Skript zur Vorlesung

Analysis 2

(Sommersemester 2015)

Dieses Geheft enthält in kompakter Form die wesentlichen Inhalte, wie sie in der Vorlesung ,,Analysis 2" vorgestellt werden.

Es ist zum Gebrauch neben der Vorlesung gedacht und erhebt nicht den Anspruch, ,,in sich selbst verständlich" oder vollständig zu sein.

S. Hilger

9 Differenzierbare Funktionen

9.10 Die Regeln von l'Hospital

9.10.1 Satz: Regel von l'Hospital für Grenzwerte im Endlichen

Es seien $f, g :]c, a[\to \mathbb{R}$ zwei auf $]c, a[$ differenzierbare Funktionen. g habe keine Nullstelle. Es seien die Voraussetzungen

$$\lim_{x \nearrow a} f(x) = 0, \qquad \lim_{x \nearrow a} g(x) = 0.$$

erfüllt. Wenn der Grenzwert G rechts existiert, dann gilt

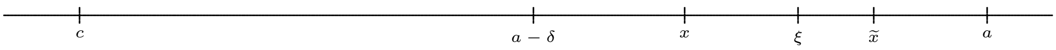
$$\lim_{x \nearrow a} \frac{f(x)}{g(x)} = \overbrace{\lim_{x \nearrow a} \frac{f'(x)}{g'(x)}}^{=G}.$$

9.10.2 Bemerkung Analoges gilt für die folgenden Situationen:

- Funktionen $f, g :]a, d[\to \mathbb{R}$ und rechtsseitige Grenzwerte $\lim\limits_{x \searrow a}$.
- Funktionen $f, g :]c, d[\setminus \{a\} \to \mathbb{R}$, wobei $c < a < d$ und Grenzwerte $\lim\limits_{x \overset{\bullet}{\to} a}$.

9.10.3 Beweis Es sei $\varepsilon > 0$ vorgegeben.

(0) Der folgende Beweis wird durch die Graphik illustriert.



(1) Zu $\frac{\varepsilon}{2} > 0$ gibt es ein $\delta > 0$, so dass

$$\left| \frac{f'(u)}{g'(u)} - G \right| < \frac{\varepsilon}{2} \qquad \text{für alle} \quad u \in]a - \delta, a[.$$

Wir betrachten jetzt ein beliebiges $x \in]a - \delta, a[$ und müssen zeigen, dass $\left| \frac{f(x)}{g(x)} - G \right| < \varepsilon$.

Ist weiter $\widetilde{x} \in]x, a[$, so gibt es nach dem verallgemeinerten Mittelwertsatz 9.8.6 ein $\xi \in]x, \widetilde{x}[\subseteq]a - \delta, a[$, so dass

$$\left| \frac{f(\widetilde{x}) - f(x)}{g(\widetilde{x}) - g(x)} - G \right| = \left| \frac{f'(\xi)}{g'(\xi)} - G \right| < \frac{\varepsilon}{2}.$$

In dieser Abschätzung lassen wir unter Fixierung von x die Stelle $\widetilde{x}$ von links gegen a laufen. Es folgt

$$\left| \frac{f(x)}{g(x)} - G \right| = \lim_{\widetilde{x} \nearrow a} \left| \frac{f(\widetilde{x}) - f(x)}{g(\widetilde{x}) - g(x)} - G \right| \leq \frac{\varepsilon}{2} < \varepsilon.$$

9.10.4 Beispiele

(1) Wir bestätigen den bereits bekannten Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 0} \dfrac{\sin(x)}{x} = 1$ (vgl. Formeln 8.8.4) mit Hilfe des Satzes von l'Hospital.

Dazu berechnen wir die Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = \cos(x), \quad g'(x) = 1, \qquad \lim_{x \overset{\bullet}{\to} 0} \frac{\cos(x)}{1} = 1.$$

(2) Existiert der Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 1} \dfrac{3x^3 - 9x + 6}{x^3 - 5x^2 + 7x - 3}$?

Wir berechnen die Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = 9x^2 - 9, \quad g'(x) = 3x^2 - 10x + 7.$$

Es liegt wieder die Situation $\frac{0}{0}$ vor. Wir können versuchen, auf diese Situation die Regel von l'Hospital erneut anzuwenden.

(3) Existiert der Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 1} \dfrac{9x^2 - 9}{3x^2 - 10x + 7}$?

Wir berechnen erneut die (zweiten) Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = 18x, \quad g'(x) = 6x - 10, \quad \lim_{x \overset{\bullet}{\to} 1} \frac{18x}{6x - 10} = \frac{18}{-4} = -\frac{9}{2}.$$

Also gilt:

$$\lim_{x \overset{\bullet}{\to} 1} \frac{9x^2 - 9}{3x^2 - 10x + 7} = -\frac{9}{2}$$

und deshalb

$$\lim_{x \overset{\bullet}{\to} 1} \frac{3x^3 - 9x + 6}{x^3 - 5x^2 + 7x - 3} = -\frac{9}{2}.$$

Man hätte dieses Ergebnis natürlich auch durch Kürzen des quadratischen Polynoms $(x - 1)^2$ mittels Polynomdivision erhalten können.

9.10.5 Andere Grenzwertsituationen

$\boxed{\frac{0}{0}}$ Der Satz und die bisherigen Beispiele betreffen die Situation

$$\lim_{x \nearrow a} f(x) = 0, \qquad \lim_{x \nearrow a} g(x) = 0,$$

die wir mit dem Symbol $\boxed{\frac{0}{0}}$ (kein mathematisches Objekt) umschreiben können. Andere Grenzwertsituationen können darauf zurückgeführt werden:

$\boxed{\frac{\infty}{\infty}}$ Existiert der Grenzwert

$$\lim_{x \overset{\bullet}{\to} a} \frac{f(x)}{g(x)}, \qquad \text{wobei} \qquad \lim_{x \overset{\bullet}{\to} a} f(x) = \infty, \quad \lim_{x \overset{\bullet}{\to} a} g(x) = \infty \quad ?$$

Diese Situation kann durch eine kleine Umformung auf die Situation $\boxed{\frac{0}{0}}$ zurückgeführt werden:

$$\frac{f(x)}{g(x)} = \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}}.$$

$\boxed{0 \cdot \infty}$ Existiert der Grenzwert

$$\lim_{x \overset{\bullet}{\to} a} f(x) \cdot g(x), \qquad \text{wobei} \qquad \lim_{x \overset{\bullet}{\to} a} f(x) = 0, \quad \lim_{x \overset{\bullet}{\to} a} g(x) = \infty \quad ?$$

Diese Situation kann ebenfalls durch eine kleine Umformung auf die Situation $\boxed{\frac{0}{0}}$ zurückgeführt werden:

$$f(x) \cdot g(x) = \frac{f(x)}{\frac{1}{g(x)}}.$$

$\boxed{\infty - \infty}$ Existiert der Grenzwert

$$\lim_{x \overset{\bullet}{\to} a} \Big[f(x) - g(x) \Big], \qquad \text{wobei} \qquad \lim_{x \overset{\bullet}{\to} a} f(x) = \infty, \quad \lim_{x \overset{\bullet}{\to} a} g(x) = \infty \quad ?$$

Diese Situation kann durch eine kleine Umformung auf die Situation $\boxed{\frac{\infty}{\infty}}$ und dann — gegebenenfalls — auf $\boxed{0 \cdot \infty}$ zurückgeführt werden:

$$f(x) - g(x) = [\frac{f(x)}{g(x)} - 1] \cdot g(x).$$

9.10.6 Satz: Regel von L'Hospital für Grenzwerte im Unendlichen

Es seien $f, g :]c, \infty[\to \mathbb{R}$ mit $c \in \mathbb{R} \cup \{-\infty\}$ zwei differenzierbare Funktionen. g habe keine Nullstelle. Es sei die Voraussetzung

$$\lim_{x\to\infty} f(x) = 0 \qquad \lim_{x\to\infty} g(x) = 0.$$

erfüllt. Wenn der Grenzwert G rechts existiert, dann gilt

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

9.10.7 Beweis Wir führen die Aussage dieses zweiten Satzes durch eine Substitution (bijektive Transformation der Definitionsmenge) auf die des ersten Satzes zurück: Mit der bijektiven Abbildung

$$\begin{cases}]\gamma, +\frac{\pi}{2}[& \to &]c, \infty[\\ u & \mapsto & x = \tan u \end{cases}$$

definieren wir die transformierten Funktionen

$$F : \begin{cases}]\gamma, +\frac{\pi}{2}[& \to & \mathbb{R} \\ u & \mapsto & f(\tan u), \end{cases} \qquad G : \begin{cases}]\gamma, +\frac{\pi}{2}[& \to & \mathbb{R} \\ u & \mapsto & g(\tan u). \end{cases}$$

Dann können wir die Voraussetzungen des aktuellen Satzes in die Voraussetzungen des vorherigen Satzes 9.10.1 übersetzen:

$$\begin{aligned} \lim_{u\nearrow\frac{\pi}{2}} F(u) &= \lim_{u\nearrow\frac{\pi}{2}} f(\tan u) = \lim_{x\to\infty} f(x) = 0, \\ \lim_{u\nearrow\frac{\pi}{2}} G(u) &= \lim_{u\nearrow\frac{\pi}{2}} g(\tan u) = \lim_{x\to\infty} g(x) = 0, \\ \lim_{u\nearrow\frac{\pi}{2}} \frac{F'(u)}{G'(u)} &= \lim_{u\nearrow\frac{\pi}{2}} \frac{[f(\tan u)]'}{[g(\tan u)]'} = \lim_{u\nearrow\frac{\pi}{2}} \frac{f'(\tan u)\cdot\frac{1}{\cos^2 u}}{g'(\tan u)\cdot\frac{1}{\cos^2 u}} \\ &= \lim_{u\nearrow\frac{\pi}{2}} \frac{f'(\tan u)}{g'(\tan u)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = G. \end{aligned}$$

Damit haben wir die Voraussetzungen von Satz 9.10.1 nachgewiesen.

Die Rücktransformation liefert jetzt die Behauptung:

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{F(\arctan x)}{G(\arctan x)} = \lim_{u\nearrow\frac{\pi}{2}} \frac{F(u)}{G(u)} \overset{\text{Satz 9.10.1}}{=} G.$$

9.11 Mehrfach stetige Differenzierbarkeit

9.11.1 Definition: Es sei $J \subseteq \mathbb{R}$ ein Intervall, $\mathbb{K}$ einer der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$, $n \in \mathbb{N}_0$.

(1) Eine Funktion $f : J \to \mathbb{K}$ heißt *n–mal differenzierbar*, wenn — der Reihe nach — für $k = 1, \ldots, n$ die Ableitungsfunktionen

$$f^{(0)} := f \qquad f^{(k)} := (f^{(k-1)})'$$

existieren. (An Intervallenden benutzen wir die einseitige Ableitung).

(2) Die Funktion f heißt *n–mal stetig differenzierbar*, wenn zusätzlich die n–te Ableitung stetig ist. Die Menge dieser Funktionen wird mit $\mathcal{C}^n(J)$ oder $\mathcal{C}^n(J, \mathbb{K})$ bezeichnet.

(3) Die Funktion f heißt *unendlich oft (stetig) differenzierbar*, wenn sie n–mal (stetig) differenzierbar ist für jedes $n \in \mathbb{N}_0$. Man schreibt für die zugehörige Menge $\mathcal{C}^\infty(J)$.

Es gelten die einfachen Regeln:

$$\mathcal{C}^0(J) \supseteq \mathcal{C}^1(J) \supseteq \mathcal{C}^2(J) \supseteq \ldots\ldots \supseteq \mathcal{C}^\infty(J)$$

und

$$f \in \mathcal{C}^n(J) \implies f^{(k)} \in \mathcal{C}^{n-k}(J) \quad \text{für} \quad k = 0, \ldots, n.$$

9.11.2 Satz: Es sei $n \in \{0, 1, \ldots, \infty\}$.

(i) Sind die Funktionen $f, g \in \mathcal{C}^n(J)$, so gehören auch die Funktionen

$$\alpha f + \beta g, \qquad f \cdot g, \qquad \frac{f}{g} \text{ (falls definiert)},$$

dazu.

