'y''} = \frac{xy'y'' + x'yy'' + x''yy'}{yy'y''}.\end{aligned}$$

Die Assoziativität der Multiplikation ist etwas einfacher:

$$\left(\frac{x}{y} \cdot \frac{x'}{y'}\right) \cdot \frac{x''}{y''} = \frac{xx'}{yy'} \cdot \frac{x''}{y''} = \frac{xx'x''}{yy'y''} = \frac{x}{y} \cdot \frac{x'x''}{y'y''} = \frac{x}{y} \cdot \left(\frac{x'}{y'} \cdot \frac{x''}{y''}\right).$$

Das Kommutativgesetz der Addition:

$$\frac{x}{y} + \frac{x'}{y'} = \frac{xy' + x'y}{yy'} = \frac{x'y + xy'}{y'y} = \frac{x'}{y'} + \frac{x}{y}.$$

Das Kommutativgesetz der Multiplikation:

$$\frac{x}{y'} \cdot \frac{x'}{y'} = \frac{xx'}{yy'} = \frac{x'x}{y'y} = \frac{x'}{y'} \cdot \frac{x}{y}.$$

Das additiv inverse Element zu x/y ist $(-x)/y$, denn es gilt

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

Das multiplikativ inverse Element zu x/y mit $x \neq 0$ ist y/x, denn es gilt

$$\frac{x}{y} \cdot \frac{y}{x} = \frac{xy}{xy} = \frac{1}{1}.$$

Schließlich findet bestätigt man noch das Distributivgesetz:

$$\begin{aligned}\frac{a}{b} \cdot \left(\frac{x}{y} + \frac{x'}{y'}\right) &= \frac{a}{b} \cdot \frac{xy' + x'y}{yy'} = \frac{axy' + ax'y}{byy'},\\ \frac{ax}{by} + \frac{ax'}{by'} &= \frac{axby' + ax'by}{byby'} = \frac{b}{b} \cdot \frac{axy' + ax'y}{byy'}.\end{aligned}$$

Hierbei beachtet man, dass $b/b = 1/1$ das multiplikativ neutrale Element ist. □

Satz 5.42 (Einbettung der ganzen Zahlen in die rationalen).
Sei $\varphi\colon \mathbb{Z} \to \mathbb{Q}$ mit $\varphi(z) := z/1$. Die Abbildung φ ist ein Eins-erhaltender Ringmonomorphismus.

Beweis. Die Erhaltung des Einselements ergibt sich trivial. Ferner findet man

$$\varphi(a+b) = \frac{a+b}{1} = \frac{a\cdot 1 + b\cdot 1}{1\cdot 1} = \frac{a}{1} + \frac{b}{1} = \varphi(a) + \varphi(b)$$

und

$$\varphi(ab) = \frac{ab}{1} = \frac{ab}{1\cdot 1} = \frac{a}{1}\cdot\frac{b}{1} = \varphi(a)\cdot\varphi(b). \qquad \square$$

Bemerkung. Vermittels der Einbettung können wir die ganze Zahl z ab jetzt mit der rationalen Zahl $z/1$ identifizieren. Das heißt, man schreibt einfach $z = z/1$ anstelle von $\varphi(z) = z/1$.

Definition 5.21 (Division rationaler Zahlen).
Wie in jedem Körper ist die Division für $a, b \in \mathbb{Q}$ definiert als $a/b := ab^{-1}$.

6 Kombinatorik

6.1 Endliche Mengen

6.1.1 Indikatorfunktion

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

$$\begin{aligned} &[A \wedge B] = [A][B], \\ &[A \vee B] = [A] + [B] - [A][B], \\ &[\neg A] = 1 - [A], \\ &[A \to B] = 1 - [A](1 - [B]). \end{aligned}$$

Beweis. Trivial mittels Wertetabelle. □

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 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

$$\begin{aligned} 1_{A\cup B}(x) &= [x \in A \cup B] = [x \in A \vee x \in B] = [x \in A] + [x \in B] - [x \in A][x \in B] \\ &= 1_A(x) + 1_B(x) - 1_{A\cap B}(x). \ \square \end{aligned}$$

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

$$\begin{aligned} |A \cup B| &= \sum_{x\in G} 1_{A\cup B}(x) = \sum_{x\in G} (1_A(x) + 1_B(x) - 1_{A\cap B}(x)) \\ &= \sum_{x\in G} 1_A(x) + \sum_{x\in G} 1_B(x) - \sum_{x\in G} 1_{A\cap B}(x) = |A| + |B| - |A \cap B|. \ \square \end{aligned}$$

Korollar 6.4. Für endliche Mengen A, B mit $A \subseteq B$ gilt $|A| \leq |B|$.

Beweis. Mithilfe der Indiktorfunktion findet sich

$$\begin{aligned} &A \subseteq B \iff (\forall x \colon 1_A(x) \leq 1_B(x)) \iff (\forall x \colon 0 \leq 1_B(x) - 1_A(x)) \\ &\implies 0 \leq \sum_{x\in B} (1_B(x) - 1_A(x)) = \sum_{x\in B} 1_B(x) - \sum_{x\in B} 1_A(x) = |B| - |A| \\ &\implies |A| \leq |B|. \ \square \end{aligned}$$

6.1.2 Endliche Abbildungen

Satz 6.5 (Anzahl der Abbildungen).
Seien X, Y endliche Mengen mit $|X| = k$ und $|Y| = n$. Die Menge der Abbildungen $X \to Y$ enthält n^k Elemente.

Beweis. Induktion über k. Im Anfang $k = 0$ ist $X = \emptyset$. Es gibt genau eine Abbildung $\emptyset \to Y$, nämlich die leere Abbildung. Gleichermaßen ist $n^0 = 1$.

Zum Induktionsschritt. Induktionsvoraussetzung sei die Gültigkeit für $k-1$. Es sei $|X| = k$ und $|Y| = n$. Gesucht ist die Anzahl der Möglichkeiten zur Festlegung der Abbildung $f : X \to Y$. Sei $x \in X$ fest. Für die Festlegung $f(x) = y$ bestehen nun genau n Möglichkeiten, nämlich so viele, wie es Elemente $y \in Y$ gibt. Für die Festlegung der übrigen Werte betrachtet man f als Abbildung

$$f : X \setminus \{x\} \to Y,$$

von denen es laut Voraussetzung n^{k-1} gibt. Wir haben also n mal n^{k-1} Möglichkeiten, das sind n^k. □

Satz 6.6 (Anzahl der Bijektionen).
Seien X, Y endliche Mengen, wobei $|X| = |Y| = n$ gelte. Die Menge der Bijektionen $X \to Y$ enthält $n!$ Elemente.

Beweis. Induktion über n. Im Anfang $n = 0$ ist $X = \emptyset$ und $Y = \emptyset$. Es existiert genau eine Bijektion $\emptyset \to \emptyset$, nämlich die leere Abbildung. Bei der Fakultät gilt ebenfalls $0! = 1$ laut Def. 6.8.

Zum Induktionsschritt. Induktionsvoraussetzung sei die Gültigkeit für $n-1$. Es sei $|X| = n$. Gesucht ist die Anzahl der Möglichkeiten zur Festlegung der Bijektion $f : X \to Y$. Sei $x \in X$ fest. Für die Festlegung $f(x) = y$ bestehen genau n Möglichkeiten, nämlich so viele, wie es Elemente $y \in Y$ gibt. Bei der Festlegung der übrigen Werte entfällt y aufgrund der Injektivität von f. Für die Festlegung betrachtet man f daher als Bijektion

$$f : X \setminus \{x\} \to Y \setminus \{y\},$$

von denen es laut Voraussetzung $(n-1)!$ gibt. Wir haben also n mal $(n-1)!$ Möglichkeiten, was gemäß Def. 6.8 gleich $n!$ ist. □

Satz 6.7 (Anzahl der Injektionen).
Seien X, Y endliche Mengen, wobei $|X| = k$ und $|Y| = n$ gelte. Die Menge der Injektionen $X \to Y$ enthält $n^{\underline{k}}$ Elemente.

Beweis. Induktion über k. Im Anfang $k = 0$ ist $X = \emptyset$. Es gibt genau eine Injektion $\emptyset \to Y$, nämlich die leere Abbildung. Gleichermaßen gilt $n^{\underline{0}} = 1$.

Zum Schritt. Voraussetzung sei die Gültigkeit für $k-1$. Es sei $|X| = k$ und $|Y| = n$. Gesucht ist die Anzahl der Möglichkeiten zur Festlegung der Injektion $f : X \to Y$. Sei $x \in X$ fest. Für die Festlegung $f(x) = y$ bestehen genau n Möglichkeiten, nämlich so viele, wie es Elemente $y \in Y$ gibt. Bei der Festlegung der übrigen entfällt y aufgrund der Injektivität von f. Für die Festlegung betrachtet man f daher als Injektion

$$f : X \setminus \{x\} \to Y \setminus \{y\},$$

von denen es laut Voraussetzung $(n-1)^{\underline{k-1}}$ gibt. Es sind also n mal $(n-1)^{\underline{k-1}}$ Möglichkeiten, was gleich $n^{\underline{k}}$ ist. □

Satz 6.8. Seien X, Y endliche Mengen und sei $|X| = k$. Sei $C_k(Y)$ die Menge der k-elementigen Teilmengen von Y. Für zwei Injektionen $X \to Y$ sei ferner die Äquivalenzrelation

$$f \sim g \;:\Longleftrightarrow\; \exists \pi \in S_k : f = g \circ \pi$$

definiert, wobei mit den $\pi \in S_k$ Permutationen gemeint sind. Zwischen der Quotientenmenge $\mathrm{Inj}(X,Y)/S_k$ und $C_k(Y)$ besteht eine kanonische Bijektion.

