Beweisarchiv

August 2019

Inhaltsverzeichnis

1 Grundlagen

1.1 Aussagenlogik

Satz 1.1. (bool-dl: Distributivgesetze). Es gilt:

$$A \wedge (B \vee C) \iff A \wedge B \vee A \wedge C, \tag{1.1}$$
$$A \vee (B \wedge C) \iff (A \vee B) \wedge (A \vee C). \tag{1.2}$$

1.2 Prädikatenlogik

Definition 1.1. (bounded: beschränkte Quantifizierung).

$$\forall x{\in}M\,(P(x)) \;:\!\iff \forall x(x \in M \implies P(x)), \tag{1.3}$$
$$\exists x{\in}M\,(P(x)) \;:\!\iff \exists x(x \in M \wedge P(x)). \tag{1.4}$$

Satz 1.2. (general-dl: allgemeine Distributivgesetze). Es gilt:

$$A \wedge \exists x(P(x)) \iff \exists x(A \wedge P(x)), \tag{1.5}$$
$$A \vee \forall x(P(x)) \iff \forall x(A \vee P(x)). \tag{1.6}$$

Satz 1.3. (exists-dl: Distributivgesetz). Es gilt:

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

Satz 1.4. (exists-asym-dl: asymmetrisches Distributivgesetz). Es gilt:

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

Satz 1.5. Es gilt:

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

Satz 1.6. (exists-cl: Kommutativgesetz). Es gilt:

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

Satz 1.7. (all-cl: Kommutativgesetz). Es gilt:

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

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

$$A \wedge \exists x{\in}M\,(P(x)) \iff \exists x{\in}M\,(A \wedge P(x)), \tag{1.7}$$
$$A \vee \forall x{\in}M\,(P(x)) \iff \forall x{\in}M\,(A \vee P(x)). \tag{1.8}$$

Beweis. Nach Def. 1.1 (bounded) und Satz 1.2 (general-dl) gilt:

$$\begin{aligned}&A \land \exists x{\in}M\,(P(x)) \iff A \land \exists x(x \in M \land P(x)) \iff \exists x(A \land x \in M \land P(x))\\ &\iff \exists x(x \in M \land A \land P(x)) \iff \exists x{\in}M\,(A \land P(x)).\end{aligned}$$

Nach Def. 1.1 (bounded) und Satz 1.2 (general-dl) gilt:

$$\begin{aligned}&A \lor \forall x{\in}M\,(P(x)) \iff A \lor \forall x(x \in M \implies P(x)) \iff A \lor \forall x(x \notin M \lor P(x))\\ &\iff \forall x(A \lor x \notin M \lor P(x)) \iff \forall x(x \in M \implies A \lor P(x))\\ &\iff \forall x{\in}M\,(A \lor P(x)). \square\end{aligned}$$

Satz 1.9. Es gilt:

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

Beweis. Nach Def. 1.1 (bounded), Satz 1.2 (general-dl) und Satz 1.6 (exists-cl) gilt:

$$\begin{aligned}&\exists x{\in}A\,\exists y{\in}B\,(P(x,y)) \iff \exists x(x \in A \land \exists y[y \in B \land P(x,y)])\\ &\iff \exists x\exists y[x \in A \land y \in B \land P(x,y)] \iff \exists y\exists x[y \in B \land x \in A \land P(x,y)]\\ &\iff \exists y(y \in B \land \exists x[x \in A \land P(x,y)]) \iff \exists y{\in}B\,\exists x{\in}A\,(P(x,y)). \square\end{aligned}$$

Satz 1.10. Es gilt:

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

Beweis. Nach Def. 1.1 (bounded), Satz 1.2 (general-dl) und Satz 1.7 (all-cl) gilt:

$$\begin{aligned}&\forall x{\in}A\,\forall y{\in}B\,(P(x,y)) \iff \forall x(x \in A \Rightarrow \forall y[y \in B \Rightarrow P(x,y)])\\ &\iff \forall x(x \notin A \lor \forall y[y \notin B \lor P(x,y)]) \iff \forall x\forall y[x \notin A \lor y \notin B \lor P(x,y)]\\ &\iff \forall y\forall x[y \notin B \lor x \notin A \lor P(x,y)] \iff \forall y(y \notin B \lor \forall x[x \notin A \lor P(x,y)])\\ &\iff \forall y(y \in B \Rightarrow \forall x[x \in A \Rightarrow P(x,y)]) \iff \forall y{\in}B\,\forall x{\in}A\,(P(x,y). \square\end{aligned}$$