(ii) Gilt $f \in \mathcal{C}^n(J_1)$ und $g \in \mathcal{C}^n(J_2)$, wobei $f(J_1) \subseteq J_2$, so ist

$$g \circ f \in \mathcal{C}^n(J_1).$$

(iii) Die bijektive Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$ gehört zu $\mathbb{C}^\infty(\mathbb{R})$, ihre Umkehrfunktion ist nicht in $\mathcal{C}^1(\mathbb{R})$ enthalten.

Wir wissen nur aus Satz 7.5.1, dass sich die $\mathcal{C}^0(\mathbb{R})$–Eigenschaft einer bijektiven Funktion auf ihre Umkehrfunktionen vererbt.

9.11.3 Beispiele

Vorbemerkung: Die Alternativdefinition der Differenzierbarkeit aus Satz 9.1.1 (C) zeigt, dass für eine reelle Funktion f mit $f(0) = 0$ in 0 die Ableitung 0 existiert, falls

$$|f(x)| \leq x^2.$$

Man setzt nämlich $\delta = \varepsilon$ und beachtet dann

$$|f(x) - 0 - 0 \cdot (x - 0)| = |f(x)| \leq x^2 = x \cdot x.$$

(1) Die Funktion

$$x \mapsto \operatorname{sgn}(x) \cdot \frac{x^2}{2}$$

ist einmal differenzierbar. Die Ableitung $x \mapsto |x|$ ist stetig, aber nicht differenzierbar.

(2) Die Funktion

$$x \mapsto \begin{cases} x^2 \cdot \sin(\frac{1}{x}), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

ist differenzierbar. Die Ableitung

$$x \mapsto \begin{cases} 2x \cdot \sin(\frac{1}{x}) - \cos(\frac{1}{x}), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

ist aber nicht stetig in 0.

(3) Polynome, die Exponentialfunktion, die trigonometrischen Funktionen sin und cos sind offenbar $\mathcal{C}^\infty$ auf $\mathbb{R}$.

(4) Die Logarithmusfunktion(en) gehören zu $\mathcal{C}^\infty(\mathbb{R}^+)$

9.11.4 Satz Für festes $n \in \mathbb{N}_0$ gilt:

$$\begin{aligned} \lim_{x\to\infty} x^{-n}e^x &= \infty & \qquad \lim_{x\to\infty} x^n e^{-x} &= 0 \\ \lim_{x\searrow 0} x^n e^{\frac{1}{x}} &= \infty & \qquad \lim_{x\searrow 0} x^{-n} e^{-\frac{1}{x}} &= 0 \end{aligned}$$

„Die Exponentialfunktion ist stärker als jedes Polynom".

9.11.5 Beweis Für $x > 0$ gilt:

$$\frac{e^x}{x^n} = \frac{\sum_{k=0}^\infty \frac{x^k}{k!}}{x^n} > \frac{\frac{x^{n+1}}{(n+1)!}}{x^n} = \frac{x}{(n+1)!} \ \to \ \infty,$$

woraus die erste Aussage folgt. Die anderen Aussagen erhält man daraus durch geeignete Kehrwertbildung

$$\begin{aligned} &\lim_{x\to\infty} x^n e^{-x} = \lim_{x\to\infty} (x^{-n}e^x)^{-1} = 0 \\ &\lim_{x\searrow 0} x^n e^{\frac{1}{x}} = \lim_{y\to\infty} y^{-n}e^y = \infty \\ &\lim_{x\searrow 0} x^{-n} e^{-\frac{1}{x}} = \lim_{y\to\infty} y^n e^{-y} = 0. \end{aligned}$$

Jetzt zu dem angekündigten Beispiel:

9.11.6 Satz: Bump-Funktion Zu $c, d \in \mathbb{R}$, $c < d$, gibt es eine $\mathcal{C}^\infty$–Funktion $\Lambda_{c,d} : \mathbb{R} \to \mathbb{R}$ mit

$$\Lambda_{c,d}(x) \begin{cases} > 0, & \text{falls } x \in \,]c,d[, \\ = 0, & \text{falls } x \notin \,]c,d[. \end{cases}$$

9.11.7 Bemerkung Graphisch anschaulich könnte man den Eindruck haben, dass die Konstruktion einer solchen Funktion vergleichsweise einfach ist. Gleichwohl muss man aber, um die Konstruktion durchzuführen, schon etwas tiefer in den „Teilekasten" der bisher bekannten Funktionen und in den „Werkzeugkasten" der Kombinationsmethoden greifen. Wir werden das gleich beim Beweis sehen.

9.11.8 Beweis Wir betrachten zunächst die Funktion

$$f : \begin{cases} \mathbb{R} & \to \quad \mathbb{R} \\ x & \mapsto \quad \begin{cases} e^{-\frac{1}{x}}, & \text{falls } x > 0, \\ 0, & \text{falls } x \leq 0 \end{cases} \end{cases}$$

und beweisen per Induktion über $k \in \mathbb{N}_0$ die Aussagen:

($\mathcal{A}_k$) Die Ableitung $f^{(k)}$ existiert und hat die Form

$$f^{(k)}(x) = \begin{cases} p_k(\frac{1}{x})\, e^{-\frac{1}{x}}, & \text{falls } x > 0, \\ 0, & \text{falls } x \leq 0, \end{cases}$$

wobei p_k ein Polynom ist.

Für $k = 0$ ist die Aussage (mit $p_k \equiv 1$) klar.

Als Induktionsvoraussetzung nehmen wir an, dass ($\mathcal{A}_k$) bereits bewiesen ist.

Für $x < 0$ ist klar, dass $f^{(k+1)}(x) = (f^{(k)})'(0) = 0$ existiert.

Für $x > 0$ ist gemäß Produkt- und Kettenregel

$$\begin{aligned} f^{(k+1)}(x) &= -\frac{1}{x^2}\, p_k'(\frac{1}{x})\, e^{-\frac{1}{x}} + \frac{1}{x^2} p_k(\frac{1}{x})\, e^{-\frac{1}{x}} \\ &= \underbrace{\frac{1}{x^2} \left[-p_k'(\frac{1}{x}) + p_k(\frac{1}{x}) \right]}_{=:p_{k+1}(\frac{1}{x})} e^{-\frac{1}{x}}. \end{aligned}$$

Für $x = 0$ schauen wir uns den Differenzenquotient von $f^{(k)}$ bei $x = 0$ an. Er hat die Form

$$\frac{f^{(k)}(h) - f^{(k)}(0)}{h - 0} = \frac{f^{(k)}(h)}{h} = \begin{cases} \frac{1}{h}\, p_k(\frac{1}{h})\, e^{-\frac{1}{h}}, & \text{falls } h > 0, \\ 0, & \text{falls } h < 0. \end{cases}$$

Für $h \searrow 0$ geht dieser Ausdruck aufgrund der vierten Aussage in 9.11.4 gegen 0, für $h \nearrow 0$ ist dies trivialerweise der Fall. Also ist

$$f^{(k+1)}(0) = \lim_{h \overset{\bullet}{\to} 0} \frac{f^{(k)}(h) - f^{(k)}(0)}{h - 0} = 0.$$

Da $(\mathcal{A}_k)$ für alle $k \in \mathbb{N}_0$ richtig ist, ist f eine $\mathcal{C}^\infty$–Funktion.

Dann gehört aber auch die durch

$$\Lambda_{c,d} = f(x-c) \cdot f(d-x)$$

definierte Funktion zu $\mathcal{C}^\infty(\mathbb{R})$. Man überzeuge sich davon, dass sie die anderen Eigenschaften aus dem Satz erfüllt.

10 Integrierbare Funktionen

10.1 Definition

10.1.1 Definition: Zerlegung Es sei ein (echtes) Intervall $[c,d]$ gegeben.

(1) Eine **endliche** (streng monoton steigende) Folge

$$c = x_0 < x_1 < \ldots < x_n = d$$

von $n+1$ Zahlen in $[c,d]$ heißt eine *Unterteilung* oder *Zerlegung* $\mathcal{Z}$ von $[c,d]$ in n Teilintervalle. Die Zahlen x_i heißen *Zerlegungsstellen.* Sie werden in der Menge $\mathcal{Z}$ zusammengefasst.

(2) Eine Zerlegung $\mathcal{Z}_2$ heißt *Verfeinerung* der Zerlegung $\mathcal{Z}_1$, wenn jede Zerlegungsstelle von $\mathcal{Z}_1$ auch eine von $\mathcal{Z}_2$ ist.

(3) Sind zwei Zerlegungen $\mathcal{Z}_1$ und $\mathcal{Z}_2$ eines Intervalls $[c,d]$ gegeben, so können wir die *gemeinsame Verfeinerung* $\mathcal{Z}_1 \sqcup \mathcal{Z}_2$ bilden, deren Menge der Zerlegungsstellen durch Vereinigung der beiden gegebenen Mengen von Zerlegungsstellen gebildet wird.

10.1.2 Definition: Riemann–Integral Es sei $f : [c,d] \to \mathbb{R}$ eine reelle beschränkte Funktion, das heißt es gibt eine Konstante $M > 0$, so dass

$$|f(x)| \leq M \qquad \text{für alle} \qquad x \in [c,d].$$

(1a) Die $\mathcal{Z}$–*Obersumme* von f ist die reelle Zahl

$$\overline{S}(f,\mathcal{Z}) \;:=\; \sum_{i=1}^{n} \sup_{x \in]x_{i-1},x_i[} \{\, f(x) \,\} \cdot (x_i - x_{i-1})$$

(1b) Die $\mathcal{Z}$–*Untersumme* von f ist die reelle Zahl

$$\underline{S}(f,\mathcal{Z}) \;:=\; \sum_{i=1}^{n} \inf_{x \in]x_{i-1},x_i[} \{\, f(x) \,\} \cdot (x_i - x_{i-1})$$

(2a) Das *Oberintegral* von f ist die reelle Zahl

$$\overline{\int_c^d} f(x)\,dx \;:=\; \inf \Big\{ \overline{S}(f,\mathcal{Z}) \Big| \mathcal{Z} \text{ ist Zerlegung von } [c,d] \Big\}.$$

(2b) Das *Unterintegral* von f ist die reelle Zahl

$$\underline{\int_c^d} f(x)\,dx \;:=\; \sup \Big\{ \underline{S}(f,\mathcal{Z}) \Big| \mathcal{Z} \text{ ist Zerlegung von } [c,d] \Big\}.$$

(3) Es ist unmittelbar klar, dass immer

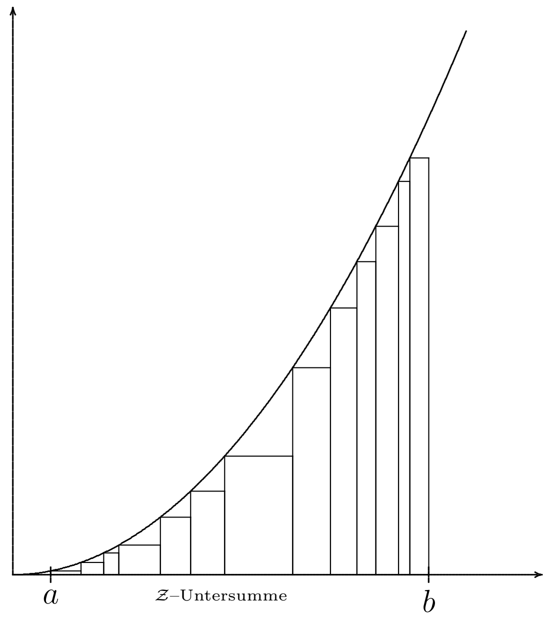
$$\underline{\int_c^d} f(x)\,dx \leq \overline{\int_c^d} f(x)\,dx.$$

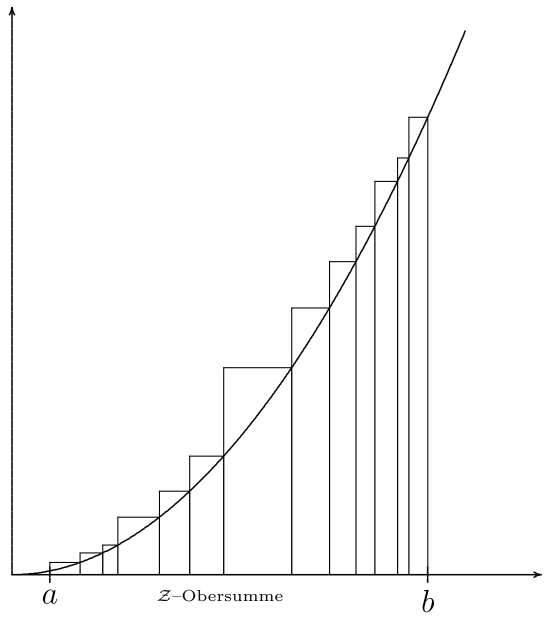
Die Funktion f heißt *(Riemann–Darboux–)integrierbar* (auf $[c,d]$), wenn Oberintegral und Unterintegral übereinstimmen. Diese Zahl

$$\int_c^d f(x)\,dx \;:=\; \overline{\int_c^d} f(x)\,dx = \underline{\int_c^d} f(x)\,dx$$

heißt dann das *(Riemann–Darboux–)Integral* (Bernhard Riemann).