Beweis. Wir definieren diese Bijektion als

$$\varphi \colon \mathrm{Inj}(X,Y)/S_k \to C_k(Y), \quad \varphi([f]) := f(X),$$

wobei $[f] = f \circ S_k$ die Äquivalenzklasse des Repräsentanten f bezeichne. Sie ist wohldefiniert, denn für $f \sim g$ gilt

$$f(X) = (g \circ \pi)(X) = g(\pi(X)) = g(X).$$

Für die Injektivität von φ ist zu zeigen, dass $f(X) = g(X)$ die Aussage $[f] = [g]$ impliziert, also die Existenz einer Permutation π mit $f = g \circ \pi$. Weil g injektiv ist, existiert eine Linksinverse g^{-1}, so dass wir $\pi := g^{-1} \circ f$ wählen können. Es verbleibt somit die Gleichung $f = g \circ g^{-1} \circ f$ zu zeigen. Zwar ist g^{-1} im Allgemeinen keine Rechtsinverse, ihre Einschränkung auf $g(X)$ aber schon. Wegen $f(X) = g(X)$ hebt sich $g \circ g^{-1}$ daher wie gewünscht auf $f(X)$ weg.

Zur Surjektivität von φ. Hier ist zu zeigen, dass es zu jeder Menge $B \in C_k(Y)$ eine Injektion f mit $f(X) = B$ gibt. Betrachten wir sie als Bijektion $f : X \to B$. Eine solche besteht, weil X und B gleichmächtig sind. □

Satz 6.9 (Anzahl der Kombinationen).
Sei Y eine $|Y| = n$ Elemente enthaltende endliche Menge und $C_k(Y)$ die Menge der k-elementigen Teilmengen von Y. Es gilt $|C_k(Y)| = \binom{n}{k}$.

Beweis 1. Sei X eine Menge mit $|X| = k$. Es gilt

$$|C_k(Y)| \overset{(1)}{=} |\mathrm{Inj}(X,Y)/S_k| \overset{(2)}{=} \frac{|\mathrm{Inj}(X,Y)|}{|S_k|} = \frac{n^{\underline{k}}}{k!} = \binom{n}{k}.$$

Gleichung (1) gilt hierbei laut Satz 6.8. Die Einsicht von (2) erhält man mit der folgenden Überlegung. Für jede Gruppe G gilt die Bahnformel $|G| = |f \circ G| \cdot |G_f|$. Ist nun die Fixgruppe G_f trivial, ist $|G_f| = 1$ und infolge $|f \circ G| = |G|$. Dies ist bei der symmetrischen Gruppe $G = S_k$ der Fall. Aus diesem Grund enthält jede Bahn $f \circ S_k$ gleich viele Elemente, $|S_k|$ an der Zahl. Weil die Bahnen außerdem paarweise disjunkt sind, erhält man die Faktorisierung

$$|\mathrm{Inj}(X,Y)| = |S_k| \cdot |\mathrm{Inj}(X,Y)/S_k|. \; \square$$

Beweis 2. Induktion über (n,k). Im Anfang ist $k = 0$ oder $k = n$. Der abstruse Fall $k = 0$ sucht nach Teilmengen ohne Elemente. Es existiert genau eine solche Menge, nämlich die leere Menge, womit $C_0(Y) = 1$ ist. Der Fall $k = n$ sucht nach Teilmengen, die so viele Elemente haben wie Y. Dies kann nur Y selbst sein, womit $C_n(Y) = 1$ gilt. Gleichermaßen gilt $\binom{n}{0} = 1$ und $\binom{n}{n} = 1$.

Induktionsvoraussetzung sei die Gültigkeit für $(n-1, k-1)$ und $(n-1, k)$. Man nimmt nun ein Element y aus Y heraus, womit darin $n-1$ verbleiben. Entscheidet man sich, y zur Teilmenge hinzuzufügen, verbleiben noch $k-1$ Elemente auszuwählen.

Entscheidet man sich dagegen, verbleibt die Teilmenge unverändert, womit nach wie vor k Elemente auszuwählen sind. Die Anzahl der Möglichkeiten ist somit

$$|C_k(Y)| = |C_{k-1}(Y \setminus \{y\})| + |C_k(Y \setminus \{y\})| \overset{\text{IV}}{=} \binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}. \square$$

Satz 6.10 (Gitterweg-Interpretation).
Ein Gitterweg auf dem Gitter $\mathbb{Z} \times \mathbb{Z}$ heißt monoton, wenn von (x, y) aus lediglich der Schritt nach $(x+1, y)$ oder der Schritt nach $(x, y+1)$ gewährt ist. Die Anzahl der monotonen Gitterwege von $(0, 0)$ zu (x, y) beträgt

$$\frac{(x+y)!}{x!y!} = \binom{x+y}{x} = \binom{x+y}{y}.$$

Beweis 1. Alle Gitterwege besitzen dieselbe Länge $x + y$. Die Knoten des jeweiligen Wegs nummerieren wir der Reihe nach mit Ausnahme des letzten. Nun ist von den $x + y$ Nummern eine Teilmenge von y Nummern auszuwählen, an denen ein Schritt nach oben stattfinden soll. Dafür gibt es $\binom{x+y}{y}$ Möglichkeiten. $\square$

Beweis 2. Es bezeichne $f(x, y)$ die Anzahl der Wege von $(0, 0)$ zu (x, y). Zum Erreichen eines Randpunktes besteht immer nur eine einzige Möglichkeit, womit $f(x, 0) = 1$ und $f(0, y) = 1$ gilt. Der nicht auf dem Rand befindliche Punkt (x, y) kann nun von $(x-1, y)$ oder von $(x, y-1)$ aus erreicht werden, womit

$$f(x, y) = f(x-1, y) + f(x, y-1)$$

gelten muss. Man sieht nun, dass diese Rekursion ein gedrehtes pascalsches Dreieck erzeugt. Wir setzen daher $f(x, y) = C(x+y, x)$ und führen die Koordinatentransformation $x + y = n$ und $x = k$ aus. Die Rekurrenz nimmt damit die Form

$$\begin{aligned} &C(x+y, x) = C(x-1+y, x-1) + C(x+y-1, x) \\ \iff\ &C(n, k) = C(n-1, k-1) + C(n-1, k). \end{aligned}$$

an. Die Randbedingungen führen zu $C(n, n) = 1$ und $C(n, 0) = 1$. Durch diese Rekurrenz ist eindeutig der Binomialkoeffizient $C(n, k) = \binom{n}{k}$ bestimmt, womit

$$f(x, y) = C(x+y, x) = \binom{x+y}{x}$$

gelten muss. $\square$

6.2 Endliche Summen

6.2.1 Allgemeine Regeln

Definition 6.2 (Summe). Sei $(G, +, 0)$ eine kommutative Gruppe und $a_k \in G$. 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 6.11. 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. Im Anfang $n = m-1$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt:

$$\begin{aligned}\sum_{k=m}^{n} (a_k + b_k) &= a_n + b_n + \sum_{k=m}^{n-1} (a_k + b_k) \overset{\text{IV}}{=} a_n + b_n + \sum_{k=m}^{n-1} a_k + \sum_{k=m}^{n-1} b_k \\ &= \sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k. \square\end{aligned}$$

Korollar 6.12. Sei R ein Ring und $c, a_k \in R$. Sei c eine Konstante. Es gilt

$$\sum_{k=m}^{n} ca_k = c \sum_{k=m}^{n} a_k.$$

Beweis. Induktion über n. Im Anfang $n = m-1$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt:

$$\sum_{k=m}^{n} ca_k = ca_n + \sum_{k=m}^{n-1} ca_k \overset{\text{IV}}{=} ca_n + c \sum_{k=m}^{n-1} a_k = c(a_n + \sum_{k=m}^{n-1} a_k) = c \sum_{k=m}^{n} a_k. \square$$

Korollar 6.13 (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:

$$\sum_{k=m}^{p} a_k = \sum_{k=m}^{p-1} a_k + p_k = \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{p} a_k.$$

Induktionsschritt:

$$\sum_{k=m}^{n} a_k = a_n + \sum_{k=m}^{n-1} a_k \overset{\text{IV}}{=} a_n + \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{n-1} a_k = \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{n} a_k. \square$$

Korollar 6.14 (Indexshift).
Für die Indexverschiebung der Distanz $d \in \mathbb{Z}$ gilt

$$\sum_{k=m}^{n} a_k = \sum_{k=m+d}^{n+d} a_{k-d}.$$

Beweis 1. Induktion über n. Im Anfang $n = m-1$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt:

$$\sum_{k=m}^{n} a_k = a_n + \sum_{k=m}^{n-1} a_k \overset{\text{IV}}{=} a_{(n+d)-d} + \sum_{k=m+d}^{n+d-1} a_{k-d} = \sum_{k=m+d}^{n+d} a_{k-d}. \qquad \square$$

Beweis 2. Mit der Substitution $k = k' - d$ findet sich die Umformung

$$\sum_{k=m}^{n} a_k \overset{(1)}{=} \sum_{m \le k \le n} a_k \overset{(2)}{=} \sum_{m \le k'-d \le n} a_{k'-d} \overset{(3)}{=} \sum_{m+d \le k' \le n+d} a_{k'-d} \overset{(4)}{=} \sum_{k'=m+d}^{n+d} a_{k'-d},$$

wobei (1), (4) gemäß Korollar 6.19 gelten und (2), (3) eine andere Schreibweise für die Substitutionsregel 6.20 ist. $\square$

Bemerkung. Der zweite Beweis ist eigentlich zirkulär, weil der Beweis der Substitutionsregel über den Beweis von Korollar 6.18 in transitiver Abhängigkeit zum generalisierten Kommutativgesetz 6.17 steht, dessen Beweis einen Indexshift enthält.