Satz 1.11. Für eine Aussage P, die nicht von x abhängt, und ein nichtleeres Diskursuniversum gilt:

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

Beweis. Nach 1.2 (general-dl) gilt:

$$\exists x(P) \iff \exists x(1 \land P) \iff \exists x(1) \land P \iff 1 \land P \iff P.$$

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

Satz 1.12. Es gilt

$$\exists x{\in}M\,(P) \iff (M \neq \emptyset) \land P.$$

Beweis. Nach Def. 1.1 (bounded) und Satz 1.2 (general-dl) gilt:

$$\exists x{\in}M\,(P) \iff \exists x(x \in M \land P) \iff \exists x(x \in M) \land P \iff (M \neq \emptyset) \land P. \square$$

1.3 Mengenlehre

1.3.1 Definitionen

Definition 1.2. (seteq: Gleichheit von Mengen).

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

Definition 1.3. (subseteq: Teilmenge).

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

Definition 1.4. (filter: beschreibende Angabe).

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

Definition 1.5. (cap: Schnitt).

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

Definition 1.6. (cup: Vereinigung).

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

Definition 1.7. (intersection: Schnitt).

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

Definition 1.8. (union: Vereinigung).

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

Definition 1.9. (cart: kartesisches Produkt).

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

1.3.2 Rechenregeln

Satz 1.13. (Kommutativgesetze). Es gilt $A \cap B = B \cap A$ und $A \cup B = B \cup A$.

Beweis. Nach Def. 1.2 (seteq) expandieren:

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

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

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

Für die Vereinigung ist das analog. □

Satz 1.14. (Assoziativgesetze). Es gilt $A \cap (B \cap C) = (A \cap B) \cap C$ und $A \cup (B \cup C) = (A \cup B) \cup C$.

Beweis. Nach Def. 1.2 (seteq) expandieren:

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

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

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

Für die Vereinigung ist das analog. □

Satz 1.15. 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 1.16. Es gilt $A \subseteq B \iff A \cap B = A$.

Beweis. Aufgrund von Satz 1.15 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 1.17. Es gilt $a = b \iff \forall x(x = a \iff x = b)$.

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

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

Die andere Implikation bringen wir zunächst in ihre Kontraposition:

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

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

Satz 1.18. Es gilt $a = b \iff \{a\} = \{b\}$.

Beweis. Es gilt:

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

Nach Satz 1.17 ist das aber äquivalent zu $a = b$. □

Satz 1.19. Es gilt:

$$\forall x \forall y(x = y \land P(x) \iff P(y))$$

Satz 1.20. Es gilt:

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

Beweis. Nach Def. 1.9 (cart) gilt:

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

Unter doppelter Anwendung von Satz 1.5 gilt weiter:

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

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

$$\implies \forall a \forall b[a \in A \land B \in B \implies P(a,b)] \iff \forall a{\in}a\;\forall b{\in}B\,(P(a,b)),$$

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

$$\exists t \exists a \exists b[t = (a,b) \land a \in A \land b \in B \land \overline{P(t)}] \implies \exists a \exists b(a \in a \land b \in B \land \overline{P(a,b)}).$$

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

$$\exists a \exists b[a \in A \land b \in B \land \overline{P(a,b)}] \implies \exists a \exists b(a \in A \land b \in B \land \overline{P(a,b)}).$$

ergibt. □

Satz 1.21. Es gilt:

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

Beweis. Nach Def. 1.9 (cart) gilt:

$$\begin{aligned}&\exists t{\in}A{\times}B\,(P(t)) \iff \exists t(t \in A \times B \land P(t))\\ &\iff \exists t(\exists a \exists b[t = (a,b) \land a \in A \land b \in B] \land P(t))\\ &\iff \exists t \exists a \exists b[a \in A \land b \in B \land t = (a,b) \land P(t)]\\ &\iff \exists a{\in}A\;\exists b{\in}B\;\exists t[t = (a,b) \land P(t)].\end{aligned}$$

Nun gilt aber ganz offensichtlich

$$\exists t[t = (a,b) \land P(t)] \iff P(a,b).$$

Nimmt man $P(a,b)$ an, dann lässt sich $\exists t[t = (a,b) \land P(t)]$ durch Wahl von $t := (a,b)$ bestätigen. Nimmt man umgekehrt $\exists t[t = (a,b) \land P(t)]$ an, lässt sich $P(a,b)$ daraus unter Anwendung von Satz 1.19 ableiten. Da $\exists t[t = (a,b) \land P(t)]$ gegen $P(a,b)$ ersetzt werden darf, folgt die Behauptung. □

Satz 1.22. 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 1.21 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 1.23. 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 1.9 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. □

1.4 Abbildungen

1.4.1 Definitionen

Definition 1.10. (app: Applikation). Für eine Abbildung f ist

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

Definition 1.11. (img: Bildmenge).
Für eine Abbildung $f\colon A \to B$ und $M \subseteq A$ wird die Menge

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

als Bildmenge von M unter f bezeichnet.