10.1.3 Veranschaulichung am Beispiel einer positiven stetigen streng monoton steigenden Funktion:





10.1.4 Bemerkungen

(1) Die $\mathcal{Z}$–Obersumme kann man sich vorstellen als die (vorzeichenbehaftete) Fläche eines Rechtecksystems, das durch die horizontale Achse, die Abszissen der Intervallgrenzen und $\mathcal{Z}$–Strecken, die von oberhalb auf den Funktionsgraphen ,,aufgesetzt" sind, gebildet wird.

(2) Entsprechend ist die $\mathcal{Z}$–Untersumme als Fläche eines Rechtecksystems aufzufassen, das durch die horizontale Achse, die Abszissen der Intervallgrenzen und $\mathcal{Z}$–Strecken, die von unterhalb an den Funktionsgraphen ,,angepasst" sind, gebildet wird.

(3) Beachte, dass die Werte der Funktion f an den Zerlegungsstellen für die Definition von Ober-, Untersumme, damit Ober-, Unterintegral und Integral ohne Belang sind. Das bedeutet, dass man eine Funktion f an endlich vielen Stellen abändern kann, ohne dass dies einen Einfluß auf die Integrierbarkeit oder den Integralwert hätte. Nützlich ist im folgenden die Abkürzung

$$\mod \mathcal{Z}.$$

Sie bedeutet ,,bis auf Zerlegungsstellen" oder ,,ohne Zerlegungsstellen".

(4) Wenn man das Oberintegral berechnen will, muss man im Prinzip erst die Obersummen bzgl. aller möglichen Zerlegungen berechnen und dann das Infimum (größte untere Grenze) all dieser Zahlen bilden. Dies scheint eine unmögliche Aufgabe zu sein. Wir werden sehen, wie man das — by the help of MATH POWER — bewältigen kann.

(5) In der Mathematik ungleich wichtiger und ,,natürlicher" ist ein anderer Integrationsbegriff, der auf HENRI LEBESGUE zurückgeht. Der wesentliche Unterschied zum Riemann–Begriff besteht darin, dass man dabei die Definitionsmenge der zu integrierenden Funktion in abzählbar unendlich viele Teilmengen zerlegen darf. Leider ist dies viel aufwändiger, wir können das hier nicht bewerkstelligen.

10.1.5 Definition (Erweiterung)

Wir ergänzen die obige Definition. Ist $f : [c, d] \to \mathbb{R}$ integrierbar und $a, b \in [c, d]$, so setzen wir

$$\int_a^b f(x)\,dx \;:=\; \begin{cases} \int_a^b f(x)\,dx, & \text{falls } a < b \quad \text{(wie oben)}, \\ 0, & \text{falls } a = b, \\ -\int_b^a f(x)\,dx, & \text{falls } a > b. \end{cases}$$

10.1.6 Satz: Es seien f, g auf $[c, d]$ integrierbar.

(i) Es gilt dann für beliebige $a, b, e \in [c, d]$

$$\int_a^e f(x)\,dx + \int_e^b f(x)\,dx \;=\; \int_a^b f(x)\,dx.$$

(ii) Falls $f(x) \leq g(x)$ für alle $x \in [c, d]$ ist, gilt

$$\int_c^d f(x)\,dx \;\leq\; \int_c^d g(x)\,dx.$$

10.2 Treppenfunktionen sind integrierbar

10.2.1 Definition: Treppenfunktionen

(1) Eine Funktion $\varphi : [c, d] \to \mathbb{R}$ heißt (naheliegend) $\mathcal{Z}$–*Treppenfunktion*, wenn sie auf jedem der n offenen Teilintervalle $]x_{i-1}, x_i[$, $i = 1, \dots, n$, konstant ist.

(2) Eine Funktion $\varphi : [c, d] \to \mathbb{R}$ heißt *Treppenfunktion*, wenn es eine Zerlegung $\mathcal{Z}$ gibt, so dass φ eine $\mathcal{Z}$–Treppenfunktion ist.

Die Werte einer Treppenfunktion an den Zerlegungsstellen sind für die im folgenden zu entwickelnden Begriffe und Sätze im wesentlichen ohne Bedeutung. Um technische Penibilitäten zu vermeiden, denken wir uns Treppenfunktionen immer mod $\mathcal{Z}$ definiert. Dabei kann man sich selbst heraussuchen, ob man die Treppenfunktionen an den Zerlegungsstellen geeignet ,,nachdefinieren" will oder ob man ständig mitbedenkt, dass die Zerlegungsstellen ohnehin keine Rolle spielen.

10.2.2 Satz: Eine Treppenfunktion $\varphi : [c, d] \to \mathbb{R}$ ist integrierbar, wobei

$$\int_c^d \varphi(x)\, dx = \sum_{k=1}^{n} \varphi(\frac{x_i + x_{i-1}}{2}) \cdot (x_i - x_{i-1})$$

gilt, falls $\mathcal{Z} = \{x_0, \dots, x_n\}$ die zugehörige Zerlegung ist.

10.2.3 Beweis Ist φ eine $\mathcal{Z}$–Treppenfunktion, so gilt für jedes $i = 1, \dots, n$

$$\inf_{x \in]x_{i-1}, x_i[} \{\varphi(x)\} = \varphi(\frac{x_i + x_{i-1}}{2}) = \sup_{x \in]x_{i-1}, x_i[} \{\varphi(x)\},$$

und deshalb

$$\underline{S}(f, \mathcal{Z}) = \sum_{k=1}^{n} \varphi(\frac{x_i + x_{i-1}}{2}) \cdot (x_i - x_{i-1}) = \overline{S}(f, \mathcal{Z}).$$

Das heißt, dass durch die Zerlegung $\mathcal{Z}$ bereits das Infimum der Obersummen und das Supremum der Untersummen realisiert ist. Es folgt

$$\underline{\int_c^d} f(x)\, dx = \sum_{k=1}^{n} \varphi(\frac{x_i + x_{i-1}}{2}) \cdot (x_i - x_{i-1}) = \overline{\int_c^d} f(x)\, dx.$$

Für die Integrationstheorie wichtig sind Treppenfunktionen wegen des folgenden Satzes.

10.2.4 Satz: Integrierbarkeits–Kriterium mittels Treppenfunktionen

Betrachte die drei Aussagen über eine Funktion $f : [c,d] \to \mathbb{R}$.

(A) Die Funktion f ist integrierbar.

(B) Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathcal{Z}$ und zwei $\mathcal{Z}$–Treppenfunktionen $\varphi, \psi : [c,d] \to \mathbb{R}$, so dass

$$\begin{aligned} &\varphi(x) \leq f(x) \leq \psi(x) \ \text{ für alle } x \in [c,d] \mod \mathcal{Z} \\ &\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \leq \varepsilon. \end{aligned}$$

(C) Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathcal{Z}$ und zwei $\mathcal{Z}$–Treppenfunktionen $\varphi, \psi : [c,d] \to \mathbb{R}$, so dass

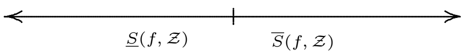
$$\left.\begin{aligned} &\varphi(x) \leq f(x) \leq \psi(x) \\ &\psi(x) - \varphi(x) \leq \varepsilon \end{aligned}\right. \qquad \text{für alle } x \in [c,d] \mod \mathcal{Z}.$$

Dann gelten die folgenden Implikationen:

$$(A) \quad \Longleftrightarrow \quad (B) \quad \Longleftarrow \quad (C).$$

10.2.5 Beweis (A) $\Rightarrow (B)$ Ist f integrierbar, so sind Ober- und Unterintegral gleich:

$$\sup_{\mathcal{Z}} \left\{ \underline{S}(f,\mathcal{Z}) \right\} = \underline{\int_c^d} f(x)\,dx = \overline{\int_c^d} f(x)\,dx = \inf_{\mathcal{Z}} \left\{ \overline{S}(f,\mathcal{Z}) \right\}.$$



Das bedeutet aber, dass es zu jedem $\varepsilon > 0$ Zerlegungen $\mathcal{Z}'$ und $\mathcal{Z}''$ gibt, so dass

$$\overline{S}(f,\mathcal{Z}') - \underline{S}(f,\mathcal{Z}'') \leq \varepsilon.$$

Geht man zu der gemeinsamen Verfeinerung $\mathcal{Z} = \mathcal{Z}' \sqcup \mathcal{Z}''$ über und definiert man dann Treppenfunktionen durch

$$\left.\begin{aligned} \varphi(x) &= \inf_{x \in]x_{i-1},x_i[} \{f(x)\} \\ \psi(x) &= \sup_{x \in]x_{i-1},x_i[} \{f(x)\} \end{aligned}\right. \qquad \text{für } x \in]x_{i-1}, x_i[\ \text{ bzgl. } \mathcal{Z},$$

so gilt $\varphi(x) \leq f(x) \leq \psi(x)$, $x \in [c,d] \mod \mathcal{Z}$, und

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx = \overline{S}(f,\mathcal{Z}) - \underline{S}(f,\mathcal{Z}) \leq \overline{S}(f,\mathcal{Z}') - \underline{S}(f,\mathcal{Z}'') \leq \varepsilon.$$

(B) $\Rightarrow (A)$ Es sei $n \in \mathbb{N}$. Zu $\varepsilon = \frac{1}{n}$ existieren $\mathcal{Z}_n$–Treppenfunktionen φ_n, ψ_n mit

$$\begin{aligned} &\varphi_n(x) \leq f(x) \leq \psi_n(x) \qquad \text{für alle } x \in [c,d] \mod \mathcal{Z}_n, \\ &\int_c^d \psi_n(x)\,dx - \int_c^d \varphi_n(x)\,dx \leq \frac{1}{n}. \end{aligned}$$

Für die zugehörigen $\mathcal{Z}_n$–Ober- und Untersummen gilt:

$$\overline{S}(f,\mathcal{Z}_n) \leq \int_c^d \psi_n(x)\,dx, \qquad \underline{S}(f,\mathcal{Z}_n) \geq \int_c^d \varphi_n(x)$$

deshalb weiter

$$\overline{\int_c^d} f(x)\,dx - \underline{\int_c^d} f(x)\,dx \leq \overline{S}(f,\mathcal{Z}_n) - \underline{S}(f,\mathcal{Z}_n) \leq \int_c^d \psi_n(x)\,dx - \int_c^d \varphi_n(x)\,dx \leq \frac{1}{n}.$$

Da die letzte Ungleichung für alle $n \in \mathbb{N}$ gilt, folgt:

$$\overline{\int_c^d} f(x)\,dx - \underline{\int_c^d} f(x) = 0.$$

Das war zu beweisen.