Korollar 6.15. Es gilt

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

Beweis. Induktion über n und n'. Im Anfang bei $n = m-1$ und $n' = m-1$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt für n:

$$\sum_{i=m}^{n} \sum_{j=m'}^{n'} a_{ij} = \sum_{j=m'}^{n'} a_{nj} + \sum_{i=m}^{n-1} \sum_{j=m'}^{n'} a_{ij} \overset{\text{IV}}{=} \sum_{j=m'}^{n'} a_{nj} + \sum_{j=m'}^{n'} \sum_{i=m}^{n-1} a_{ij} = \sum_{j=m'}^{n'} (a_{nj} + \sum_{i=m}^{n-1} a_{ij}) = \sum_{j=m'}^{n'} \sum_{i=m}^{n} a_{ij}.$$

Induktionsschritt für n':

$$\sum_{i=m}^{n} \sum_{j=m'}^{n'} a_{ij} = \sum_{i=m}^{n} (a_{in'} + \sum_{j=m'}^{n'-1} a_{ij}) = \sum_{i=m}^{n} a_{in'} + \sum_{i=m}^{n} \sum_{j=m'}^{n'-1} a_{ij} \overset{\text{IV}}{=} \sum_{i=m}^{n} a_{in'} + \sum_{j=m'}^{n'-1} \sum_{i=m}^{n} a_{ij} = \sum_{j=m'}^{n'} \sum_{i=m}^{n} a_{ij}.$$

Weil immer ein Schritt nach rechts oder ein Schritt nach oben durchführbar ist, werden alle Punkte (n, n') im Gitter $\mathbb{Z}_{\ge m-1} \times \mathbb{Z}_{\ge m'-1}$ erreicht. $\square$

Korollar 6.16 (Umkehrung der Reihenfolge).
Es gilt $\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} a_{n-k}$.

Beweis. Induktion über n. Im Anfang $n = -1$ haben beide Seiten der Gleichung den Wert null. Der Induktionsschritt ist

$$\begin{aligned} \sum_{k=0}^{n} a_{n-k} &= a_{n-n} + \sum_{k=0}^{n-1} a_{n-k} \overset{\text{IV}}{=} a_0 + \sum_{k=0}^{n-1} a_{n-(n-1-k)} a_k \\ &= a_0 + \sum_{k=0}^{n-1} a_{k-1} \overset{(1)}{=} \sum_{k=0}^{0} a_k + \sum_{k=1}^{n} a_k \overset{(2)}{=} \sum_{k=0}^{n} a_k, \end{aligned}$$

wobei (1) gemäß Indexshift 6.14 und (2) gemäß Aufteilung 6.13 gilt. $\square$

Satz 6.17 (Generalisiertes Kommutativgesetz).
Sei $M = \{k \in \mathbb{Z} \mid m \leq k \leq n\}$. Für jede Permutation $\pi\colon M \to M$ gilt

$$\sum_{k=m}^{n} a_k = \sum_{k=m}^{n} a_{\pi(k)}.$$

Beweis. Induktiv. Sei ohne Beschränkung der Allgemeinheit $m = 1$. Im Induktionsanfang $n = 0$ und $n = 1$ ist die Gleichung offenkundig erfüllt.

Induktionsschritt. Induktionsvoraussetzung sei die Gültigkeit für M. Zu zeigen ist die Gültigkeit für $M \cup \{n+1\}$.

Sei t ein fester Parameter mit $1 \leq t \leq n+1$. Im Fall $\pi(t) = n+1$ geht man wie folgt vor. Man setze $\sigma(k) := \pi(k)$ für $1 \leq k \leq t-1$. Man setze $\sigma(k) := \pi(k+1)$ für $t \leq k \leq n$. Weil $n+1$ kein Wert von σ ist, muss σ eine Permutation $\sigma\colon M \to M$ sein. Ergo gilt

$$\begin{aligned}\sum_{k=1}^{n+1} a_{\pi(k)} &= \sum_{k=1}^{t-1} a_{\pi(k)} + a_{\pi(t)} + \sum_{k=t+1}^{n+1} a_{\pi(k)} = a_{\pi(t)} + \sum_{k=1}^{t-1} a_{\pi(k)} + \sum_{k=t}^{n} a_{\pi(k+1)} \\ &= a_{n+1} + \sum_{k=1}^{t-1} a_{\sigma(k)} + \sum_{k=t}^{n} a_{\sigma(k)} = a_{n+1} + \sum_{k=1}^{n} a_{\sigma(k)} \\ &\overset{\text{IV}}{=} a_{n+1} + \sum_{k=1}^{n} a_k = \sum_{k=1}^{n+1} a_k.\end{aligned}$$

Man beachte, dass in den beiden Randfällen $t = 1$ und $t = n+1$ die jeweilige Randsumme den Wert null hat und somit verschwindet. □

Definition 6.3. Für eine endliche Menge M definiert man

$$\sum_{k \in M} a_k := \sum_{i=m}^{n} a_{f(i)},$$

wobei $f\colon \{m, \ldots, n\} \to M$ eine frei wählbare Bijektion ist.

Korollar 6.18. Der Wert Summe auf der rechten Seite von Def. 6.3 ist unabhängig von der gewählten Bijektion.

Beweis. Seien f, g zwei solche Bijektionen. Dann existiert π mit $f = g \circ \pi$, womit

$$\sum_{i=m}^{n} a_{f(i)} = \sum_{i=m}^{n} a_{g(\pi(i))} = \sum_{i=m}^{n} a_{g(i)}$$

laut Satz 6.17 gilt. □

Korollar 6.19. Für $M = \{k \in \mathbb{Z} \mid m \leq k \leq n\}$ gilt

$$\sum_{m \leq k \leq n} a_k := \sum_{k \in M} a_k = \sum_{k=m}^{n} a_k.$$

Beweis. Es gilt $\sum_{k \in M} a_k = \sum_{k=m}^{n} a_{\mathrm{id}(k)} = \sum_{k=m}^{n} a_k$. □

Korollar 6.20 (Substitutionsregel). Ist $\varphi\colon M' \to M$ eine Bijektion, gilt

$$\sum_{k\in M} a_k = \sum_{k'\in M'} a_{\varphi(k')}.$$

Beweis. Zur Bijektion $f\colon \{1,\dots,|M|\} \to M$ existiert die Bijektion g mit $f = \varphi \circ g$. Infolge gilt

$$\sum_{k\in M} a_k = \sum_{i=1}^{|M|} a_{f(i)} = \sum_{i=1}^{|M|} a_{\varphi(g(i))} = \sum_{k'\in M'} a_{\varphi(k')}. \square$$

Korollar 6.21. Es gilt $\sum\limits_{k\in M} ca_k = c \sum\limits_{k\in M} a_k$ und $\sum\limits_{k\in M} (a_k + b_k) = \sum\limits_{k\in M} a_k + \sum\limits_{k\in M} b_k$.

Beweis. Laut Definition gilt

$$\sum_{k\in M} ca_k = \sum_{i=1}^{|M|} ca_{f(i)} = c\sum_{i=1}^{|M|} a_{f(i)} = c\sum_{k\in M} a_k,$$

$$\sum_{k\in M} (a_k + b_k) = \sum_{i=1}^{|M|} (a_{f(i)} + b_{f(i)}) = \sum_{i=1}^{|M|} a_{f(i)} + \sum_{i=1}^{|M|} b_{f(i)} = \sum_{k\in M} a_k + \sum_{k\in M} b_k. \square$$

Korollar 6.22. Es gilt

$$\sum_{k\in M}\sum_{l\in N} a_{kl} = \sum_{l\in N}\sum_{k\in M} a_{kl}.$$

Beweis. Laut Definition gilt

$$\sum_{k\in M}\sum_{l\in N} a_{kl} = \sum_{i=1}^{|M|}\sum_{j=1}^{|N|} a_{f(i),g(j)} = \sum_{j=1}^{|N|}\sum_{i=1}^{|M|} a_{f(i),g(j)} = \sum_{l\in N}\sum_{k\in M} a_{k,l}.$$

Korollar 6.23. Für $M \cap N = \varnothing$ gilt

$$\sum_{k\in M\cup N} a_k = \sum_{k\in M} a_k + \sum_{k\in N} a_k.$$

Beweis. Sei $m := |M|$ und $n := |N|$. Laut Prämisse existiert eine Bijektion $f\colon \{1,\dots,m+n\} \to M \cup N$ mit $f(i) \in M$ für $1 \le i \le m$ und $f(i) \in N$ für $m+1 \le i \le m+n$. Das macht

$$\sum_{k\in M\cup N} a_k = \sum_{i=1}^{m+n} a_{f(i)} = \sum_{i=1}^{m} a_{f(i)} + \sum_{i=m+1}^{m+n} a_{f(i)} = \sum_{k\in M} a_k + \sum_{k\in N} a_k. \square$$

Korollar 6.24. Für eine disjunkte Zerlegung $M = \bigcup_{i\in I} M_i$ gilt

$$\sum_{k\in M} a_k = \sum_{i\in I}\sum_{k\in M_i} a_k.$$

Beweis. Induktion über I. Im Anfang $I = \varnothing$ haben beide Seiten den Wert null. Induktionsvoraussetzung sei die Gültigkeit für I. Zu zeigen ist die Gültigkeit für $I \cup \{n\}$ mit $n \notin I$. Der Induktionsschritt ist

$$\sum_{k\in M_n\cup M} a_k = \sum_{k\in M_n} a_k + \sum_{k\in M} a_k \overset{\text{IV}}{=} \sum_{k\in M_n} a_k + \sum_{i\in I}\sum_{k\in M_i} a_k = \sum_{i\in I\cup\{n\}}\sum_{k\in M_i} a_k. \square$$

Korollar 6.25. Es gilt

$$\sum_{t\in M\times N} a_t = \sum_{k\in M}\sum_{l\in N} a_{(k,l)}.$$

Beweis. Es ist $M = \bigcup_{k\in M}\{k\}$ und weiter $M\times N = \bigcup_{k\in M}(\{k\}\times N)$ eine disjunkte Zerlegung. Hiermit findet sich die Umformung

$$\sum_{t\in M\times N} a_t \overset{(1)}{=} \sum_{k\in M}\sum_{t\in\{k\}\times N} a_t \overset{(2)}{=} \sum_{k\in M}\sum_{l\in N} a_{(k,l)},$$

wobei (1) laut Korollar 6.24 gilt und (2) per Substitutionsregel 6.20 mit der Bijektion $\varphi\colon N\to\{k\}\times N$ mit $\varphi(l) := (k,l)$ und $t = \varphi(l)$.