Definition 1.12. (preimg: Urbildmenge). Für eine Abbildung $f\colon A \to B$ wird

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

als Urbildmenge von M unter f bezeichnet.

Definition 1.13. (inj: Injektion).
Eine Abbildung $f\colon A \to B$ 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

$$\forall x_1 \forall x_2 (x_1 \ne x_2 \implies f(x_1) \ne f(x_2)).$$

Definition 1.14. (sur: Surjektion).
Eine Abbildung $f\colon A \to B$ heißt genau dann surjektiv, wenn gilt:

$$B \subseteq f(A).$$

Definition 1.15. (composition: Verkettung).
Für Abbildungen $f\colon A \to B$ und $g\colon B \to C$ heißt

$$(g \circ f)\colon A \to C, \quad (g \circ f)(x) := g(f(x))$$

Verkettung von f und g.

1.4.2 Grundlagen

Satz 1.24. (feq: Gleichheit von Abbildungen). Zwei Abbildungen $f\colon A \to B$ und $g\colon C \to D$ sind genau dann gleich, kurz $f = g$, wenn $A = C$ und $B = D$ und

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

Beweis. Nach Definition gilt $f = g$ genau dann, wenn $(G_f, A, B) = (G_g, C, D)$, was äquivalent zu $G_f = G_g \land A = C \land B = D$ ist. Nach Def. 1.2 (seteq) gilt

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

Nach Satz 1.17 und Def. 1.10 (app) gilt

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

Da die Quantifizerung auf $x \in A$, $y \in B$ und $t \in A \times B$ beschränkt ist, konnte im letzten Schritt Satz 1.20 angewendet werden. □

Satz 1.25. (preimg-dl: Distributivität der Urbildoperation).
Für $f\colon A \to B$ und beliebige Mengen M_i gilt:

$$f^{-1}(M_1 \cap M_2) = f^{-1}(M_1) \cap f^{-1}(M_2), \tag{1.9}$$

$$f^{-1}(M_1 \cup M_2) = f^{-1}(M_1) \cup f^{-1}(M_2), \tag{1.10}$$

$$f^{-1}(\bigcap_{i\in I} M_i) = \bigcap_{i\in I} f^{-1}(M_i), \tag{1.11}$$

$$f^{-1}(\bigcup_{i\in I} M_i) = \bigcup_{i\in I} f^{-1}(M_i). \tag{1.12}$$

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

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

Für die Vereinigung ist das analog.
Schnitt von beliebig vielen Mengen. Nach Def. 1.2 (seteq) expandieren:

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

Nach Def. 1.12 (preimg) und Def. 1.7 (intersection) zusammen mit Def. 1.4 (filter) gilt:

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

Für die Vereinigung ist das analog. □

Satz 1.26. (img-cup-dl: Distributivität der Bildoperation über die Vereinigung). Für $f\colon A \to B$ und Mengen $M_i \subseteq A$ gilt:

$$f(M_1 \cup M_2) = f(M_1) \cup f(M_2), \tag{1.13}$$

$$f(\bigcup_{i\in I} M_i) = \bigcup_{i\in I} f(M_i). \tag{1.14}$$

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. 1.11 (img), Def. 1.6 (cup), Satz 1.1 (bool-dl) und Satz 1.3 (exists-dl) gilt:

$$\begin{aligned}&y \in f(M_1 \cup M_2) \iff \exists x[x \in M_1 \cup M_2 \land y = f(x)]\\ &\iff \exists x[(x \in M_1 \lor x \in M_2) \land y = f(x)]\\ &\iff \exists x[x \in M_1 \land y = f(x) \lor x \in M_2 \land y = f(x)]\\ &\iff \exists x[x \in M_1 \land y = f(x)] \lor \exists x[x \in M_2 \land y = f(x)]\\ &\iff y \in f(M_1) \lor y \in f(M_2) \iff y \in f(M_1) \cup f(M_2).\end{aligned}$$

Nach Def. 1.2 (seteq) expandieren:

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

Nach Def. 1.11 (img), Def. 1.8 (union), Satz 1.2 (general-dl) und Satz 1.6 (exists-cl) gilt:

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

Satz 1.27. Es gilt:

$$f(M_1 \cap M_2) \subseteq f(M_1) \cap f(M_2), \tag{1.15}$$

$$f(\bigcap_{i\in I} M_i) \subseteq \bigcap_{i\in I} f(M_i). \tag{1.16}$$

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. 1.11 (img), Def. 1.5 (cap) und Satz. 1.4 (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. 1.11 (img) und Def. 1.7 (intersection) gilt:

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

Satz 1.28. Es gilt $M \subseteq N \implies f^{-1}(M) \subseteq f^{-1}(N)$.

Beweis 1. Gemäß Satz 1.16 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 1.25 (preimg-dl) dann

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

Nochmalige Anwendung von Satz 1.16 liefert das gewünschte Resultat

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

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. □

Satz 1.29. Es gilt $M \subseteq N \implies f(M) \subseteq f(N)$.

Beweis. Gemäß Satz 1.16 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 1.27 dann

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

Laut Satz 1.15 ist folglich $f(M) = f(M) \cap f(N)$. Nochmalige Anwendung von Satz 1.16 bringt das gewünschte Resultat $f(M) \subseteq f(N)$. □

Satz 1.30. Es gilt:

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

Beweis. Nach Def. 1.11 (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 1.31. Es gilt $(g \circ f)^{-1}(M) = f^{-1}(g^{-1}(M))$.

Beweis. Nach Def. 1.12 (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. 1.15 (composition) anwenden und rechts nochmals Def. 1.4 (filter):

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

Satz 1.32. Es gilt $(g \circ f)(M) = g(f(M))$.

Beweis. Nach Def. 1.11 (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. 1.11 (img) expandieren und Def. 1.4 (filter) anwenden. Unter Anwendung von Satz 1.2 (general-dl) und Satz 1.6 (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 1.33. Sei $f\colon A \to B$ eine Abbildung und $A \neq \emptyset$. Man nennt eine Funktion $g\colon B \to A$ 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 $g\colon B \to A$ 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

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

Satz 1.34. Für jede Abbildung f gilt $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$.

Beweis. Ergibt sich sofort gemäß Definition:

$$f^{-1}(A) \setminus f^{-1}(B) = \{x \mid x \in f^{-1}(A) \land \neg x \in f^{-1}(B)\}$$
$$= \{x \mid f(x) \in A \land f(x) \notin B\} = \{x \mid f(x) \in A \setminus B\} = f^{-1}(A \setminus B).$$

Satz 1.35. Für jede Abbildung f gilt $f(f^{-1}(N) \subseteq N$.

Beweis. Gemäß Definition bekommt man

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

Leicht ersichtlich ist nun, dass

$$\exists x (f(x) \in N \land y = f(x)) \implies y \in N. \square$$

Satz 1.36. Für jede Abbildung $f\colon A \to B$ gilt $f(f^{-1}(N)) = N$, sofern $N \subseteq f(A)$ ist.

Beweis. Laut Satz 1.35 bleibt zu zeigen

$$y \in N \implies \exists_{x \in A} (f(x) \in N \land y = f(x)).$$

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

$$\exists_{x \in A} (f(x) \in N \land y = f(x)) \iff \exists_{x \in A} (y = f(x)) \iff y \in f(A).$$

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

Satz 1.37. Für jede Abbildung $f\colon A \to B$ 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 1.36 gilt dann $f(M) = N$. □

Satz 1.38. Ist f injektiv, dann gilt $f(A \setminus B) = f(A) \setminus f(B)$.

Beweis. Da f injektiv ist, gibt es nach Satz 1.33 eine Linksinverse f^{-1}. Nach Satz 1.32 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 1.34 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 1.36 das gesuchte Resultat $f(A) \setminus f(B) = f(A \setminus B)$. □

Satz 1.39. Ist f eine bijektive Abbildung und f^{-1} die Umkehrabbildung von f, dann stimmt das Urbild $f^{-1}(N)$ mit der Bildmenge von N unter der Umkehrabbildung – zur Unterscheidung $(f^{-1})(N)$ geschrieben – überein.

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

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

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

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

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 $(f^{-1})(\{1\}) = \{0\}$.

1.4.3 Kardinalzahlen

Satz 1.40. (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 1.16. (equipotent: Gleichmächtigkeit). Zwei Mengen A, B heißen genau dann gleichmächtig, wenn eine Bijektion $f\colon A\to B$ existiert.