(C) $\Rightarrow$ (B). Es sei $\varepsilon > 0$ (wie in (B)) vorgegeben. Zu $\frac{\varepsilon}{d-c}$ existieren gemäß (C) $\mathcal{Z}$–Treppenfunktionen φ, ψ, so dass

$$\left.\begin{array}{l} \varphi(x) \leq f(x) \leq \psi(x) \\ \psi(x) - \varphi(x) \leq \frac{\varepsilon}{d-c} \end{array}\right. \qquad \text{für alle } x \in [c,d] \mod \mathcal{Z}.$$

Die Funktion $\psi - \varphi$ ist eine $\mathcal{Z}$–Treppenfunktion, deren Betrag $\leq \frac{\varepsilon}{d-c}$ ist. Wir wollen nicht noch einmal technische Einzelheiten nachvollziehen und schließen etwas schneller als oben:

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx = \int_c^d [\psi - \varphi](x)\,dx \leq \frac{\varepsilon}{d-c} \cdot (d-c) = \varepsilon.$$

10.2.6 Beispiel Die Dirichlet–Funktion

$$\chi_{\mathbb{Q}} : \begin{cases} [0,1] & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 1, & \text{falls } x \in [0,1] \cap \mathbb{Q}, \\ 0, & \text{falls } x \in [0,1] \cap \mathbb{R} \setminus \mathbb{Q}. \end{cases} \end{cases}$$

ist nicht Riemann–integrierbar. (Sie ist Lebesgue–integrierbar.)

10.2.7 Beweis Für jedes Paar von Treppenfunktionen φ, ψ mit

$$\varphi(x) \leq \chi_{\mathbb{Q}}(x) \leq \psi(x) \qquad x \in [0,1] \mod \mathcal{Z}$$

gilt nämlich zwangsläufig:

$$\varphi(x) \leq 0 \qquad \text{und} \qquad \psi(x) \geq 1 \qquad x \in [0,1] \mod \mathcal{Z}$$

und damit

$$\int_0^1 \psi(x)\,dx - \int_0^1 \varphi(x)\,dx \geq 1.$$

Die letzte Differenz kann also nicht kleiner als $\varepsilon = \frac{1}{2}$ (beispielsweise) gemacht werden.

10.3 Regelfunktionen sind integrierbar

10.3.1 Definition: Eine auf einem kompakten Intervall $[c, d]$ definierte Funktion $f : [c, d] \to \mathbb{R}$ heißt *Regelfunktion* oder *sprungstetige Funktion*, wenn für alle $a \in [c, d]$ die einseitigen Grenzwerte existieren, genauer

- $\lim_{x \nearrow a} f(x)$ existiert für alle $a \in]c, d]$,
- $\lim_{x \searrow a} f(x)$ existiert für alle $a \in [c, d[$.

Beachte, dass für $a \in [c, d]$ die drei Werte $\lim_{x \nearrow a} f(x)$, $\lim_{x \searrow a} f(x)$ und $f(a)$ **nicht** übereinstimmen müssen.

10.3.2 Satz: Regelfunktionen sind integrierbar

Eine Regelfunktion $f : [c, d] \to \mathbb{R}$ ist integrierbar.

Der Beweis erfordert ein wenig Vorbereitung.

10.3.3 Präposition

Es sei $f : [c, d] \to \mathbb{R}$ eine Funktion und $a \in [c, d]$.

(i) Existiert in a der linksseitige Grenzwert, so gibt es zu jedem $\varepsilon > 0$ Konstanten $\delta > 0$, $M^+, M^- \in \mathbb{R}$, so dass

$$\begin{aligned} &M^- \leq f(x) \leq M^+ \qquad \text{für alle } x \in [a - \delta, a[\cap]c, d[\\ &M^+ - M^- \leq \varepsilon. \end{aligned}$$

(ii) Existiert in a der rechtsseitige Grenzwert, so gibt es zu jedem $\varepsilon > 0$ Konstanten $\delta > 0$, $M^+, M^- \in \mathbb{R}$, so dass

$$\begin{aligned} &M^- \leq f(x) \leq M^+ \qquad \text{für alle } x \in]a, a + \delta] \cap]c, d[\\ &M^+ - M^- \leq \varepsilon. \end{aligned}$$

10.3.4 Beweis Es genügt die zweite Aussage zu beweisen. Wegen der Existenz des rechtsseitigen Grenzwerts von f an der Stelle a gibt es zu $\frac{\varepsilon}{2}$ ein $\delta > 0$, so dass

$$|f(x) - f(a)| \leq \frac{\varepsilon}{2} \qquad \text{für alle } x \in]a, a + \delta] \cap]c, d[.$$

Setzt man

$$\begin{aligned} M^- &:= f(a) - \frac{\varepsilon}{2} \\ M^+ &:= f(a) + \frac{\varepsilon}{2}, \end{aligned}$$

so sind die Aussagen des Satzes erfüllt: $M^+ - M^- = \varepsilon$ und

$$M^- = f(a) - \frac{\varepsilon}{2} \leq f(x) \leq f(a) + \frac{\varepsilon}{2} = M^+ \qquad \text{für } x \in]a, a + \delta] \cap]c, d[.$$

10.3.5 Beweis (von Satz 10.3.2) Wir wollen die Implikation $(C) \Rightarrow (A)$ aus Satz 10.2.4 anwenden. Dazu sei ein $\varepsilon > 0$ vorgegeben.

(1) Wir betrachten die Zahlen $y \in [c, d]$ mit der folgenden Eigenschaft:

($\mathcal{A}_y$) Es gibt eine Zerlegung $\mathcal{Z}$ von $[c, y]$ und zwei $\mathcal{Z}$–Treppenfunktionen $\varphi, \psi : [c, y] \to \mathbb{R}$ mit

$$\left.\begin{array}{l} \varphi(x) \leq f(x) \leq \psi(x) \\ \psi(x) - \varphi(x) \leq \varepsilon \end{array}\right. \qquad \text{für alle } x \in [c, y] \bmod \mathcal{Z}.$$

(2) Es sei $Y \subseteq [c, d]$ die Menge der Zahlen, die diese Eigenschaft ($\mathcal{A}_y$) erfüllen. Wir setzen

$$a := \sup Y = \sup \Big\{ y \in [c, d] \;\Big|\; (\mathcal{A}_y) \text{ ist erfüllt } \Big\}.$$

Es gilt offenbar $a \geq c$. Wir müssen noch $a = d$ zeigen. Mit dem Ziel der Herleitung eines Widerspruchs nehmen wir an, dass $a < d$.

(3) Wir zeigen zunächst, dass die Eigenschaft ($\mathcal{A}_a$) erfüllt ist. Gemäß Präposition 10.3.3(i) gibt es $\delta > 0$ und Konstanten M^+, M^- mit

$$\begin{array}{ll} M^- \leq f(x) \leq M^+ & \text{für alle } x \in [a - \delta, a[\,\cap\,]c, d[\\ M^+ - M^- \leq \varepsilon. & \end{array}$$

Gemäß Definition von a in Schritt (2) ist die Eigenschaft $\mathcal{A}_{a-\delta}$ erfüllt. Das heißt, auf dem Intervall $[c, a - \delta]$ existieren bereits die richtigen Treppenfunktionen. Ergänzt man diese Treppenfunktionen mit Hilfe der Konstanten M^-, M^+ zu Treppenfunktionen auf dem Intervall $[c, a]$, so ist die Eigenschaft $\mathcal{A}_y$ für alle $c \leq y \leq a$ erfüllt. Also gilt ($\mathcal{A}_a$).

(4) Gemäß Präposition 10.3.3(ii) gibt es $\delta > 0$ und Konstanten M^+, M^- mit

$$\begin{array}{ll} M^- \leq f(x) \leq M^+ & \text{für alle } x \in]a, a + \delta] \;\cap\;]c, d[\\ M^+ - M^- \leq \varepsilon. & \end{array}$$

Das bedeutet aber, dass man die für das Intervall $[c, a]$ bereits bestehenden Treppenfunktionen mit Hilfe dieser Konstanten M^+, M^- zu Treppenfunktionen auf dem Intervall $[c, a + \delta]$ erweitern kann. Dann ist aber $\mathcal{A}_y$ für alle $y \in [c, a + \delta]$ erfüllt im WIDERSPRUCH zur Definition von a.

*** Zeichnung ***

10.3.6 Korollar: Verschiedenen Typen integrierbarer Funktionen

Für eine beschränkte Funktion $f : [c, d] \to \mathbb{R}$ gilt

$$\left.\begin{array}{l} f \text{ Treppenfunktion} \\ \text{ODER} \\ f \text{ stetig} \\ \text{ODER} \\ f \text{ monoton} \end{array}\right\} \Longrightarrow f \text{ Regelfunktion} \Longrightarrow f \text{ integrierbar.}$$

10.3.7 Beweis Stetige Funktionen sind zweifelsohne Regelfunktionen. Ist f monoton (o.B.d.A. wachsend), so existieren

$$\begin{aligned}\lim_{x\nearrow a} f(x) &= \sup_{x<a}\{f(x)\} \qquad \text{für } a \in [c,d[\\ \lim_{x\searrow a} f(x) &= \inf_{x>a}\{f(x)\} \qquad \text{für } a \in]c,d],\end{aligned}$$

also ist f eine Regelfunktion.

10.4 Beispiel

Es sei die Funktion $f : [c,d] \to \mathbb{R}$, $x \mapsto x^2$ gegeben. Wir berechnen für $c \geq 0$ die Obersumme bzgl. der äquidistanten Zerlegung $\mathcal{Z}_n$

$$x_i = c + \frac{i}{n}(d-c).$$

$$\begin{aligned}
\overline{\mathcal{S}}(f,\mathcal{Z}_n) &= \sum_{i=1}^{n} \sup_{x\in]x_{i-1},x_i[} \{f(x)\} \cdot (x_i - x_{i-1}) = \sum_{i=1}^{n} x_i^2 \cdot \frac{d-c}{n} \\
&= \frac{d-c}{n} \cdot \sum_{i=1}^{n} [c + \frac{i}{n}(d-c)]^2 \\
&= \frac{d-c}{n} \cdot \sum_{i=1}^{n} [c^2 + 2\frac{i}{n}c(d-c) + i^2 \left(\frac{d-c}{n}\right)^2] \\
&= \frac{d-c}{n} \cdot \left[\sum_{i=1}^{n} c^2 + \frac{2}{n}c(d-c)\sum_{i=1}^{n} i + \left(\frac{d-c}{n}\right)^2 \sum_{i=1}^{n} i^2\right] \\
&= \frac{d-c}{n} \cdot \left[n \cdot c^2 + \frac{2}{n}c(d-c)\frac{n(n+1)}{2} + \left(\frac{d-c}{n}\right)^2 \frac{n(n+1)(2n+1)}{6}\right] \\
&= (d-c)c^2 + c(d-c)^2\frac{n+1}{n} + (d-c)^3\frac{(n+1)(2n+1)}{6n^2}
\end{aligned}$$

Diese Werte sind für größer werdendes n monoton fallend. Es ist deshalb weiter

$$\begin{aligned}
\inf_{n\in\mathbb{N}} \overline{\mathcal{S}}(f,\mathcal{Z}_n) &= \lim_{n\to\infty} \overline{\mathcal{S}}(f,\mathcal{Z}_n) = (d-c)c^2 + c(d-c)^2 + \frac{(d-c)^3}{3} \\
&= (d-c) \cdot \left[c^2 + cd - c^2 + \frac{d^2 - 2cd + c^2}{3}\right] \\
&= \frac{d-c}{3} \cdot \left[cd + d^2 + c^2\right] = \frac{d^3 - c^3}{3}.
\end{aligned}$$

und dann weiter

$$\overline{\int_c^d} x^2\,dx = \inf_{\mathcal{Z}} \overline{\mathcal{S}}(f,\mathcal{Z}) \leq \inf_{n\in\mathbb{N}} \overline{\mathcal{S}}(f,\mathcal{Z}_n) = \frac{d^3-c^3}{3}.$$

Analog zu oben kann man nachrechnen, dass

$$\underline{\int_c^d} x^2\,dx \geq \frac{d^3-c^3}{3}$$

und es folgt

$$\frac{d^3-c^3}{3} \leq \underline{\int_c^d} x^2\,dx \leq \overline{\int_c^d} x^2\,dx \leq \frac{d^3-c^3}{3}.$$

Also stimmen Unter- und Oberintegral überein (wissen wir wegen Korollar 10.3.6 schon) und es gilt

$$\int_c^d x^2\,dx = \frac{d^3-c^3}{3}.$$

10.5 Linearkombinationen sind integrierbar

10.5.1 Satz Sind φ, ψ Treppenfunktionen und $\alpha, \beta \in \mathbb{R}$, so ist auch die Funktion

$$\alpha\,\varphi + \beta\,\psi$$

eine Treppenfunktion.