Korollar 6.26. Mit der Indikatorfunktion $1_A\colon M\to\{0,1\}$ für $A\subseteq M$ gilt

$$\sum_{k\in M} 1_A(k)a_k = \sum_{k\in A} a_k.$$

Beweis. Mit disjunkter Zerlegung $M = A\cup(M\setminus A)$ und Korollar 6.23 gilt

$$\sum_{k\in M} 1_A(k)a_k = \sum_{k\in A}\underbrace{1_A(k)}_{1}a_k + \sum_{k\in M\setminus A}\underbrace{1_A(k)}_{0}a_k = \sum_{k\in A} a_k. \qquad \square$$

Korollar 6.27. Allgemein gilt

$$\sum_{k\in A\cup B} a_k = \sum_{k\in A} a_k + \sum_{k\in B} a_k - \sum_{k\in A\cap B} a_k.$$

Beweis. Sei $G = A\cup B$ die Grundmenge. Gemäß Korollar 6.2 darf man rechnen

$$\begin{aligned}\sum_{k\in G} a_k &= \sum_{k\in G} 1_{A\cup B}(k)a_k = \sum_{k\in G} 1_A(k)a_k + \sum_{k\in G} 1_B(k)a_k - \sum_{k\in G} 1_{A\cap B}(k)a_k\\ &= \sum_{k\in A} a_k + \sum_{k\in B} a_k - \sum_{k\in A\cap B} a_k. \qquad \square\end{aligned}$$

Definition 6.4 (Differenzenfolge). Zu einer Folge (a_k) definiert man

$$(\Delta a)_k := a_{k+1} - a_k.$$

Korollar 6.28 (Teleskopsumme). Es gilt

$$\sum_{k=m}^{n-1} (\Delta a)_k = \sum_{k=m}^{n-1} (a_{k+1} - a_k) = a_n - a_m.$$

Beweis 1. Induktion über n. Im Anfang $n = m$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt:

$$\sum_{k=m}^{n} (a_{k+1} - a_k) = (a_{n+1} - a_n) + \sum_{k=m}^{n-1} (a_{k+1} - a_n) \overset{\text{IV}}{=} a_{n+1} - a_n + a_n - a_m = a_{n+1} - a_m. \qquad \square$$

Beweis 2. Per Indexshift 6.14 gilt $\sum_{k=m}^{n-1} a_{k+1} = \sum_{k=m+1}^{n} a_k = a_n - a_m + \sum_{k=m}^{n-1} a_k$. Somit ist

$$\sum_{k=m}^{n-1} (a_{k+1} - a_k) = \sum_{k=m}^{n-1} a_{k+1} - \sum_{k=m}^{n-1} a_k = a_n - a_m + \sum_{k=m}^{n-1} a_k - \sum_{k=m}^{n-1} a_k = a_n - a_m. \qquad \square$$

Korollar 6.29. Zum Beweis einer Formel

$$\sum_{k=m}^{n-1} a_k = s_n$$

genügt es, $s_m = 0$ und $(\Delta s)_n = a_n$ zu zeigen.

Beweis 1. Induktion über n. Im Anfang $n = m$ haben beide Seiten der Gleichung laut der Prämisse den Wert null. Induktionsschritt:

$$\sum_{k=m}^{n} a_k = a_n + \sum_{k=m}^{n-1} a_k \overset{\text{IV}}{=} a_n + s_n = (\Delta s)_n + s_n = s_{n+1} - s_n + s_n = s_{n+1}. \square$$

Beweis 2. Spezialisierung von Korollar 6.28. □

Korollar 6.30. Der Differenzoperator ist linear. Das heißt, für alle Folgen $(a_n), (b_n)$ und jede Konstante c gilt

$$\Delta(a+b) = \Delta a + \Delta b, \qquad ((a+b)_n := a_n + b_n)$$
$$\Delta(ca) = c\Delta a. \qquad ((ca)_n := ca_n)$$

Beweis. Man findet

$$\begin{aligned}(\Delta(a+b))_n &= (a+b)_{n+1} - (a+b)_n = (a_{n+1} + b_{n+1}) - (a_n + b_n) \\ &= a_{n+1} - a_n + b_{n+1} - b_n = (\Delta a)_n + (\Delta b)_n = (\Delta a + \Delta b)_n\end{aligned}$$

und

$$(\Delta(ca))_n = (ca)_{n+1} - (ca)_n = ca_{n+1} - ca_n = c(a_{n+1} - a_n) = c(\Delta a)_n = (c\Delta a)_n. \square$$

Definition 6.5 (Shiftoperator). Man definiert

$$(Ta)_n := a_{n+1}.$$

Korollar 6.31 (Iterierter Differenzoperator).
Für jede Folge (a_n) und $m \in \mathbb{Z}_{\geq 0}$ gilt

$$(\Delta^m a)_n = (-1)^m \sum_{k=0}^{m} \binom{m}{k} (-1)^k a_{n+k}.$$

Beweis. Es gilt $\Delta = T - \mathrm{id}$. Weil T und id kommutieren, ist der binomische Lehrsatz anwendbar. Es ergibt sich

$$\Delta^m = (T - \mathrm{id})^m = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} T^k \,\mathrm{id}^{m-k} = (-1)^m \sum_{k=0}^{m} \binom{m}{k} (-1)^k T^k. \square$$

Korollar 6.32. Sei f eine Polynomfunktion. Dann ist $\Delta_h f$ eine Polynomfunktion mit niedrigerem Grad.

Beweis. Für $f(x) = \sum_{n=0}^{m} a_n x^n$ gilt

$$\begin{aligned}\Delta_h f(x) &= f(x+h) - f(x) = \sum_{n=0}^{m} a_n (x+h)^n - \sum_{n=0}^{m} a_n x^n = \sum_{n=0}^{m} a_n ((x+h)^n - x^n) \\ &= \sum_{n=0}^{m} a_n (x^n + \sum_{k=0}^{n-1} \binom{n}{k} x^k h^{n-k} - x^n) = \sum_{n=0}^{m} a_n \sum_{k=0}^{n-1} \binom{n}{k} x^k h^{n-k}.\end{aligned}$$

In der Summe treten nur Monome bis x^{m-1} auf. □

Satz 6.33. Sei f ein Polynom vom Grad N. Für $n, a \in \mathbb{Z}$ und $n \geq a$ gilt

$$f(n) = \sum_{k=0}^{N} \frac{(\Delta^k f)(a)}{k!} (n-a)^{\underline{k}} = \sum_{k=0}^{N} \binom{n-a}{k} (\Delta^k f)(a).$$

Beweis. Es gilt $T = \Delta + \mathrm{id}$. Für jede nichtnegative ganze Zahl m gilt

$$T^m = (\Delta + \mathrm{id})^m = \sum_{k=0}^{m} \binom{m}{k} \Delta^k$$

mit dem binomischen Lehrsatz, da Δ und id kommutieren. Das macht

$$f(a+m) = \sum_{k=0}^{m} \binom{m}{k} (\Delta^k f)(a).$$

Man substituiere nun $n = a + m$. Für $n \geq a$ gilt dann

$$f(n) = \sum_{k=0}^{n-a} \binom{n-a}{k} (\Delta^k f)(a) = \sum_{k=0}^{N} \binom{n-a}{k} (\Delta^k f)(a).$$

Der Indexbereich der Summierung durfte auf bis $k = N$ geändert werden, weil $\Delta^k f = 0$ für $k > N$ laut Korollar 6.32 gilt. Dass nun Summanden mit $k > n - a$ auftreten können, ist nicht weiter schlimm, weil in diesem Fall $\binom{n-a}{k} = 0$ ist. □

6.2.2 Klassische Partialsummen

Korollar 6.34 (Partialsummen der konstanten Folge).
Es gilt $\sum_{k=m}^{n} 1 = n - m + 1$.

Beweis. Induktion über n. Im Anfang $n = m - 1$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt:

$$\sum_{k=m}^{n} 1 = 1 + \sum_{k=m}^{n-1} 1 \overset{\text{IV}}{=} 1 + n - 1 - m + 1 = n - m + 1. \square$$

Satz 6.35 (Partialsummen der arithmetischen Folge).
Es gilt $\sum_{k=0}^{n} k = \frac{n}{2}(n+1)$.

Beweis 1. Induktion über n. Im Anfang $n = -1$ haben beide Seiten der Gleichung den Wert null. Induktionsschritt:

$$\sum_{k=0}^{n} k = n + \sum_{k=0}^{n-1} k \overset{\text{IV}}{=} n + \frac{n-1}{2}(n-1+1) = \frac{n}{2}(2+n-1) = \frac{n}{2}(n+1). \square$$

Beweis 2. Klassischer Beweis. Man findet die Umformung

$$2\sum_{k=0}^{n} k = \sum_{k=0}^{n} k + \sum_{k=0}^{n} k \overset{(1)}{=} \sum_{k=0}^{n} k + \sum_{k=0}^{n} (n-k) \overset{(2)}{=} \sum_{k=0}^{n} (k+n-k) = \sum_{k=0}^{n} n \overset{(3)}{=} n \sum_{k=0}^{n} 1 \overset{(4)}{=} n(n+1),$$

wobei (1), (2), (3), (4) gemäß Korollar 6.16, 6.11, 6.12, 6.34 gelten. $\square$

Satz 6.36 (Partialsummen der geometrischen Folge).
Für $m \geq 0$ und $z \in \mathbb{C} \setminus \{1\}$ gilt $\sum_{k=m}^{n-1} z^k = \frac{z^n - z^m}{z-1}$.