Satz 1.41. Sei M eine beliebige Menge. Die Potenzmenge 2^M ist zur Menge $\{0,1\}^M$ gleichmächtig.

Beweis. Für eine Aussage A sei

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

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

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

Die Abbildung

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

ist eine kanonische Bijektion.

Zur Injektivität. Nach Def. 1.13 (inj) muss gelten:

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

Nach Satz 1.24 (feq) und Def. 1.2 (seteq) wird die Aussage expandiert zu:

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

Es gilt aber nun:

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

Zur Surjektivität. Wir müssen nach Def. 1.14 (sur) prüfen, dass $\{0,1\}^M\subseteq\varphi(2^M)$ gilt. Expansion nach Def. 1.3 (subseteq) und Def. 1.11 (img) ergibt:

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

Um dem Existenzquantor zu genügen, wähle

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

Es gilt $f=\chi_A$, denn

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

Da φ eine Bijektion ist, müssen 2^M und $\{0,1\}^M$ nach Def. 1.16 (equipotent) gleichmächtig sein. □

Satz 1.42. Man setze Axiom 1.40 (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 1.40 (acc) kann aus jeder Menge B_n eine Bijektion $f_n\colon \mathbb{N} \to A_n$ ausgewählt werden. Man betrachte nun

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

Die Abbildung φ ist surjektiv, denn nach Satz 1.30 und Satz 1.22 gilt

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

Daher gilt $|\bigcup_{n\in\mathbb{N}} A_n| \le |\mathbb{N}\times\mathbb{N}| = |\mathbb{N}|$. Für eine beliebige der Bijektionen $f_n \in B_n$ lässt sich die Zielmenge erweitern, so dass man eine Injektion $f\colon \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n$ 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 1.43. 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 1.42 gilt

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

Satz 1.44. Es gibt nur abzählbar unendlich viele algebraische Zahlen.

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

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

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 1.43 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

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

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. □

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:

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

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

Definition 2.2. (lim: konvergente Folge, Grenzwert).

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

bzw.

$$\lim_{n\to\infty} a_n = a \;:\!\iff\; \forall\varepsilon{>}0\; \exists n_0\; \forall n{\ge}n_0\; (\|a_n - a\| < \varepsilon).$$

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:

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

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

Um dem Existenzquantor zu genügen, wählt man nun $\varepsilon = \frac{1}{2}d(g_1,g_2)$. Nach Def. 3.3 (metric-space) gilt $d(g_1,g_2) > 0$, daher ist auch $\varepsilon > 0$. Nach Satz 3.2 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 $n \ge n_0$. Dann gibt es für jedes beliebig große n_0 aber auch $n \ge n_0$ mit $a_n \notin U_\varepsilon(g_2)$. □

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

$$\lim_{n\to\infty} a_n = a \iff \forall\varepsilon{>}0\; \exists n_0\; \forall n{\ge}n_0\; (\|a_n - a\| < R\varepsilon),$$

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 3.4 (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 ergibt

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

Nach der Dreiecksungleichung, das ist Axiom (N3) in Def. 3.5 (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:

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

Beweis. Sei $\varepsilon > 0$ fest aber beliebig. Es gibt nun ein n_0, so dass $\|a_n - a\| < \varepsilon$ für $n \geq n_0$. Multipliziert man auf beiden Seiten mit $|r|$ und zieht Def. 3.5 (normed-space) Axiom (N2) heran, dann ergibt sich

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

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

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

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

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

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

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

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

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

Satz 2.10. (Satz zur Fixpunktgleichung). Sei M ein metrischer Raum und sei $f\colon M \to M$. Sei $x_{n+1} := f(x_n)$ eine Fixpunktiteration. Wenn die Folge (x_n) zu einem Startwert x_0 konvergiert mit $x_n \to x$, 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:

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

2.1.2 Wachstum und Landau-Symbole

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

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

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

Beweis. Nach Def. 2.4 ist

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

Man hat nun

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

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

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

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

2.2 Stetige Funktionen

Definition 2.5. (Grenzwert einer Funktion). Sei $f\colon D \to \mathbb{R}$ 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 $f\colon D \to \mathbb{R}$ 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 $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ heißt Lipschitz-stetig, wenn eine Konstante L existiert, so dass

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

für alle $a, b \in D$.

Definition 2.8. (Lipschitz-stetig an einer Stelle).
Eine Funktion $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}$ heißt Lipschitz-stetig an der Stelle $x_0 \in D$, wenn eine Konstante L existiert, so dass

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

für alle $a \in D$.