10.5.2 Beweis Ist φ eine $\mathcal{Z}'$–Treppenfunktion und ψ eine $\mathcal{Z}''$–Treppenfunktion, so ist $\alpha\,\varphi + \beta\,\psi$ eine $\mathcal{Z}$–Treppenfunktion, wobei $\mathcal{Z} = \mathcal{Z}' \sqcup \mathcal{Z}''$ die gemeinsame Verfeinerung ist. Der genauere Beweis ist einfach, ein bißchen technisch.

10.5.3 Satz: Linearität Sind die beiden Funktionen $f, g : [c,d] \to \mathbb{R}$ integrierbar und $\alpha, \beta \in \mathbb{R}$, so ist auch die Funktion $\alpha f + \beta g$ integrierbar und es gilt:

$$\int_c^d [\alpha f + \beta g](x)\,dx = \alpha \int_c^d f(x)\,dx + \beta \int_c^d g(x)\,dx.$$

Das bedeutet, die Menge der integrierbaren Funktionen ist ein (∞-dimensionaler) $\mathbb{R}$–Vektorraum. Die Menge der Treppenfunktionen ist ein Untervektorraum.

10.5.4 Beweis Wir zeigen zunächst, dass $f + g$ integrierbar ist und wenden dafür das Kriterium aus Satz 10.2.4 an. Dazu sei ein $\varepsilon > 0$ vorgegeben. Zu $\frac{\varepsilon}{2}$ existieren $\mathcal{Z}_1$–Treppenfunktionenen φ_1, ψ_1, so dass

$$\varphi_1(x) \le f(x) \le \psi_1(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}_1$$
$$\int_c^d \psi_1(x)\,dx - \int_c^d \varphi_1(x)\,dx \le \frac{\varepsilon}{2}.$$

und $\mathcal{Z}_2$–Treppenfunktionen φ_2, ψ_2, so dass

$$\varphi_2(x) \le g(x) \le \psi_2(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}_2$$
$$\int_c^d \psi_2(x)\,dx - \int_c^d \varphi_2(x)\,dx \le \frac{\varepsilon}{2}.$$

Dann sind $\varphi_1 + \varphi_2$ und $\psi_1 + \psi_2$ $\mathcal{Z}_1 \sqcup \mathcal{Z}_2$–Treppenfunktionen, für die gilt:

$$\varphi_1(x) + \varphi_2(x) \le f(x) + g(x) \le \psi_1(x) + \psi_2(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}_1 \sqcup \mathcal{Z}_2$$
$$\int_c^d [\psi_1 + \psi_2](x)\,dx - \int_c^d [\varphi_1 + \varphi_2](x)\,dx \le \varepsilon.$$

Als nächstes überlegen wir, dass mit einer Funktion f auch die Funktion $-f$ integrierbar ist. Dies folgt aber mit Hilfe des Kriteriums, wenn man alle beteiligten Funktionen mit -1 multipliziert und die Rollen von $-\varphi$ und $-\psi$ vertauscht.

Ist f integrierbar und $\alpha > 0$, $\varepsilon > 0$ vorgegeben, so existieren $\mathcal{Z}$–Treppenfunktionen mit der Eigenschaft, dass

$$\varphi(x) \le f(x) \le \psi(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}$$
$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \frac{\varepsilon}{\alpha}.$$

Multipliziert man diese beiden Gleichungen mit $\alpha > 0$, so entsteht das Kriterium für die Integrierbarkeit von αf

Da die Funktion $\alpha f + \beta g$ schrittweise aus f und g durch die obigen Operationen aufgebaut werden kann, ist der Satz gezeigt.

10.6 Produkte sind integrierbar

10.6.1 Definition: Positiv- und Negativteil einer Funktion

Für eine Funktion $f : D \to \mathbb{R}$ definieren wir den *Positiv*- und *Negativteil*

$$f_+(x) := \begin{cases} f(x), & \text{falls } f(x) > 0, \\ 0, & \text{falls } f(x) \le 0. \end{cases} \qquad f_-(x) := \begin{cases} -f(x), & \text{falls } f(x) < 0, \\ 0, & \text{falls } f(x) \ge 0. \end{cases}$$

Es gilt dann

$$f = f_+ - f_- \qquad |f| = f_+ + f_-.$$

10.6.2 Satz: Betrags–Integrierbarkeit Es sei $f : [c, d] \to \mathbb{R}$ integrierbar. Dann sind auch $f_+, f_-, |f|$ integrierbar. Es gilt:

$$\Big| \int_c^d f(x)\,dx \Big| \le \int_c^d |f(x)|\,dx.$$

(Beachte die Ähnlichkeit zur Dreiecksungleichung)

10.6.3 Beweis Nach Voraussetzung gibt es zu $\varepsilon > 0$ Treppenfunktionen φ, ψ mit

$$\varphi(x) \le f(x) \le \psi(x) \ \text{ für alle } x \in [c, d] \mod \mathcal{Z}_1$$
$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \varepsilon.$$

Für die zugehörigen Positivteile gilt:

$$\varphi_+(x) \le f_+(x) \le \psi_+(x) \ \text{ für alle } x \in [c, d] \mod \mathcal{Z}_1$$
$$\int_c^d \psi_+(x)\,dx - \int_c^d \varphi_+(x)\,dx \le \int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \varepsilon.$$

Also ist auch f_+ integrierbar. Wegen $f_- = (-f)_+$ ist auch f_- integrierbar. Wegen $|f| = f_+ + f_-$ ist auch $|f|$ integrierbar. Dabei kann man schließen

$$\begin{aligned} f \le |f| \quad &\Longrightarrow \quad \int_c^d f(x)\,dx \le \int_c^d |f(x)|\,dx \\ -f \le |f| \quad &\Longrightarrow \quad -\int_c^d f(x)\,dx = \int_c^d [-f(x)]\,dx \le \int_c^d |f(x)|\,dx \end{aligned}$$

In jedem Fall gilt also

$$\Big| \int_c^d f(x)\,dx \Big| \le \int_c^d |f(x)|\,dx.$$

10.6.4 Satz: Integrierbarkeit von Produkten Es seien $f, g : [c, d] \to \mathbb{R}$ integrierbar.

(i) Für jedes $p \in [1, \infty[$ ist die Funktion $|f|^p$ integrierbar.

(ii) Dann ist auch die Produktfunktion $f \cdot g$ integrierbar.

10.6.5 Beweis

(i) Es sei $M := \sup\{|f(x)| \,\big|\, x \in [c, d]\}$. O.B.d.A. ist $M > 0$.

Ist ein $\varepsilon > 0$ vorgegeben, so existieren (vgl. Satz 10.6.2) $\mathcal{Z}$–Treppenfunktionen φ, ψ, so dass

$$0 \le \varphi(x) \le |f(x)| \le \psi(x) \le M \text{ für alle } x \in [c, d] \bmod \mathcal{Z} \qquad (*)$$

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \frac{\varepsilon}{p \cdot M^{p-1}}.$$

Sind φ_i, ψ_i die Werte der beiden Treppenfunktionen im Intervall $]x_{i-1}, x_i[$ der Zerlegung $\mathcal{Z}$, so gilt

$$|\psi_i^p - \varphi_i^p| \le p \cdot \xi^{p-1} |\psi_i - \varphi_i|,$$

wobei $\xi \in [\varphi_i, \psi_i] \subseteq [0, M]$ ein Mittelwert ist.

Hier haben wir den Mittelwertsatz der Differentialrechnung 9.8.8 angewandt auf die Funktion $x \mapsto x^p$ und die beiden Stellen φ_i, ψ_i. Beachte, dass dieser Schritt für $p < 1$ nicht durchgeführt werden kann.

Wir formulieren um und schätzen weiter ab:

$$|\psi^p(x) - \varphi^p(x)| \le p \cdot M^{p-1} |\psi(x) - \varphi(x)| \qquad \text{für alle } x \in [c, d] \bmod \mathcal{Z}$$

Daraus folgt aber

$$\begin{aligned} \int_c^d |\psi^p(x) - \varphi^p(x)|\,dx &\le p \cdot M^{p-1} \int_c^d |\psi(x) - \varphi(x)|\,dx \\ &\le p \cdot M^{p-1} \cdot \frac{\varepsilon}{p \cdot M^{p-1}} = \varepsilon. \end{aligned}$$

Die Ungleichung

$$\varphi^p(x) \le |f|^p(x) \le \psi^p(x) \text{ für alle } x \in [c, d] \bmod \mathcal{Z}$$

folgt direkt aus $(*)$.

(ii) Dies ist ganz einfach, da

$$fg = \frac{(f+g)^2 - (f-g)^2}{4}.$$

10.7 Integration und Abschätzung, der Mittelwertsatz

10.7.1 Satz (Abschätzungssatz, Mittelwertsatz)

Es sei die Funktionen $\mu : [c,d] \to \mathbb{R}$ eine nicht–negative integrierbare Funktion („eine Gewichtung auf dem Intervall $[c,d]$"). Dann gilt für eine integrierbare Funktion $f : [c,d] \to \mathbb{R}$:

(i) $$\inf_{x\in[c,d]}\{f(x)\} \cdot \int_c^d \mu(x)\,dx \le \int_c^d f(x)\mu(x)\,dx \le \sup_{x\in[c,d]}\{f(x)\} \cdot \int_c^d \mu(x)\,dx.$$

(ii) $$\inf_{x\in[c,d]}\{f(x)\} \cdot (d-c) \le \int_c^d f(x)\,dx \le \sup_{x\in[c,d]}\{f(x)\} \cdot (d-c).$$

(iii) Ist f stetig, so gibt es ein $\xi \in [c,d]$, so dass

$$\int_c^d f(x)\mu(x)\,dx = f(\xi) \cdot \int_c^d \mu(x)\,dx.$$

(iv) Ist f stetig, so gibt es ein $\xi \in [c,d]$, so dass

$$\int_c^d f(x)\,dx = f(\xi) \cdot (d-c).$$

10.7.2 Beweis Es gilt einfach:

$$\inf_{x\in[c,d]}\{f(x)\} \cdot \mu(x) \le f(x)\mu(x) \le \sup_{x\in[c,d]}\{f(x)\} \cdot \mu(x)$$

und deshalb mit Satz 10.1.6(ii) die Behauptung (i). Die Behauptung (ii) ist ein Spezialfall von (i) für $\mu \equiv 1$.

Die Behauptungen (iii) und (iv) ergeben sich wie folgt aus dem Zwischenwertsatz.

Für den Fall $\int_c^d \mu(x)\,dx = 0$ folgt die Aussage (iii) sofort aus (i).