Beweis. Induktion über n. Im Anfang $n = m - 1$ haben beiden Seiten der Gleichung den Wert null. Induktionsschritt:

$$\sum_{k=m}^{n} z^k = z^n + \sum_{k=m}^{n-1} z^k \overset{\text{IV}}{=} z^n + \frac{z^n - z^m}{z-1} = \frac{(z-1)z^n + z^n - z^m}{z-1} = \frac{z^{n+1} - z^m}{z-1}. \square$$

Korollar 6.37. Für $m \geq 0$ und $z \in \mathbb{C} \setminus \{1\}$ gilt

$$\sum_{k=m}^{n-1} k z^k = \frac{(nz^n - mz^m)(z-1) - (z^n - z^m)z}{(z-1)^2}.$$

Beweis. Die Gleichung von Satz 6.36 für $m \geq 1$ auf beiden Seiten nach z ableiten und anschließend beide Seiten mit z multiplizieren. Den Fall $m = 0$ und in diesem den Summand zu $k = 0$ explizit betrachten, sonst aber auf dieselbe Weise vorgehen. $\square$

Satz 6.38. Es gilt

$$\sum_{k=1}^{n}(-1)^k k = (-1)^n \left\lfloor \frac{n+1}{2} \right\rfloor.$$

Beweis. Induktion über n. Im Anfang $n = 0$ haben beiden Seiten den Wert null. Induktionsschritt:

$$\sum_{k=1}^{n}(-1)^k k = (-1)^n n + \sum_{k=1}^{n-1}(-1)^k k \overset{\text{IV}}{=} (-1)^n n + (-1)^{n-1} \left\lfloor \frac{n}{2} \right\rfloor = (-1)^n (n - \left\lfloor \frac{n}{2} \right\rfloor).$$

Zu zeigen verbleibt die Gleichung

$$n - \left\lfloor \frac{n}{2} \right\rfloor = \left\lfloor \frac{n+1}{2} \right\rfloor \iff n = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n+1}{2} \right\rfloor.$$

Wir nehmen die Fallunterscheidung zwischen geraden und ungeraden Zahlen vor, um Korollar 6.39 und 6.40 nutzen zu können. Im geraden Fall $n = 2k$ bestätigt sich

$$\left\lfloor \frac{2k}{2} \right\rfloor + \left\lfloor \frac{2k+1}{2} \right\rfloor = \lfloor k \rfloor + \left\lfloor k + \frac{1}{2} \right\rfloor = k + k = 2k.$$

Im ungeraden Fall $n = 2k + 1$ bestätigt sich

$$\left\lfloor \frac{2k+1}{2} \right\rfloor + \left\lfloor \frac{2k+1+1}{2} \right\rfloor = \left\lfloor k + \frac{1}{2} \right\rfloor + \lfloor k+1 \rfloor = k + k + 1 = 2k + 1. \square$$

6.3 Funktionen

6.3.1 Floor und Ceil

Definition 6.6 (Floorfunktion). Für $x \in \mathbb{R}$ definiert man

$$y = \lfloor x \rfloor \ :\iff y \in \mathbb{Z} \wedge 0 \leq x - y < 1.$$

Definition 6.7 (Ceilfunktion). Für $x \in \mathbb{R}$ definiert man

$$y = \lceil x \rceil \ :\iff y \in \mathbb{Z} \wedge 0 \leq y - x < 1.$$

Korollar 6.39. Für jede ganze Zahl k gilt $\lfloor k + x \rfloor = k + \lfloor x \rfloor$.

Beweis. Aufgrund der Prämisse $k \in \mathbb{Z}$ ist $y \in \mathbb{Z}$ äquivalent zu $y - k \in \mathbb{Z}$. Unter dieser Gegebenheit findet sich mit Def. 6.6 die äquivalente Umformung

$$\begin{aligned} y = \lfloor k + x \rfloor &\iff y \in \mathbb{Z} \wedge 0 \leq (k + x) - y < 1 \iff y - k \in \mathbb{Z} \wedge 0 \leq x - (y - k) < 1 \\ &\iff y - k = \lfloor x \rfloor \iff y = k + \lfloor x \rfloor. \square \end{aligned}$$

Korollar 6.40. Für $0 \leq x < 1$ gilt $\lfloor x \rfloor = 0$.

Beweis. Dies folgt unmittelbar aus Def. 6.6. $\square$

6.3.2 Faktorielle

Definition 6.8 (Fakultät).
Für eine nichtnegative ganze Zahl n definiert man $n!$ rekursiv durch

$$0! := 1, \quad (n+1)! := (n+1)n!.$$

Definition 6.9 (Fallende Faktorielle).
Für $k \in \mathbb{Z}_{\geq 0}$ und $n \in \mathbb{Z}$ (oder allgemeiner $n \in \mathbb{C}$) definiert man $n^{\underline{k}}$ rekursiv durch

$$n^{\underline{0}} := 1, \quad n^{\underline{k+1}} := n(n-1)^{\underline{k}}.$$

Definition 6.10 (Steigende Faktorielle).
Für $k \in \mathbb{Z}_{\geq 0}$ und $n \in \mathbb{Z}$ (oder allgemeiner $n \in \mathbb{C}$) definiert man $n^{\overline{k}}$ rekursiv durch

$$n^{\overline{0}} := 1, \quad n^{\overline{k+1}} := n(n+1)^{\underline{k}}.$$

Korollar 6.41. Für $n, k \in \mathbb{Z}_{\geq 0}$ und $k \leq n$ gilt

$$n^{\underline{k}} = \frac{n!}{(n-k)!}.$$

Beweis. Induktion über k. Im Anfang $k = 0$ resultieren beide Seiten der Gleichung im gleichen Wert 1. Der Induktionsschritt ist

$$n^{\underline{k}} = n(n-1)^{\underline{k-1}} \overset{\text{IV}}{=} n\frac{(n-1)!}{((n-1)-(k-1))!} = \frac{n(n-1)!}{(n-k)!} = \frac{n!}{(n-1)!}. \quad \square$$

Korollar 6.42. Für ganze Zahlen n, k mit $n \geq 1$ und $k \geq 1-n$ gilt

$$n^{\overline{k}} = \frac{(n+k-1)!}{(n-1)!}.$$

Beweis. Induktion über k. Im Anfang $k = 0$ resultieren beide Seiten der Gleichung im gleichen Wert 1. Der Induktionsschritt ist

$$\begin{aligned} n^{\overline{k}} &= n(n+1)^{\overline{k-1}} \overset{\text{IV}}{=} n\frac{(n+1+k-1-1)!}{(n+1-1)!} = \frac{n(n+k-1)!}{n!} \\ &= \frac{n(n+k-1)!}{n(n-1)!} = \frac{(n+k-1)!}{(n-1)!}. \quad \square \end{aligned}$$

6.3.3 Binomialkoeffizient

Definition 6.11 (Binomialkoeffizient).
Für $k \in \mathbb{Z}_{\geq 0}$ und $n \in \mathbb{Z}$ (oder allgemeiner $n \in \mathbb{C}$) definiert man

$$\binom{n}{k} := \frac{n^{\underline{k}}}{k!}.$$

Korollar 6.43. Für $n \in \mathbb{Z}$ mit $k \leq n$ gilt

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

Beweis. Folgt direkt aus Def. 6.11 und Korollar 6.41. □

Korollar 6.44. Für $k \geq 1$ und $n \in \mathbb{Z}$ (oder allgemeiner $n \in \mathbb{C}$) gilt

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}.$$

Beweis. Es findet sich die Umformung

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n(n-1)^{\underline{k-1}}}{k(k-1)!} = \frac{n}{k}\binom{n-1}{k-1}. \quad \square$$

Satz 6.45. Für $k \geq 1$ und $n \in \mathbb{Z}$ (oder allgemeiner $n \in \mathbb{C}$) gilt

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Beweis. Es findet sich die Umformung

$$\begin{aligned}\binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)^{\underline{k-1}}}{(k-1)!} + \frac{(n-1)^{\underline{k}}}{k!} = \frac{k(n-1)^{\underline{k-1}}}{k!} + \frac{(n-1)^{\underline{k-1}}(n-k)}{k!} \\ &= \frac{(n-1)^{\underline{k-1}}}{k!}(k+n-k) = \frac{n(n-1)^{\underline{k-1}}}{k!} = \frac{n^{\underline{k}}}{k!} = \binom{n}{k}. \quad \square\end{aligned}$$

7 Wahrscheinlichkeitsrechnung

7.1 Diskrete Wahrscheinlichkeitsräume

Definition 7.1 (Diskreter Wahrscheinlichkeitsraum).
Sei Ω eine höchstens abzählbare Menge. Das Paar (Ω, P) nennt man diskreten Wahrscheinlichkeitsraum, wenn

$$P\colon 2^{\Omega} \to [0,1], \quad P(A) := \sum_{\omega \in A} P(\{\omega\})$$

die Eigenschaft $P(\Omega) = 1$ besitzt.

Bemerkung: Man schreibt auch $P(\omega) := P(\{\omega\})$.

Definition 7.2 (Reelle Zufallsgröße).
Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum. Eine Funktion $X\colon \Omega \to \mathbb{R}$ 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

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)}$$

die bedingte Wahrscheinlichkeit von A, gegeben B.