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

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

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

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

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

Definition 2.9. (lokal Lipschitz-stetig).
Eine Funktion $f\colon D \to \mathbb{R}$ 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 2.13. Ist die Funktion $f\colon D \to \mathbb{R}$ 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 3.4 gilt

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

Daraus ergibt sich

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

und somit erst recht

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

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

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

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

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

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

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

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

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 $|f'(x)| \le L$ für alle x. Es ergibt sich

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

Nun darf auch $a = b$ gewählt werden. □

Satz 2.15. 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 2.14 muss f Lipschitz-stetig sein. □

Korollar 2.16. 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 2.15. Dann ist auch die Einschränkung von f auf $U_\varepsilon(x_0) \subseteq [a, b]$ Lipschitz-stetig. □

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

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

Ist $f: D \to \mathbb{R}$ 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. □

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

$$f + g \text{ dort differenzierbar mit } (f+g)'(x) = f'(x) + g'(x), \tag{2.1}$$
$$f - g \text{ dort differenzierbar mit } (f-g)'(x) = f'(x) - g'(x), \tag{2.2}$$
$$fg \text{ dort differenzierbar mit } (fg)'(x) = f'(x)g(x) + f(x)g'(x). \tag{2.3}$$

Beweis. Es gilt

$$(f+g)'(x) = \lim_{h\to 0} \frac{(f+g)(x+h) - (f+g)(x)}{h} \tag{2.4}$$
$$= \lim_{h\to 0} \frac{(f(x+h) + g(x+h)) - (f(x) + g(x))}{h} \tag{2.5}$$
$$= \lim_{h\to 0} \left(\frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h}\right) \tag{2.6}$$
$$= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} + \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} = f'(x) + g'(x). \tag{2.7}$$

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:

$$g(x)f'(x) + f(x)g'(x) = g(x)\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} + f(x)\lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \tag{2.8}$$
$$= \lim_{h\to 0} \left[g(x+h)\frac{f(x+h) - f(x)}{h}\right] + \lim_{h\to 0} \left[f(x)\frac{g(x+h) - g(x)}{h}\right] \tag{2.9}$$
$$= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h)}{h} + \lim_{h\to 0} \frac{f(x)g(x+h) - f(x)g(x)}{h} \tag{2.10}$$
$$= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \tag{2.11}$$
$$= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} = \lim_{h\to 0} \frac{(fg)(x+h) - (fg)(x)}{h} = (fg)'(x). \tag{2.12}$$

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

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \tag{2.13}$$

Beweis. Nach der Produktregel (2.3) gilt

$$0 = 1' = \left(g \cdot \frac{1}{g}\right)' = g' \cdot \frac{1}{g} + g \cdot \left(\frac{1}{g}\right)'. \tag{2.14}$$

Umstellen bringt $(1/g)'(x) = -g'(x)/g(x)^2$. Nochmalige Anwendung der Produktregel (2.3) bringt

$$\left(\frac{f}{g}\right)'(x) = \left(f \cdot \frac{1}{g}\right)'(x) = f'(x) \cdot \frac{1}{g(x)} + f(x)\left(\frac{1}{g}\right)'(x) \tag{2.15}$$

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

Satz 2.20. Für $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) := x^n$ mit $n \in \mathbb{N}$ gilt $f'(x) = nx^{n-1}$.

Beweis. Heranziehung des binomischen Lehrsatzes bringt

$$f'(x) = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} = \lim_{h\to 0} \frac{\sum_{k=0}^{n} \binom{n}{k} x^{n-k} h^k - x^n}{h} \tag{2.17}$$

$$= \lim_{h\to 0} \left(nx^{n-1} + \sum_{k=2}^{n} \binom{n}{k} x^{n-k} h^{k-1}\right) = nx^{n-1}. \;\square \tag{2.18}$$

Satz 2.21. Für $f\colon \mathbb{R}\setminus\{0\} \to \mathbb{R}$, $f(x) := x^n$ 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 2.20 gezeigt. Sei nun $a \in \mathbb{N}$ und $n = -a$. Nach der Produktregel (2.3) und Satz 2.20 gilt

$$0 = \frac{\mathrm{d}}{\mathrm{d}x} 1 = \frac{\mathrm{d}}{\mathrm{d}x}(x^a x^{-a}) = x^{-a} \frac{\mathrm{d}}{\mathrm{d}x} x^a + x^a \frac{\mathrm{d}}{\mathrm{d}x} x^{-a} = x^{-a} a x^{a-1} + x^a \frac{\mathrm{d}}{\mathrm{d}x} x^{-a}. \tag{2.19}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}x} x^{-a} = -a x^{-a-1} \implies \frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}. \;\square \tag{2.20}$$

2.3.2 Glatte Funktionen

Satz 2.22. Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Funktion mit der Eigenschaft $f(x) = 0$ für $x \le 0$ und $f(x) > 0$ für $x > 0$. Es gibt glatte Funktionen mit dieser Eigenschaft, jedoch keine analytischen.