Im Fall $\int_c^d \mu(x)\,dx > 0$ gilt mit (i)

$$\min_{x\in[c,d]}\{f(x)\} = \inf_{x\in[c,d]}\{f(x)\} \le \frac{\int_c^d f(x)\mu(x)\,dx}{\int_c^d \mu(x)\,dx} \le \sup_{x\in[c,d]}\{f(x)\} = \max_{x\in[c,d]}\{f(x)\}.$$

Also existiert eine Zwischenstelle $\xi \in [c,d]$ mit

$$f(\xi) = \frac{\int_c^d f(x)\mu(x)\,dx}{\int_c^d \mu(x)\,dx}.$$

10.8 Riemann'sche Summen

10.8.1 Definition: Es sei $[c, d]$ ein Intervall von $\mathbb{R}$.

(1) Eine Folge aus $2n + 1$ Zahlen

$$\begin{array}{ccccccccccccccccc} c & = & x_0 & \leq & \xi_1 & \leq & x_1 & \leq & \xi_2 & \leq & x_2 & \ldots\ldots & x_{n-1} & \leq & \xi_n & \leq & x_n & = & d \\ & & x_0 & & < & & x_1 & & < & & x_2 & \ldots\ldots & x_{n-1} & & < & & x_n & & \end{array}$$

nennen wir kurz eine *Besetzung* $\mathcal{B}$ des Intervalls $[c, d]$.

(2) Die Zahlen ξ_i, $i = 1, \ldots, n$ heißen in diesem Zusammenhang *Stützstellen*.

(3) Die Zahlen x_i, $i = 0, \ldots, n$ bilden eine Zerlegung $\mathcal{Z}_{\mathcal{B}}$ des Intervalls, die Zahl

$$\eta(\mathcal{Z}_{\mathcal{B}}) := \max\{x_i - x_{i-1} | i = 1, \ldots, n\}$$

heißt *Feinheit* der Zerlegung oder der Besetzung.

10.8.2 Definition:

Es sei $f : [c, d] \to \mathbb{R}$ eine Funktion, $\mathcal{B}$ eine Besetzung des Intervalls $[c, d]$.

Die Zahl

$$\mathcal{R}(f, \mathcal{B}) := \sum_{i=1}^{n} f(\xi_i) \cdot (x_i - x_{i-1})$$

heißt *Riemann'sche Summe (von f bezüglich der Besetzung $\mathcal{B}$).*

10.8.3 Beispiel Es sei $f : [c, d] \to \mathbb{R}$ stetig und $\mathcal{Z}$ eine Zerlegung. Wählt man dann als Stützstellen ξ_i die Maximumsstellen von f in $[x_{i-1}, x_i]$ — sie existieren aufgrund von Satz 7.4.6 — , so gilt $\overline{\mathcal{S}}(f, \mathcal{Z}) = \mathcal{R}(f, \mathcal{B})$.

Die Obersumme kann also als Riemann'sche Summe dargestellt werden.

10.8.4 Satz: Riemann'sche Summen Es sei $f : [c, d] \to \mathbb{R}$ integrierbar.

Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\left| \int_c^d f(x)\, dx - \mathcal{R}(f, \mathcal{B}) \right| \leq \varepsilon,$$

falls $\mathcal{B}$ eine Besetzung des Intervalls $[c, d]$ mit Feinheit $\eta(\mathcal{Z}_{\mathcal{B}}) \leq \delta$ ist.

In Worten: Die Riemann'schen Summen approximieren das Integral, wobei die ,,Qualität der Approximation" allein durch die Feinheit (und nicht durch die konkrete Auswahl) der Besetzung gesteuert wird.

10.8.5 Beweis (1) Wir beweisen die Aussage zunächst für den Fall, dass f eine $\widetilde{\mathcal{Z}}$–Treppenfunktion bzgl. einer Zerlegung

$$\widetilde{\mathcal{Z}} = \{y_0, y_1, \ldots, y_m\}$$

ist. (2) Zu einem vorgegebenen $\varepsilon > 0$ wähle

$$\delta := \frac{\varepsilon}{4mM}, \qquad \text{wobei} \quad M := \sup_{[c,d]}\{|f(x)|\}.$$

Es sei jetzt $\mathcal{B}$ eine (beliebige) Besetzung von $[c,d]$ mit Feinheit $\eta(\mathcal{Z}_\mathcal{B}) \leq \delta$.

(3) Wir definieren eine (andere) $\mathcal{Z}_\mathcal{B}$–Treppenfunktion $r_\mathcal{B} : [c,d] \to \mathbb{R}$, deren Integral mit der Riemann Summe von f bzgl. $\mathcal{B}$ übereinstimmt,

$$\int_c^d r_\mathcal{B}(x)\,dx = \mathcal{R}(f,\mathcal{B}),$$

durch

$$r_\mathcal{B}(x) := \begin{cases} f(\xi_i), & \text{falls } x \in\ [x_{i-1}, x_i[\ \text{ für ein } i \in \{1,\ldots,n\}, \\ f(d), & \text{falls } x = d. \end{cases}$$

(4) Je nach *Art* der Stelle $x \in [c,d]$ kann man die Differenz zwischen der $\widetilde{\mathcal{Z}}$–Treppenfunktion f und der $\mathcal{Z}_\mathcal{B}$–Treppenfunktion $r_\mathcal{B}$ abschätzen:

$$|f(x) - r_\mathcal{B}(x)| \leq \begin{cases} 0, & \text{falls } x \in [x_{i-1}, x_i[\\ & \text{und } y_j \notin [x_{i-1}, x_i] \\ & \qquad \text{für alle } j \in \{1,\ldots,m\}, \\ |f(x)| + |r_\mathcal{B}(x)| \leq 2M, & \text{sonst.} \end{cases}$$

(5) Der untere Fall kann für höchstens $2m$ Indices $i \in \{1,\ldots,n\}$ eintreten. Deshalb gilt

$$\begin{aligned} \left| \int_c^d f(x)\,dx - \mathcal{R}(f,\mathcal{B}) \right| &\leq \left| \int_c^d [f(x) - r_\mathcal{B}(x)]\,dx \right| \leq \int_c^d |f(x) - r_\mathcal{B}(x)|\,dx \\ &= \sum_{i=1}^n \sup_{x \in]x_{i-1}, x_i[} \{|f(x) - r_\mathcal{B}(x)|\} \cdot (x_i - x_{i-1}) \\ &\leq 2m \cdot 2M \cdot \eta(\mathcal{Z}_\mathcal{B}) \leq 2m \cdot 2M \cdot \delta = \varepsilon. \end{aligned}$$

(6) Jetzt sei f eine beliebige integrierbare Funktion, $\varepsilon > 0$ vorgegeben. Es gibt Treppenfunktionen φ, ψ mit der Eigenschaft

$$\varphi(x) \leq f(x) \leq \psi(x), \qquad x \in [c,d] \qquad (*_1)$$

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \leq \frac{\varepsilon}{2} \qquad (*_2)$$

(7) Für diese Treppenfunktionen existieren $\delta_\varphi > 0$ und $\delta_\psi > 0$, so dass

$$\left| \int_c^d \varphi(x)\,dx - \mathcal{R}(\varphi,\mathcal{B}) \right| \leq \frac{\varepsilon}{2}$$

$$\left| \int_c^d \psi(x)\,dx - \mathcal{R}(\psi,\mathcal{B}) \right| \leq \frac{\varepsilon}{2},$$

falls $\mathcal{B}$ eine Besetzung der Feinheit

$$\eta(\mathcal{Z}_{\mathcal{B}}) \leq \delta \ := \ \min\{\delta_\varphi, \delta_\psi\}$$

ist.

(8) Für eine solche Besetzung zieht $(*_1)$ nach sich, dass

$$\mathcal{R}(\varphi, \mathcal{B}) \leq \mathcal{R}(f, \mathcal{B}) \leq \mathcal{R}(\psi, \mathcal{B}).$$

Das aber bedeutet dann:

$$\begin{aligned}
\int_c^d f(x)\,dx - \mathcal{R}(f,\mathcal{B}) &\overset{(*_1)}{\leq} \int_c^d \psi(x)\,dx - \mathcal{R}(\varphi,\mathcal{B}) \\
&\leq \int_c^d \psi(x)\,dx - [\int_c^d \varphi(x)\,dx - \frac{\varepsilon}{2}] \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

$$\begin{aligned}
\mathcal{R}(f,\mathcal{B}) - \int_c^d f(x)\,dx &\overset{(*_1)}{\leq} \mathcal{R}(\psi,\mathcal{B}) - \int_c^d \varphi(x)\,dx \\
&\leq \int_c^d \psi(x)\,dx + \frac{\varepsilon}{2} - \int_c^d \varphi(x)\,dx \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Damit ist der Beweis fertig.

11 Der Hauptsatz der Differential- und Integralrechnung

11.1 Der Satz

11.1.1 Hauptsatz der Differential- und Integralrechnung (HDI)

Es sei $f : J \to \mathbb{R}$ eine auf dem Intervall J definierte stetige Funktion und $a \in J$ eine fest gewählte Stelle.

Für eine weitere Funktion $F : J \to \mathbb{R}$ sind die beiden folgenden Aussagen äquivalent:

(A) F ist die **Integralfunktion (Unbestimmtes Integral)** zu f mit unterer Grenze a, das heißt

$$F(x) = \int_a^x f(t)\,dt.$$

(B) F ist die **Stammfunktion** von f mit Nullstelle bei a, das heißt: F ist auf J differenzierbar und es gilt

$$F'(x) = f(x) \quad \text{für alle } x \in J \qquad \text{und} \qquad F(a) = 0.$$

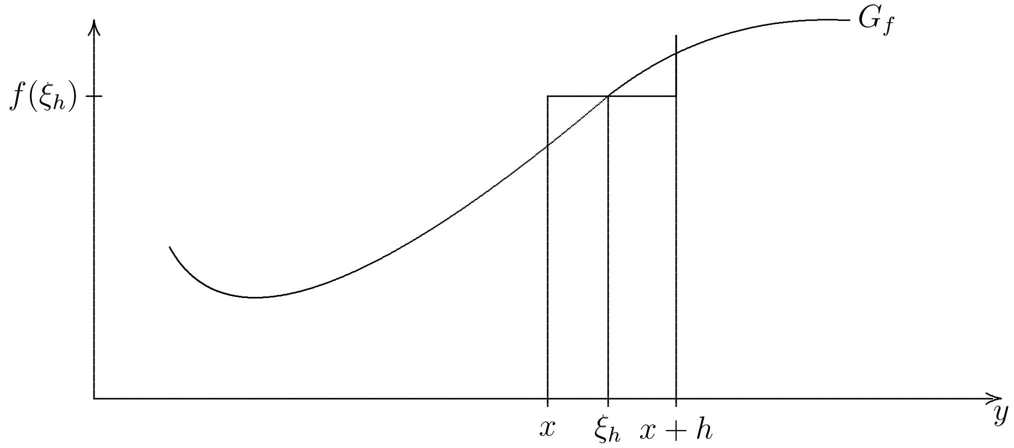
11.1.2 Bemerkung Ist x ein Endpunkt des Intervalls J, so ist die Ableitung als einseitige Ableitung definiert.

11.1.3 Beweis von $(A) \Rightarrow (B)$

Gemäß Aussage (A) sei

$$F(x) = \int_a^x f(t)\,dt$$

eine Integralfunktion zu f mit unterer Grenze a im Intervall J. Für ein beliebiges $x \in J$ müssen wir $F'(x) = f(x)$ beweisen.