Lemma 7.1. Seien A, B disjunkt. Seien die A_i disjunkt. Dann gilt

$$1_{A \cup B} = 1_A + 1_B,$$
$$1\{\bigcup_{i \in I} A_i\} = \sum_{i \in I} 1\{A_i\}.$$

Beweis. Es gilt

$$\begin{aligned} 1_{A \cup B}(\omega) &= [\omega \in A \cup B] = [\omega \in A \vee \omega \in B] \overset{(P)}{=} [\omega \in A \oplus \omega \in B] \\ &= [\omega \in A] + [\omega \in B] = 1_A(\omega) + 1_B(\omega). \end{aligned}$$

Die allgemeine Rechnung ist

$$1\{\bigcup_{i \in I} A_i\}(\omega) = [\omega \in \bigcup_{i \in I} A_i] = [\exists i \in I\colon \omega \in A_i] \overset{(P)}{=} \sum_{i \in I} [\omega \in A_i] = \sum_{i \in I} 1\{A_i\}(\omega).$$

Gleichung (P) gilt hierbei laut Prämisse. □

Korollar 7.2 (Additivität des Wahrscheinlichkeitsmaßes).
Seien die A_i paarweise disjunkte Mengen. Dann gilt

$$P(\bigcup_{i\in I} A_i) = \sum_{i\in I} P(A_i).$$

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

$$\begin{aligned} P(A \cup B) &= \sum_{\omega\in A\cup B} P(\{\omega\}) = \sum_{\omega\in\Omega} [\omega \in A \cup B] P(\{\omega\}) \\ &= \sum_{\omega\in\Omega} ([\omega \in A] + [\omega \in B]) P(\{\omega\}) = \sum_{\omega\in\Omega} [\omega \in A] P(\{\omega\}) + \sum_{\omega\in\Omega} [\omega \in B] P(\{\omega\}) \\ &= \sum_{\omega\in A} P(\{\omega\}) + \sum_{\omega\in B} P(\{\omega\}) = P(A) + B(B). \end{aligned}$$

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

$$\begin{aligned} P(\bigcup_{i\in I} A_i) &= \sum_{\omega\in\Omega} [\omega \in \bigcup_{i\in I} A_i] P(\{\omega\}) = \sum_{\omega\in\Omega} \sum_{i\in I} [\omega \in A_i] P(\{\omega\}) \\ &= \sum_{i\in I} \sum_{\omega\in\Omega} [\omega \in A_i] P(\{\omega\}) = \sum_{i\in I} \sum_{\omega\in A_i} P(\{\omega\}) = \sum_{i\in I} P(A_i). \quad \square \end{aligned}$$

Beweis 2. Laut Korollar 7.12, Lemma 7.1 und Korollar 7.8 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

$$P(\bigcup_{i\in I} A_i) = E(1\{\bigcup_{i\in I} A_i\}) = E(\sum_{i\in I} 1\{A_i\}) = \sum_{i\in I} E(1\{A_i\}) = \sum_{i\in I} P(A_i).$$

Korollar 7.3. Sei (Ω, 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, $(\forall A \colon P(A) \geq 0)$ und $P(\Omega) = 1$, gelten per Definition. Das dritte Axiom, die Additivität, gilt laut Korollar 7.2. Dass es sich bei 2^Ω 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 Ω in paarweise disjunkte nichtleere Mengen $B \in Z$. Dann gilt

$$P(A) = \sum_{B\in Z} P(A \mid B) P(B).$$

Beweis. Es gilt

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap \bigcup_{B\in Z} B) = P(\bigcup_{B\in Z} (A \cap B)) \\ &= \sum_{B\in Z} P(A \cap B) = \sum_{B\in Z} P(A \mid B) P(B). \quad \square \end{aligned}$$

Korollar 7.5 (Gesetz der totalen Wahrscheinlichkeit für Zufallsgrößen).
Seien $X, Y : \Omega \to \Omega'$ 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

$$\Omega = X^{-1}(Z) = X^{-1}(\bigcup_{x \in Z} \{x\}) = \bigcup_{x \in Z} X^{-1}(x)$$

gemäß Satz 1.42 (preimg-dl), und gemäß Korollar 1.45 ist das eine Vereinigung nichtleerer 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 (ω_k) eine beliebige Abzählung von Ω. 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. Für $g : \mathbb{R} \to \mathbb{R}$ gilt

$$E(g \circ X) = \sum_{x \in X(\Omega)} g(x) P(X^{-1}(x)) = \sum_{x \in X(\Omega)} g(x) P(X = x).$$

Beweis. Zunächst gilt

$$\sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P(\omega) = P(\bigcup_{\substack{\omega \in \Omega \\ X(\omega) = x}} \{\omega\}) = P(\{\omega \in \Omega \mid X(\omega) = x\}) = P(X^{-1}(x)).$$

Da die Reihe zu $E(g \circ X)$ nach Def. 7.4 absolut konvergent ist, darf sie beliebig umgeordnet werden und man bekommt

$$\begin{aligned} E(g \circ X) &= \sum_{\omega \in \Omega} g(X(\omega)) P(\omega) = \sum_{x \in X(\Omega)} \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} g(x) P(\omega) = \sum_{x \in X(\Omega)} g(x) \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P(\omega) \\ &= \sum_{x \in X(\Omega)} g(x) P(X^{-1}(x)). \quad \square \end{aligned}$$

Korollar 7.7. Es gilt

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

Beweis. Spezialisierung von Satz 7.6 mit $g := \mathrm{id}$. □

Korollar 7.8. 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

$$\begin{aligned} E(X + Y) &= \sum_{\omega \in \Omega} (X(\omega) + Y(\omega))P(\omega) = \sum_{\omega \in \Omega} (X(\omega)P(\omega) + Y(\omega)P(\omega)) \\ &= \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega) = E(X) + E(Y). \ \square \end{aligned}$$

Korollar 7.9. 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

$$X \leq Y \implies 0 \leq E(Y - X) = E(Y) - E(X) \iff E(X) \leq E(Y). \ \square$$

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 $X, Y : \Omega \to \mathbb{R}$ heißen unabhängig, wenn die Ereignisse $\{X \in A\}$ und $\{X \in B\}$ für alle Mengen $A, B \subseteq \mathbb{R}$ unabhängig sind.

Satz 7.10. Zwei Zufallsgrößen $X, Y : \Omega \to \mathbb{R}$ 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

$$P(\{X\in A\}\cap\{Y\in B\}) = P(\bigcup_{x\in A}\{X=x\}\cap\bigcup_{y\in B}\{Y=y\})$$
$$= P(\bigcup_{x\in A}\bigcup_{y\in B}(\{X=x\}\cap\{Y=y\})) = \sum_{x\in A}\sum_{y\in B}P(\{X=x\}\cap\{Y=y\})$$
$$= \sum_{x\in A}\sum_{y\in B}P(X=x)P(Y=y) = \sum_{x\in A}P(X=x)\sum_{y\in B}P(Y=y)$$
$$= P(\bigcup_{x\in A}\{X=x\})P(\bigcup_{y\in B}\{Y=y\}) = P(X\in A)P(Y\in B). \square$$

Definition 7.7 (Bedingter Erwartungswert).

$$E(X\mid A) = \frac{E(1_A X)}{P(A)} = \frac{1}{P(A)}\sum_{\omega\in A}X(\omega)P(\{\omega\}).$$

Satz 7.11. Es gilt

$$E(X\mid A) = \frac{1}{P(A)}\sum_x xP(\{X=x\}\cap A) = \sum_x xP(X=x\mid A),$$

wobei sich die Summe über alle $x\in X(\Omega)$ erstreckt.

Beweis. Man kann rechnen

$$E(1_A X) = \sum_{\omega\in\Omega}1_A(\omega)X(\omega)P(\{\omega\}) = \sum_x\sum_{\omega\in X^{-1}(x)}1_A(\omega)X(\omega)P(\{\omega\})$$
$$= \sum_x x\sum_{\omega\in X^{-1}(x)}1_A(\omega)P(\{\omega\}) = \sum_x x\sum_{\omega\in X^{-1}(x)\cap A}P(\{\omega\})$$
$$= \sum_x xP(X^{-1}(x)\cap A),$$

wobei $X^{-1}(x) = \{X=x\}$. $\square$

Korollar 7.12. Es gilt $P(A) = E(1_A)$, wobei 1_A die Indikatorfunktion ist.

Beweis. Gemäß Definition des Erwartungswertes ist

$$E(1_A) = \sum_{\omega\in\Omega}1_A(\omega)P(\{\omega\}) = \sum_{\omega\in A}P(\{\omega\}) = P(A). \square$$

Korollar 7.13. Es gilt $P(A\mid B) = E(1_A\mid B)$, wobei 1_A die Indikatorfunktion ist.

Beweis. Gemäß Definition 7.7 und Korollar 7.12 ist

$$E(1_A\mid B) = \frac{E(1_A 1_B)}{P(B)} = \frac{E(1_{A\cap B})}{P(B)} = \frac{P(A\cap B)}{P(B)} = P(A\mid B). \square$$

Korollar 7.14. Der Erwartungswert ist der Schwerpunkt der Wahrscheinlichkeitsmassen.

Beweis. Der Schwerpunkt s ist charakterisiert durch die folgende Gleichgewichtsbedingung, die wie folgt äquivalent umgeformt werden darf:

$$\sum_{x\in X(\Omega)}(s-x)P(X=x) = 0 \iff s\underbrace{\sum_{x\in X(\Omega)}P(X=x)}_{=1} - \sum_{x\in X(\Omega)}xP(X=x) = 0$$
$$\iff s = \sum_{x\in X(\Omega)}xP(X=x) \iff s = E(X). \square$$

7.2 Allgemeine Wahrscheinlichkeitsräume

Korollar 7.15. Der Erwartungswert ist der Schwerpunkt der Dichtefunktion.