Beweis. Wegen $f(x) = 0$ für $x \le 0$ 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:

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

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

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

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. □

2.4 Fixpunkt-Iterationen

Definition 2.11. (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 2.23. (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 2.24. (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 2.14 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 $L \le r$ sein, und somit $L < 1$. D. h. φ ist eine Kontraktion. Die Konvergenz der Folge (x_n) ist gemäß Satz 2.23 gewährleistet. □

Satz 2.25. (Hinreichendes Konvergenzkriterium zum Newton-Verfahren).
Sei $f : [a, b] \to \mathbb{R}$ zweimal stetig differenzierbar und $f'(x) \neq 0$ für alle x. Sei

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

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 2.24 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. (nhfilter: Umgebungsfilter).

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

Definition 3.2. (int: offener Kern).

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

Satz 3.1. Der offene Kern 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. 3.2 (int) expandieren:

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

Den äußeren Allquantor brauchen wir nicht weiter mitschreiben, da alle freien Variablen automatisch allquantifiziert werden. Nach Def. 3.1 (nhfilter) 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:

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

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

$$\exists O(x \in M \land O \in T \land x \in O \land O \subseteq M).$$

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. □

3.2 Metrische Räume

3.2.1 Metrischer Räume

Definition 3.3. (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)	$d(x,y) \le d(x,z) + d(z,y).$	(Dreiecksungleichung)

Definition 3.4. (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 3.2. (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 $a\le A$ und $b\le B$ 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

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

Gemäß der Dreiecksungleichung Def. 3.3 Axiom (M3) gilt nun aber

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

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

Korollar 3.3. (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 3.2. Wähle speziell z. B. $a=b=d(p,q)/2$. □

3.2.2 Normierte Räume

Definition 3.5. (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)	$\|\lambda x\| = \lvert\lambda\rvert\,\|x\|,$	(betragsmäßige Homogenität)
(N3)	$\|x+y\| \le \|x\|+\|y\|.$	(Dreiecksungleichung)

Satz 3.4. (umgekehrte Dreiecksungleichung). In jedem normierten Raum gilt

$$\big|\|x\|-\|y\|\big| \le \|x-y\|.$$

Beweis. Auf beiden Seiten von Def. 3.5 (normed-space) Axiom (N3) wird $\|y\|$ subtrahiert. Es ergibt sich

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

Substitution $x := x-y$ bringt nun

$$\|x\|-\|y\| \le \|x-y\|.$$

Vertauscht man nun x und y, dann ergibt sich

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

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

3.2.3 Homöomorphien

Satz 3.5. (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 3.6. 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 1.38. Da $\mathbb{R}_{>0}$ zusammenhängend ist, $\mathbb{R} \setminus \{f(0)\}$ aber nicht, kann f laut Satz 3.5 nicht stetig sein. □

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:

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

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

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

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

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

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. □

4.2.1 Quadratische Matrizen

Satz 4.9. 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). □

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 der selben 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. □

5.2 Polynomringe

5.2.1 Einsetzungshomomorphismus

Satz 5.4. Die Abbildung $\Phi\colon \mathbb{R}[X] \to \mathrm{Abb}(X,X)$ mit $\Phi(f)(x) := f(x)$ ist injektiv.

Beweis. Sei $f = \sum_{k=0}^{n} a_k X^k$ und $g = \sum_{k=0}^{n} b_k X^k$, wobei $n = \max(\deg f, \deg g)$. Zu zeigen ist

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

d. h.