Wir betrachten für $h > 0$, wobei $x, x+h \in J$, den Differenzenquotienten

$$\frac{F(x+h) - F(x)}{h} = \frac{\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt}{h} = \frac{\int_x^{x+h} f(t)\,dt}{h}.$$

Im Intervall $[x; x+h]$ gibt es gemäß Mittelwertsatz 10.7.1(iv) eine Stelle ξ, so dass

$$\int_x^{x+h} f(t)\,dt \quad = \quad f(\xi) \cdot h,$$

insgesamt also

$$\frac{F(x+h)-F(x)}{h} \quad = \quad \frac{\int_x^{x+h} f(t)\,dt}{h} \quad = \quad \frac{f(\xi)\cdot h}{h} \quad = \quad f(\xi).$$

Wir lassen jetzt $h \searrow 0$ gehen. Wegen $\xi \in [x; x+h]$ gilt

$$\lim_{h \searrow 0} \xi = x$$

und, da die Funktion f als stetig vorausgesetzt ist,

$$\lim_{h \searrow 0} f(\xi) = f(x).$$

Mit der Gleichung oben folgt insgesamt:

$$\lim_{h \searrow 0} \frac{F(x+h)-F(x)}{h} = \lim_{h \searrow 0} f(\xi) = f(x).$$

Eine analoge Überlegung mit $h < 0$ führt auf

$$\lim_{h \nearrow 0} \frac{F(x+h)-F(x)}{h} = f(x),$$

insgesamt also

$$F'(x) = \lim_{h \overset{\bullet}{\to} 0} \frac{F(x+h)-F(x)}{h} = f(x).$$

Zum Schluss bemerken wir noch, dass

$$F(a) = \int_a^a f(t)\,dt = 0.$$

Damit ist die Aussage (B) gezeigt.

11.1.4 Beweis von $(B) \Rightarrow (A)$

Gemäß Aussage (B) sei F eine Stammfunktion von f im Intervall J mit Nullstelle a:

$$F'(x) = f(x) \quad \text{für} \quad x \in J, \qquad F(a) = 0.$$

Nach der (bereits bewiesenen) Folgerung $(A) \Rightarrow (B)$ ist

$$G(x) = \int_a^x f(t)\,dt$$

ebenfalls eine Stammfunktion zu f mit $G(a) = 0$. Für die Differenz $H(x) = F(x) - G(x)$ dieser beiden Stammfunktionen von f gilt nun:

$$H'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

Nach Folgerung 9.8.14 ist diese Funktion konstant, und zwar wegen

$$H(a) = F(a) - G(a) = 0 - 0 = 0$$

gleich der Nullfunktion. Das aber bedeutet

$$F(x) = G(x) = \int_a^x f(t)\,dt.$$

Das ist die zu beweisende Aussage (A).

11.2 Anwendungen des HDI: Berechnung von Integralen

Man sagt, eine Funktion ist *elementar*, wenn sie als Verknüpfung (Grundrechenarten, Hintereinanderausführung) von Polynom–, Exponential–, trigonometrischen Funktionen und deren Umkehrfunktionen gebildet werden kann. Bei Zugrundelegung geeigneter Definitionsmengen sind elementare Funktionen differenzierbar, also

$$\text{elementar} \quad \Longrightarrow \quad \text{differenzierbar} \quad \Longrightarrow \quad \text{stetig} \quad \Longrightarrow \quad \text{integrierbar.}$$

Aufgrund der Differentiationsregeln ist einsichtig, dass die Ableitungsfunktion einer beliebigen elementaren Funktion wieder eine elementare Funktion ist.

Beachte aber, dass es elementare Funktionen gibt, deren Stammfunktion nicht elementar ist. Beispiele dafür sind die Funktionen $\sin(x^2)$ oder $x^\alpha \cdot (1-x)^\beta$.

Der HDI erleichtert erheblich die Berechnung von Integralen vieler elementarer Funktionen.

Als abkürzende Schreibweise für die Differenz von zwei Funktionswerten führen wir ein:

$$\Big[f(x)\Big]_{x=a}^{x=b} = \Big[f(x)\Big]_a^b = f(x)\Big|_{x=a}^{x=b} = f(x)\Big|_a^b := f(b) - f(a).$$

11.2.1 Beispiele Die Formel

$$\int_c^d x^\alpha \, dx = \frac{x^{\alpha+1}}{\alpha+1}\Big|_c^d$$

ist gültig, wenn die Ausdrücke auf den beiden Seiten wohldefiniert sind, das heißt, eine der folgenden Bedingungen erfüllt ist:

$$\begin{array}{l} \alpha \geq 0 \\ \alpha \in \mathbb{Z}, \alpha \leq -2 \quad \text{und} \quad (c,d \in \mathbb{R}^+ \text{ oder } c,d \in \mathbb{R}^-), \\ \alpha \in \mathbb{R} \setminus \{-1\} \quad \text{und} \quad c,d \in \mathbb{R}^+. \\ \alpha > -1 \qquad (**) \end{array}$$

Der Nachweis, dass die letzte Bedingung $(**)$ ebenfalls hinreichend ist für die obige Formel, erfordert eine Ausweitung des Integralbegriffs auf Funktionen f, die im Integrationsintervall $[c,d]$ Polstellen haben (Uneigentliche Integrale: Später).

Für den Sonderfall $\alpha = -1$ und ($c,d \in \mathbb{R}^+$ oder $c,d \in \mathbb{R}^-$) gilt:

$$\int_c^d x^{-1} \, dx = \ln|x| \Big|_c^d.$$

Für die natürliche Exponentialfunktion und die allgemeinere Exponentialfunktion $(a > 0)$ gelten:

$$\int_c^d \exp(x) \, dx = \exp(x)\Big|_c^d \qquad \int_c^d a^x \, dx = \frac{a^x}{\ln(a)}\Big|_c^d.$$

Aus den Ableitungsformeln für trigonometrischen Funktionen ergeben sich die folgenden Integrationsformeln:

$$\begin{aligned}
\int_c^d \sin(x)\,dx &= -\cos(x)\Big|_c^d \\
\int_c^d \cos(x)\,dx &= \sin(x)\Big|_c^d \\
\int_c^d \frac{1}{\cos^2(x)}\,dx &= \tan(x)\Big|_c^d, \qquad \text{wobei } (k-\frac{1}{2})\pi < c,d < (k+\frac{1}{2})\pi \text{ für ein } k \in \mathbb{Z}.
\end{aligned}$$

Aus den Ableitungsformeln für die Umkehrfunktionen trigonometrischen Umkehrfunktionen ergeben sich die folgenden Integrationsformeln:

$$\begin{aligned}
\int_c^d \frac{1}{1+x^2}\,dx &= \arctan(x)\Big|_c^d \\
\int_c^d \frac{1}{\sqrt{1-x^2}}\,dx &= \arcsin(x)\Big|_c^d, \qquad \text{wobei } -1 < c,d < +1.
\end{aligned}$$

11.2.2 Satz Zu $c_1, c_2, d_2, d_1 \in \mathbb{R}$, $c_1 < c_2 < d_2 < d_1$, gibt es eine $\mathcal{C}^\infty$–Funktion $\Omega_{c_1,c_2,d_2,d_1} : \mathbb{R} \to [0,1]$ mit

$$\Omega_{c_1,c_2,d_2,d_1}(x) \begin{cases} = 0, & \text{falls } x \le c_1 \text{ oder } x \ge d_1, \\ > 0, & \text{falls } x \in \,]c_1, c_2[\, \cup \,]d_2, d_1[, \\ = 1, & \text{falls } x \in [c_2, d_2]. \end{cases}$$

Man nennt solche Funktionen auch *bump–Funktionen*. Sie spielen in vielen Teilgebieten der Analysis eine wichtige Rolle.

11.2.3 Beweis Wir setzen zunächst für $c, d \in \mathbb{R}$, $c < d$

$$\lambda_{c,d}(x) := \begin{cases} = 0, & \text{falls } x \le c, \\ = \frac{\int_c^x \Lambda_{c,d}(t)\,dt}{\int_c^d \Lambda_{c,d}(t),dt}, & \text{falls } x \ge c, \end{cases}$$

wobei $\Lambda_{c,d}$ die in Sätzchen 9.11.6 konstruierte Funktion ist. Diese $\mathcal{C}^\infty$–Funktion ist konstant gleich 0 für $x \le c$ und konstant gleich 1 für $x \ge d$, dazwischen positiv, streng monoton steigend. Die durch

$$\Omega_{c_1,c_2,d_2,d_1}(x) = \lambda_{c_1,c_2}(x) \cdot [1 - \lambda_{d_2,d_1}(x)]$$

definierte Funktion leistet dann das gewünschte.

11.3 Partielle Integration

Mit Hilfe des HDI kann die Leibniz'sche Produktregel in eine Integrationstechnik umgeschrieben werden.

11.3.1 Satz: Partielle Integration

Für zwei stetig differenzierbare Funktionen $f, g : [c, d] \to \mathbb{R}$ gilt:

$$\int_c^d f(x)g'(x)\,dx + \int_c^d f'(x)g(x)\,dx \quad = \quad \Big[f(x)g(x)\Big]_{x=c}^{x=d}.$$

11.3.2 Beweis Ist die Funktion F definiert durch $F(x) = f(x)g(x) - f(c)g(c)$, so gilt:

$$F'(x) = f(x)g'(x) + f'(x)g(x) \qquad \text{und} \qquad F(c) = 0.$$

Aufgrund der Implikation $(B) \Rightarrow (A)$ aus dem HDI gilt:

$$\int_c^x [f(t)g'(t) + f'(t)g(t)]\,dt = F(x) = f(x)g(x) - f(c)g(c).$$

Für $x = d$ (und Ersetzung der Integrationsvariablen t durch x) ist dies die Aussage des Satzes.

11.3.3 Beispiele

(1) Für $c, d > 0$ ist

$$\begin{aligned}\int_c^d \ln(x)\,dx &= \int_c^d \underbrace{1}_{f'(x)} \cdot \underbrace{\ln(x)}_{g(x)}\,dx = [\underbrace{x}_{f(x)} \cdot \underbrace{\ln(x)}_{g(x)}]_c^d - \int_c^d \underbrace{x}_{f(x)} \cdot \underbrace{\frac{1}{x}}_{g'(x)}\,dx \\ &= [x \cdot \ln(x) - x]_c^d.\end{aligned}$$

(2) Für $c, d \in \mathbb{R}$ ist

$$\begin{aligned}\int_c^d \sin^2(x)\,dx &= \int_c^d \underbrace{\sin(x)}_{f'(x)} \cdot \underbrace{\sin(x)}_{g(x)}\,dx = [\underbrace{-\cos(x)}_{f(x)} \cdot \underbrace{\sin(x)}_{g(x)}]_c^d - \int_c^d \underbrace{-\cos(x)}_{f(x)} \cdot \underbrace{\cos(x)}_{g'(x)}\,dx \\ &= [\underbrace{-\cos(x)}_{f(x)} \cdot \underbrace{\sin(x)}_{g(x)}]_c^d + \int_c^d [1 - \sin^2(x)]\,dx.\end{aligned}$$

Daraus folgt:

$$2 \cdot \int_c^d \sin^2(x)\,dx = [x - \cos(x) \cdot \sin(x)]_c^d.$$

Allgemeiner kann man mit Hilfe der Zerlegung

$$\sin^m(x) = \sin(x) \cdot \sin^{m-1}(x)$$

Integrale von Sinus–Potenzen rekursiv bestimmen.

11.4 Die Substitutionsregel

11.4.1 Satz: Substitutionsregel

Zur Übersicht ein Diagramm:

$$\begin{array}{ccccc} [c,d] & \xrightarrow{\vartheta} & [a,b] & \xrightarrow{f} & \mathbb{R} \\ t & & x & & \end{array}$$

Es sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion.

Weiter sei $\vartheta : [c,d] \to [a,b]$ eine (in den Endpunkten einseitig) stetig differenzierbare Funktion, die das „Intervall $[c,d]$ in das Intervall $[a,b]$ (evtl. mehrfach übereinander) hineinlegt".

Dann gilt:

$$\int_c^d f(\vartheta(t)) \cdot \vartheta'(t)\, dt = \int_{\vartheta(c)}^{\vartheta(d)} f(x)\, dx.$$

11.4.2 Beweis Ist die Funktion F irgendeine Stammfunktion von f, so gilt gemäß Kettenregel

$$(F \circ \vartheta)'(t) \;=\; F'(\vartheta(t)) \cdot \vartheta'(t) \;=\; f(\vartheta(t)) \cdot \vartheta'(t)$$

Die Implikation $(B) \Rightarrow (A)$ aus dem HDI liefert:

$$\int_c^x f(\vartheta(t)) \cdot \vartheta'(t)\, dt \;=\; F(\vartheta(x)) - F(\vartheta(c)) \;=\; \int_{\vartheta(c)}^{\vartheta(x)} f(t)\, dt$$

Mit $x = d$ (und Ersetzung der Integrationsvariablen t durch x auf der rechten Seite) folgt der Satz.