Beweis. Der Schwerpunkt s ist charakterisiert durch die folgende Gleichgewichtsbedingung, die wie folgt äquivalent umgeformt werden darf:

$$\int_{\mathbb{R}} (s-x)f(x)\,\mathrm{d}x = 0 \iff s\underbrace{\int_{\mathbb{R}} f(x)\,\mathrm{d}x}_{=1} - \int_{\mathbb{R}} xf(x)\,\mathrm{d}x = 0$$

$$\iff s = \int_{\mathbb{R}} xf(x)\,\mathrm{d}x = E(X). \qquad \square$$

Satz 7.16. Sei $g\colon \mathbb{R} \to \mathbb{R}$ 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

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y)).$$

Gemäß der Kettenregel findet man

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}F_Y(y) = f_X(g^{-1}(y))\frac{\mathrm{d}}{\mathrm{d}y}g^{-1}(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

Sei g nun streng monoton fallend. Dann kann man rechnen

$$F_Y(y) = P(g(X) \le y) = P(X \ge g^{-1}(y)) = 1 - P(X < g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

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 $\operatorname{sgn}(g'(x))$ konstant für alle x und wir haben allgemein

$$f_Y(y) = \operatorname{sgn}(g'(x))\frac{f_X(x)}{g'(x)} = \frac{f_X(x)}{|g'(x)|}$$

mit $x = g^{-1}(y)$. $\square$

Satz 7.17 (»LOTUS: Law of the unconscious statistican«).
Ist $g\colon \mathbb{R} \to \mathbb{R}$ streng monoton, dann muss gelten

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f_X(x)\,\mathrm{d}x.$$

Beweis. Sei $Y = g(X)$. Mit Satz 7.16 und Substitution $y = g(x)$ kann man rechnen

$$\begin{aligned} E(g(X)) = E(Y) &= \int_{-\infty}^{\infty} yf_Y(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} y\frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}\,\mathrm{d}y \\ &= \int_{g^{-1}(-\infty)}^{g^{-1}(\infty)} g(x)\frac{f_X(x)}{|g'(x)|}g'(x)\,\mathrm{d}x \\ &= \operatorname{sgn}(g')\int_{-\operatorname{sgn}(g')\infty}^{\operatorname{sgn}(g')\infty} g(x)f_X(x)\mathrm{d}x = \int_{-\infty}^{\infty} g(x)f_X(x)\mathrm{d}x. \qquad \square \end{aligned}$$

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 7.18 (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

$$Q := P'(0) = \lim_{h \to 0} \frac{P(h) - E}{h}$$

die Intensitätsmatrix, wobei mit E die Einheitsmatrix gemeint ist.

Bemerkung. Weil $P(t)$ für $t \geq 0$ definiert ist, stimmt der gewöhnliche Grenzwert für $h \to 0$ mit dem rechtsseitigen Grenzwert für $h \searrow 0$ überein.

Satz 7.19. Der Grenzwert in Def. 7.10 existiert.

Korollar 7.20. 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 7.19 an der Stelle 0 differenzierbar ist, besteht die Linearisierung

$$P(h) = P(0) + hP'(0) + o(h) = E + hQ + o(h). \square$$

Satz 7.21 (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 7.19 gilt zudem

$$P'(t) = \lim_{h\to 0} \frac{P(t+h) - P(t)}{h} = \lim_{h\to 0} \frac{P(t)P(h) - P(t)}{h} = P(t) \lim_{h\to 0} \frac{P(h) - E}{h} = P(t)Q. \square$$

Satz 7.22. Die kolmogorowsche Vorwärtsgleichung besitzt die eindeutige Lösung $P(t) = \mathrm{e}^{tQ}$, wobei mit $\mathrm{e}^A = \exp(A)$ das Matrixexponential gemeint ist.

Beweis. Analog zur gewöhnlichen Exponentialfunktion gilt

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tQ} = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{\infty}\frac{(tQ)^k}{k!} = \sum_{k=1}^{\infty}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t^kQ^k}{k!}\right) = \sum_{k=1}^{\infty} k\frac{t^{k-1}Q^k}{k!}$$
$$= \sum_{k=0}^{\infty}(k+1)\frac{t^kQ^{k+1}}{(k+1)!} = \sum_{k=0}^{\infty}\frac{t^kQ^{k+1}}{k!} = \mathrm{e}^{tQ}Q = Q\mathrm{e}^{tQ}.$$

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 $P(t) = R(t)\mathrm{e}^{tQ}$ schreiben mit $R(t) := P(t)\mathrm{e}^{-tQ}$. Da P die Differentialgleichung erfüllen soll, muss gelten

$$0 = P'(t) - P(t)Q = R'(t)\mathrm{e}^{tQ} + R(t)\mathrm{e}^{tQ}Q - R(t)\mathrm{e}^{tQ}Q = R'(t)\mathrm{e}^{tQ}.$$

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 $P(t) = \mathrm{e}^{tQ}$. $\square$

Korollar 7.23. 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

$$0 = \tfrac{\mathrm{d}}{\mathrm{d}t}1 = \sum_{j\in S} p'_{ij}(t) = \sum_{j\in S} q_{ij}. \square$$

Korollar 7.24 (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 7.23 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.

7.4 Mathematische Statistik

7.4.1 Schätzfunktionen

Definition 7.11 (Arithmetischer Mittelwert).
Das arithmetische Mittel von unabhängigen und identisch verteilten Zufallsgrößen $X_k : \Omega \to \mathbb{R}$ ist definiert als

$$\overline{X} := \tfrac{1}{n} \textstyle\sum_{k=1}^{n} X_k.$$

Satz 7.25. Sei $X_k = w + \varepsilon_k$ eine Streuung um einen festen wahren Wert w. Die Streuung sei unverzerrt, das heißt, es gelte $E(\varepsilon_k) = 0$ für alle k. Dann ist das arithmetische Mittel der X_k ein erwartungstreuer Schätzer von w, das heißt, es gilt $E(\overline{X}) = w$.

Beweis. Aufgrund von $E(X_k) = E(w + \varepsilon_k) = w + E(\varepsilon_k) = w$ gilt

$$E(\overline{X}) = E\left(\frac{1}{n} \sum_{k=1}^{n} X_k\right) = \frac{1}{n} \sum_{k=1}^{n} E(X_k) = \frac{1}{n} \sum_{k=1}^{n} w = \frac{1}{n} n w = w. \quad \square$$

8 Zahlentheorie

8.1 Kongruenzen und Teilbarkeit

Definition 8.1 (Teiler). Für $a, m \in \mathbb{Z}$ ist die Relation »m teilt a« definiert als

$$m \mid a \ :\Longleftrightarrow\ \exists k \in \mathbb{Z} \colon a = km.$$

Definition 8.2 (Kongruenz). Für $a, b, m \in \mathbb{Z}$ ist die Relation »a ist kongruent zu b modulo m« definiert als

$$a \equiv b \pmod{m} \ :\Longleftrightarrow\ m \mid (a - b).$$

Korollar 8.1. Es gilt $m \mid a$ genau dann, wenn $a \equiv 0 \pmod{m}$.

Beweis. Spezialisierung von Def. 8.2 mit $b := 0$. □

Definition 8.3 (Teilermenge). Die Teilermenge einer ganzen Zahl a ist definiert durch

$$m \in T(a) \ :\Longleftrightarrow\ m \mid a.$$

Bemerkung. Wir setzen $T_{\geq 1}(a) := T(a) \cap \mathbb{Z}_{\geq 1}$ und $T_{\geq 0}(a) := T(a) \cap \mathbb{Z}_{\geq 0}$.

Korollar 8.2. Es gilt

$$a \mid b \iff T(a) \subseteq T(b).$$

Beweis. Zur Implikation von rechts nach links. Die Aussage $T(a) \subseteq T(b)$ ist laut ihrer Definition äquivalent zu $\forall m \colon m \mid a \Rightarrow m \mid b$. Die Anwendung dieser Prämsse auf den Ansatz $a \mid a$ liefert sofort $a \mid b$.

Zur Implikation von links nach rechts. Expandiert man die Aussage durch Einsetzen der Definitionen, ist

$$(\exists k \colon b = ka) \implies (\forall m \colon (\exists x \colon a = xm) \implies (\exists y \colon b = ym)).$$

zu zeigen. Wir haben einen Zeugen k für $b = ka$ gegeben. Gemäß Prämisse liegt ein Zeuge x für $a = xm$ vor, wonach $ka = kxm$ gilt, also $b = kxm$. Ergo ist $y := kx$ ein Zeuge für $\exists y \colon b = ym$. □

Korollar 8.3. Die Teilerrelation ist transitiv, das heißt, $a \mid b$ und $b \mid c$ impliziert $a \mid c$.

Beweis. Unter Nutzung von Korollar 8.2 nimmt die Aussage die Gestalt

$$T(a) \subseteq T(b) \land T(b) \subseteq T(c) \implies T(a) \subseteq T(c)$$

an. Diese Aussage ist nun aber offenkundig, da die Relation »ist Teilmenge von« eine transitive ist. □

Korollar 8.4. Gilt $m \mid a$ und $b \in \mathbb{Z}$, so gilt auch $m \mid (ab)$.

Beweis. Die Prämisse liefert einen Zeugen x mit $a = xm$, womit $ab = bxm$ gilt. Ergo existiert mit $y := bx$ ein Zeuge für $\exists y : ab = ym$. □

Korollar 8.5. Die Kongruenzrelation ist eine Äquivalenzrelation. Das heißt, es gilt

$$\begin{aligned} &a \equiv a \pmod m, && \text{(Reflexivität)}\\ &a \equiv b \implies b \equiv a \pmod m, && \text{(Symmetrie)}\\ &a \equiv b \wedge b \equiv c \implies a \equiv c \pmod m. && \text{(Transitivität)} \end{aligned}$$

Beweis. Zur Reflexivität. Es gilt die äquivalente Umformung

$$a \equiv a \pmod m \iff m \mid (a-a) \iff m \mid 0 \iff (\exists k : 0 = km).$$

Die letzte Aussage ist durch $k = 0$ erfüllt.