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

Die Umformung der Voraussetzung ergibt $\sum_k (b_k - a_k)x^k = 0$. D. h. jedes der $(b_k - a_k)$ muss verschwinden. Zu zeigen ist also lediglich

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

Wenn $f(x) = 0$ für alle x ist, muss auch die Ableitung $D^m f(x) = 0$ sein. Es gilt $D^k x^k = k!$, und daher

$$D^n \sum_{k=0}^{n} c_k x^k = n! \cdot c_n = 0 \implies c_n = 0.$$

Demnach ergibt sich dann aber auch

$$D^{n-1} \sum_{k=0}^{n} c_k x^k = (n-1)! \cdot c_{n-1} = 0 \implies c_{n-1} = 0$$

usw. Man erhält $c_k = 0$ für alle k. □

6 Wahrscheinlichkeitsrechnung

6.1 Diskrete Wahrscheinlichkeitsräume

Definition 6.1. (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 $\sum_{\omega\in\Omega} P(\{\omega\}) = 1$ besitzt.

Bemerkung: Man schreibt auch $P(\omega) := P(\{\omega\})$.

Definition 6.2. (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 6.3. (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 6.1. Es gilt

$$E(X) = \sum_{x\in X(\Omega)} xP(X^{-1}(x)) = \sum_{x\in X(\Omega)} xP(X=x).$$

Beweis. Zunächst gilt

$$\sum_{\substack{\omega\in\Omega\\ X(\omega)=x}} P(\omega) = P(\bigcup_{\substack{\omega\in\Omega\\ X(\omega)=x}} \{\omega\}) = P(\{\omega\in\Omega \mid X(\omega)=x\}) = P(X^{-1}(x)).$$

Da die Reihe zu $E(X)$ nach Def. 6.3 absolut konvergent ist, darf sie beliebig umgeordnet werden und man bekommt

$$E(X) = \sum_{\omega\in\Omega} X(\omega)P(\omega) = \sum_{x\in X(\Omega)} \sum_{\substack{\omega\in\Omega\\ X(\omega)=x}} xP(\omega) = \sum_{x\in X(\Omega)} x \sum_{\substack{\omega\in\Omega\\ X(\omega)=x}} P(\omega)$$
$$= \sum_{x\in X(\Omega)} xP(X^{-1}(x)). \square$$

Korollar 6.2. Der Erwartungswertoperator ist ein lineares Funktional, d. h. es gilt $E(\alpha X) = \alpha E(X)$ und $E(X+Y) = E(X) + E(Y)$.

Beweis. Aufgrund der Konvergenz der Reihen gilt

$$E(\alpha X) = \sum_{\omega\in\Omega} \alpha X(\omega)P(\omega) = \alpha \sum_{\omega\in\Omega} X(\omega)P(\omega) = \alpha E(X)$$

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 6.3. Ist $X \le Y$, dann ist auch $E(X) \le E(Y)$.

Beweis. Gemäß $P(\omega) \ge 0$ ist

$$X \le Y \iff X(\omega) \le Y(\omega) \iff 0 \le Y(\omega) - X(\omega) \iff 0 \le (Y(\omega)-X(\omega))P(\omega).$$

Somit hat man

$$X \le Y \implies 0 \le E(Y-X) = \sum_{\omega\in\Omega} (Y(\omega)-X(\omega))P(\omega),$$

und gemäß Linearität daher

$$X \le Y \implies 0 \le E(Y-X) = E(Y)-E(X) \iff E(X) \le E(Y). \;\square$$

Definition 6.4. (Unabhängige Ereignisse).
Zwei Ereignisse A, B heißen unabhängig, falls $P(A\cap B) = P(A)P(B)$.

Definition 6.5. (Unabhängige Zufallsgrößen).
Zwei Zufallsgrößen $X, Y\colon \Omega \to \mathbb{R}$ heißen unabhängig, wenn die Ereignisse $\{X \in A\}$ und $\{X \in B\}$ für alle Mengen $A, B \subseteq \mathbb{R}$ unabhängig sind.

Satz 6.4. Zwei Zufallsgrößen $X, Y\colon \Omega \to \mathbb{R}$ 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

$$\begin{aligned} &P(\{X\in A\}\cap\{Y\in B\}) = P(\bigcup_{x\in A}\{X=x\}\cap\bigcup_{y\in B}\{Y=y\}) \\ &= P(\bigcup_{x\in A}\bigcup_{y\in B}(\{X=x\}\cap\{Y=y\})) = \sum_{x\in A}\sum_{y\in B} P(\{X=x\}\cap\{Y=y\}) \\ &= \sum_{x\in A}\sum_{y\in B} P(X=x)P(Y=y) = \sum_{x\in A} P(X=x)\sum_{y\in B} P(Y=y) \\ &= P(\bigcup_{x\in A}\{X=x\})P(\bigcup_{y\in B}\{Y=y\}) = P(X\in A)P(Y\in B). \;\square \end{aligned}$$

Index