Zur Symmetrie. Die Prämisse liefert einen Zeugen x mit $a - b = xm$, womit $b - a = -xm$ gilt. Ergo ist $y := -x$ ein Zeuge für die Existenzaussage $\exists y : b - a = ym$.

Zur Reflexivität. Die Prämissen liefern Zeugen x mit $a - b = xm$ und y mit $b - c = ym$. Infolge gilt

$$a - c = (a-b) + (b-c) = xm + ym = (x+y)m.$$

Ergo ist $z := x + y$ ein Zeuge für $\exists z : a - c = zm$. □

Korollar 8.6. Für ganze Zahlen a, b, c gilt

$$\begin{aligned} a \equiv b \pmod m &\iff a + c \equiv b + c \pmod m,\\ a \equiv b \pmod m &\iff a - c \equiv b - c \pmod m. \end{aligned}$$

Beweis. Unter Nutzung der Definition findet sich die äquivalente Umformung

$$\begin{aligned} a + c \equiv b + c \pmod m &\iff m \mid ((a+c)-(b+c)) \iff m \mid (a-b)\\ &\iff a \equiv b \pmod m. \end{aligned}$$

Der Beweis der zweiten Aussage verläuft analog. □

Korollar 8.7. Für ganze Zahlen a, b, c gilt

$$a \equiv b \pmod m \implies ac = bc \pmod m.$$

Beweis. Die Prämisse liefert einen Zeugen x mit $a - b = xm$, womit $ac - bc = cxm$ gilt. Ergo existiert mit $y := cx$ ein Zeuge für $\exists y : ac - bc = ym$. □

Korollar 8.8. Gilt $a \equiv a' \pmod m$ und $b \equiv b' \pmod m$, so gilt auch

$$\begin{aligned} &a + b \equiv a' + b' \pmod m,\\ &a - b \equiv a' - b' \pmod m,\\ &ab \equiv a'b' \pmod m. \end{aligned}$$

Beweis. Laut Prämisse liegen Zeugen x mit $a = a' + xm$ und y mit $b = b' + ym$ vor. Infolge gilt

$$a + b = a' + xm + b' + ym = a' + b' + (x+y)m.$$

Ergo existiert mit $z := x + y$ ein Zeuge für $\exists z : (a + b) = (a' + b') + zm$. Der Beweis der zweiten Regel verläuft analog. Bei der Multiplikation gilt

$$ab = (a' + xm)(b' + ym) = a'b' + a'ym + b'xm + xym^2 = a'b' + (a'y + b'x + xym)m.$$

Ergo existiert mit $z := a'y + b'x + xym$ ein Zeuge für $\exists z : ab = a'b' + zm$. □

Korollar 8.9. Gilt $a \equiv a' \pmod{m}$ und $n \in \mathbb{Z}_{\geq 0}$, so gilt auch $a^n \equiv (a')^n \pmod{m}$.

Beweis. Induktion über n mit Induktionsanfang bei $n = 0$. Mit $a^0 = 1$ und $(a')^0 = 1$ wird die Behauptung in diesem Fall zu $1 \equiv 1$, die aufgrund der Reflexivität gilt. Induktionsschritt. Wendet man Korollar 8.8 auf die Prämisse $a \equiv a'$ und die Induktionsvoraussetzung $a^n \equiv (a')^n$ an, findet sich $aa^n \equiv a'(a')^n$, also $a^{n+1} \equiv (a')^{n+1}$. □

Korollar 8.10. Gilt $a_k \equiv a'_k \pmod{m}$ für alle k, so gilt auch $\sum_{k=0}^{n-1} a_k \equiv \sum_{k=0}^{n-1} a'_k \pmod{m}$.

Beweis. Induktion über n. Für $n = 0$ wird die Behauptung zu $0 \equiv 0$, die aufgrund der Reflexivität gilt. Induktionsschritt. Wendet man Korollar 8.8 auf die Prämisse $a_n \equiv a'_n$ und die Induktionsvoraussetzung $\sum_{k=0}^{n-1} a_k \equiv \sum_{k=0}^{n-1} a'_k$ an, folgt

$$\sum_{k=0}^{n-1} a_k = a_n + \sum_{k=0}^{n-1} a_k \equiv a'_n + \sum_{k=0}^{n-1} a'_k = \sum_{k=0}^{n-1} a'_k. \qquad \square$$

Korollar 8.11. Sei $p \in \mathbb{Z}[X]$, also ein Polynom mit ganzzahligen Koeffizienten. Gilt $x \equiv x' \pmod{m}$, so gilt auch $p(x) \equiv p(x') \pmod{m}$.

Beweis. Laut Korollar 8.9 gilt $x^k \equiv (x')^k$ für jedes $k \geq 0$. Im weiteren Fortgang gilt $a_k x^k \equiv a_k (x')^k$ wegen Korollar 8.6. Mit Korollar 8.10 erhält man schließlich

$$p(x) = \sum_{k=0}^{n} a_k x^k \equiv \sum_{k=0}^{n} a_k (x')^k = p(x'). \qquad \square$$

Satz 8.12. Für jede ganze Zahl n gilt $2 \mid n \iff 2 \mid n^2$.

Beweis. Die Implikation von links nach rechts gilt gemäß Korollar 8.4 mit $a := n$ und $b := n$. Zur Implikation von rechts nach links. Wir zeigen die Kontraposition

$$\neg(2 \mid n) \implies \neg(2 \mid n^2).$$

Laut Prämisse ist n ungerade, also von der Form $n = 2k + 1$ mit $k \in \mathbb{Z}$. Nun gilt

$$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

wonach n^2 ebenfalls ungerade sein muss. □

8.2 Primzahlen

Definition 8.4 (Teilerfunktion).
Für eine positive ganze Zahl n definiert man

$$\sigma_k(n) := \sum_{d \mid n} d^k,$$

wobei mit $d \mid n$ die positiven Teiler $d \in T_{\geq 1}(n)$ gemeint sind.

Definition 8.5 (Primzahl).
Eine positive ganze Zahl n wird Primzahl genannt, wenn sie zwei unterschiedliche positive Teiler besitzt, womit $\sigma_0(n) = 2$ gemeint ist.

Korollar 8.13. Eine Zahl n ist genau dann eine Primzahl, wenn $n \geq 2$ ist und ihre einzigen beiden positiven Teiler 1 und n selbst sind.

Beweis. Jede ganze Zahl besitzt 1 und sich selbst als Teiler. Somit muss $\sigma_0(n) = 2$ für $n \geq 2$ äquivalent zu $T_{\geq 1}(n) = \{1, n\}$ sein. □

Satz 8.14 (Satz des Euklid). Es gibt unendlich viele Primzahlen.

Klassischer Beweis. Sei $M = \{p_1, \ldots, p_n\}$ eine endliche Menge von Primzahlen. Es wird gezeigt, dass eine weitere Primzahl $p \notin M$ existiert. Man bilde dazu das Produkt $m = \prod_{k=1}^{n} p_k$. Nun ist $m + 1$ entweder prim oder nicht. Falls $m + 1$ prim ist, ist mit $p = m + 1$ eine weitere Primzahl gefunden. Sei also $m + 1$ nicht prim, womit mindestens ein Primfaktor p enthalten ist. Angenommen, es wäre $p = p_k$ für eines der k. Dann gälte $p_k \mid m + 1$. Es gilt gemäß Konstruktion von m aber auch $p_k \mid m$. Ergo wäre p_k ebenso ein Teiler der Differenz $(m + 1) - m = 1$. Das ist absurd, weil keine Primzahl ein Teiler der Zahl 1 ist. □

Index

Abbildung, 17
Ableitung, 36
abzählbares Auswahlaxiom, 24
adjungierte Matrix, 49
algebraische Zahlen
    Kardinalität, 25
Assoziativgesetz
    Mengen, boolesche Algebra, 12
Aussagenlogik, 5
Auswahlaxiom
    abzählbares, 24

Banach
    Fixpunktsatz von, 40
bedingte Wahrscheinlichkeit, 91
bedingter Erwartungswert, 95
beschränkte Folge, 29
Bildmenge, 17

differenzierbar, 36
Distributivgesetz
    boolesche Algebra, 6
    Urbildoperation, 18
Dreiecksungleichung, 46
    umgekehrte, 47

Epsilon-Umgebung, 29
Erwartungswert, 93
    bedingter, 95

Fixpunkt-Iteration, 40
Fixpunktgleichung, 32
Fixpunktsatz von Banach, 40
folgenstetig, 32

Gesetz der totalen Wahrsch., 92
Gleichheit
    von Abbildungen, 17
    von Mengen, 11
gleichmächtig, 24
Grenzwert, 29
Grenzwertsätze, 31

homogener Markow-Prozess, 97

Indikatorfunktion, 24
Injektion, 17
inverse Matrix, 49

kartesisches Produkt, 11
Kolmogorow-Chapman-Gleichung, 97
Kommutativgesetz
    boolesche Algebra, 5
    Mengen, boolesche Algebra, 11
Komposition, 17
konjugierte Matrix, 49
Kontraktion, 40
konvergente Folge, 29

Markow-Prozess, 97
Mengenlehre, 11
metrischer Raum, 46

Newton-Verfahren, 40
normierter Raum, 46

offene Epsilon-Umgebung, 29
offener Kern, 41

Prädikatenlogik, 7
Produktregel, 36

Relation, 15

Schnittmenge, 11
stetig
    folgenstetig, 32
stochastischer Prozess, 97
Surjektion, 17

Teilmenge, 11
transponierte Matrix, 49

Umgebungsfilter, 41
umgekehrte Dreiecksungleichung, 47
unabhängige Ereignisse, 94
unabhängige Zufallsgrößen, 94
unitäre Matrix, 51
Urbildmenge, 17

Vereinigungsmenge, 11
Verkettung, 17

Zufallsgröße, 91