11.4.3 Beispiele

(1) Die lineare Transformation $\vartheta : t \mapsto x = \alpha t + \beta$ für die Definitionsmenge führt auf die Formel

$$\int_c^d \alpha \cdot f(\alpha t + \beta)\, dt \;=\; \int_{\alpha c+\beta}^{\alpha d+\beta} f(x)\, dx.$$

(1b) Beachte dabei den Spezialfall $\alpha = -1$ / $\beta = 0$, er entspricht der Spiegelung des Graphen an der y-Achse.

$$\int_c^d f(t)\, dt \;=\; -\int_c^d (-f(t))\, dt \;=\; -\int_{-c}^{-d} f(-x)\, dx \;=\; \int_{-d}^{-c} f(-x)\, dx$$

(1c) Beachte weiter den Spezialfall $\alpha = 1$ (Verschiebung).

$$\int_c^d f(t+\beta)\, dt \;=\; \int_{c+\beta}^{d+\beta} f(x)\, dx.$$

(2) Beispiel: Für $\alpha \neq 0$ ist

$$\int_c^d (\alpha t + \beta)^2\, dt = \frac{1}{\alpha} \int_{\alpha c+\beta}^{\alpha d+\beta} x^2\, dx = \frac{(\alpha d + \beta)^3 - (\alpha c + \beta)^3}{3\alpha}.$$

11.4.4 Satz: Logarithmische Integration

Es sei $g : [c, d] \to \mathbb{R}^+$ (oder $\mathbb{R}^-$) differenzierbar mit stetiger Ableitung. Dann gilt

$$\int_c^d \frac{g'(x)}{g(x)}\, dx = (\ln|g(x)|)\Big|_c^d.$$

11.4.5 Beweis Dies ist eine direkte Anwendung der Substitutionsregel mit den Ersetzungen $f(x) \rightsquigarrow \frac{1}{x}$ und $\vartheta(t) \rightsquigarrow g(t)$.

$$\int_c^d \frac{g'(t)}{g(t)}\, dt = \int_c^d f(g(t)) \cdot g'(t)\, dt = \int_{g(c)}^{g(d)} \frac{1}{x}\, dx = \ln|x|\Big|_{x=g(c)}^{x=g(d)} = \ln|g(t)|\Big|_{t=c}^{t=d}.$$

Wechsle zum Schluss einfach wieder die Buchstaben aus.

11.4.6 Beispiel Es gilt für $c, d \in \,]-\frac{\pi}{2}, +\frac{\pi}{2}[$

$$\int_c^d \tan(x)\, dx = -\int_c^d \frac{-\sin(x)}{\cos(x)}\, dx = -(\ln|\cos(x)|)\Big|_c^d.$$

11.5 Nachtrag: Das uneigentliche Integral

11.5.1 Definition: Uneigentliches Integral

(1) Es sei die (nicht notwendig beschränkte) Funktion $f : [c, v[\to \mathbb{R}$ (mit $v > c$ oder $v = +\infty$) gegeben. Ist

- $f : [c, d] \to \mathbb{R}$ integrierbar für alle $d \in [c, v[$ und
- existiert der (endliche) Funktionsgrenzwert

$$\lim_{d \nearrow v} \int_c^d f(x)\, dx,$$

so heißt dieser Grenzwert das *uneigentliche Integral von f über* $[c, v[$. Man sagt dazu auch, dass das uneigentliche Integral

$$\int_c^v f(x)\, dx$$

existiert.

(2) Die duale Definition gilt für eine Funktion $f :]u, d] \to \mathbb{R}$ mit $u < d$ oder $u = -\infty$.

(3) Ist $f :]u, v[\to \mathbb{R}$ mit $u < v$ oder $u = -\infty$ oder $v = +\infty$, so ist das uneigentliche Integral

$$\int_u^v f(x)\, dx \ := \int_u^E f(x)\, dx + \int_E^v f(x)\, dx \qquad (*)$$

mit irgendeinem $E \in]u, v[$ — per definitionem — existent, wenn die beiden Integrale auf der rechten Seite in $(*)$ existieren.

11.5.2 Beachte

- Es existiert der Folgengrenzwert

$$\lim_{n \to \infty} \int_0^{2\pi n} \sin(x)\, dx,$$

nicht aber das uneigentliche Integral $\int_0^\infty \sin(x)\, dx$.

- Es existiert der Folgengrenzwert

$$\lim_{n \to \infty} \int_{-n}^{+n} x\, dx,$$

nicht aber das uneigentliche Integral $\int_{-\infty}^{+\infty} x\, dx$.

11.5.3 Beispiele (1) Es sei $\alpha \in \mathbb{R}$ ein Exponent. Es gilt für $d > 1$:

$$\int_1^d x^\alpha\,dx = \begin{cases} \frac{x^{\alpha+1}}{\alpha+1}\Big|_1^d = \frac{d^{\alpha+1}-1}{\alpha+1} \stackrel{d\to\infty}{\longrightarrow} \frac{1}{|\alpha+1|}, & \text{falls } \alpha < -1, \\ \ln x\Big|_1^d = \ln d \stackrel{d\to\infty}{\longrightarrow} +\infty, & \text{falls } \alpha = -1, \\ \frac{x^{\alpha+1}}{\alpha+1}\Big|_1^d = \frac{d^{\alpha+1}-1}{\alpha+1} \stackrel{d\to\infty}{\longrightarrow} +\infty, & \text{falls } \alpha > -1, \end{cases}$$

also im Fall $\alpha < -1$

$$\int_1^\infty x^\alpha\,dx = \frac{1}{|\alpha+1|}.$$

„Eine nicht-beschränkte Fläche hat endlichen Flächeninhalt".

(2) Für $\alpha \in \mathbb{R}$ und $c > 0$ ist

$$\int_c^1 x^\alpha\,dx = \begin{cases} \frac{x^{\alpha+1}}{\alpha+1}\Big|_c^1 = \frac{1-c^{\alpha+1}}{\alpha+1} \stackrel{c\searrow 0}{\longrightarrow} +\infty, & \text{falls } \alpha < -1, \\ \ln x\Big|_c^1 = -\ln c \stackrel{c\searrow 0}{\longrightarrow} +\infty, & \text{falls } \alpha = -1, \\ \frac{x^{\alpha+1}}{\alpha+1}\Big|_c^1 = \frac{1-c^{\alpha+1}}{\alpha+1} \stackrel{c\searrow 0}{\longrightarrow} \frac{1}{\alpha+1}, & \text{falls } \alpha > -1. \end{cases}$$

(3) Für alle $\alpha \in \mathbb{R}$ gilt

$$\int_0^\infty x^\alpha\,dx = +\infty.$$

11.5.4 Satz: Reihen–Integral–Vergleich

Es sei $f : [1,\infty[\to \mathbb{R}_0^+$ eine monoton fallende Funktion. Dann gilt

$$\sum_{k=1}^\infty f(k) \quad \text{konvergent} \qquad \Longleftrightarrow \qquad \int_1^\infty f(x)\,dx \quad \text{konvergent.}$$

11.5.5 Beweis Für festes n seien $\mathcal{Z}_n = \{0, 1, \ldots, n\}$ die „Länge–Eins–Zerlegung" des Intervalls $[0, n]$. Dann gilt einfach

$$\sum_{k=2}^n f(k) = \underline{\mathcal{S}}(f, \mathcal{Z}_n) \le \underline{\int_1^n} f(x)\,dx \le \int_1^n f(x)\,dx \le \overline{\int_1^n} f(x)\,dx \le \overline{\mathcal{S}}(f, \mathcal{Z}_n) = \sum_{k=1}^{n-1} f(k).$$

Reihe und uneigentliches Integral majorisieren sich wechselseitig.

11.5.6 Beispiel

Beispiel (1) oben sagt aus, dass $\int_1^\infty x^\alpha\,dx$ genau dann konvergiert, wenn $\alpha < -1$. Also konvergiert auch $\sum_{k=1}^\infty k^\alpha$ genau für $\alpha < -1$. Das ist uns schon aus der Übungsaufgabe über den Cauchy–Verdichtungssatz bekannt.

12 Gleichmäßige Konvergenz von Funktionenfolgen

In diesem Abschnitt sei M ein metrischer Raum, $\mathbb{K}$ sei einer der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$. Denken Sie einfach, dass M ein Intervall oder eine andere beliebige Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$ ist.

Wir untersuchen eine Folge von Funktionen $(f_n)_{n\in\mathbb{N}}$ mit $f_n : M \to \mathbb{K}$. Es stellt sich dabei vor allem die Frage nach der Konvergenz.

12.1 Definition

Eine weniger wichtige Definition ist zunächst:

12.1.1 Definition: Punktweise konvergente Funktionenfolge Die Funktionenfolge (f_n) heißt *punktweise konvergent* in einem festen $x \in M$, wenn die Folge $(f_n(x))_{n\in\mathbb{N}}$ konvergent in $\mathbb{K}$ ist. Sie heißt *punktweise konvergent* (schlechthin), wenn dies für alle $x \in M$ zutrifft. Es kann dann die *Grenzfunktion*

$$f(x) \ := \ \lim_{n\to\infty} f_n(x)$$

mit Definitionsmenge M gebildet werden.

12.1.2 Beispiele

(1) Es sei $M = \,]-1,+1]$ und $f_n(x) \ := \ x^n$. Es gilt

$$\lim_{n\to\infty} x^n = \begin{cases} 0, & \text{falls } -1 < x < +1, \\ 1, & \text{falls } x = +1. \end{cases}$$

Die Grenzfunktion $f : \,]-1,+1] \to \mathbb{R}$ ist nicht stetig in $x = 1$, obwohl dies für alle Funktionen der Folge zutrifft.

(2a) Für die auf $[0,1]$ definierten Treppenfunktionen

$$\varphi_n(x) \ := \ \begin{cases} n, & \text{falls } x \in \,]0, \frac{1}{n}[, \\ 0, & \text{sonst,} \end{cases}$$

gilt $\int_0^1 \varphi_n(x)\,dx = 1$. Für jedes feste $x \in [0,1]$ gilt aber

$$\varphi(x) \ := \ \lim_{n\to\infty} \varphi_n(x) = 0$$

und daher $\int_0^1 \varphi(x)\,dx = 0$.

(2b) Die Tatsache, dass die Treppenfunktionen φ_n nicht stetig sind, ist dabei ohne Bedeutung. Auch die $\mathcal{C}^\infty$–glatten normierten „Bump–Funktionen" (vgl. Satz 11.2.2)

$$\widetilde{\Omega}_n : \begin{cases} [0,1] & \to & \mathbb{R} \\ x & \mapsto & \dfrac{\Omega_{\frac{1}{n},\frac{2}{n},\frac{3}{n},\frac{4}{n}}(x)}{\int_0^1 \Omega_{\frac{1}{n},\frac{2}{n},\frac{3}{n},\frac{4}{n}}(t)\,dt} \end{cases}$$

haben die in (2a) beschriebenen Eigenschaften.

(3) Die Funktionen

$$f_n(x) = |x|^{1+\frac{1}{n}}, \qquad n \in \mathbb{N},$$

sind alle differenzierbar. Für $n = 1$ ergibt sich als Graph die Normalparabel, für $n \to \infty$ konvergiert die Funktionenfolge punktweise gegen die nicht differenzierbare Betragsfunktion.

Wie kann man gewährleisten, dass sich bestimmte Eigenschaften der Funktionen f_n in der Folge auf die Grenzfunktion f übertragen?

12.1.3 Definition:

(1) Ist $f : M \to \mathbb{K}$ eine beschränkte Funktion, so heißt

$$\|f\| \ := \ \sup_{x \in M} \{|f(x)|\}$$

die Supremumsnorm von f.
(Lineare Algebra: Das ist eine Norm auf dem Vektorraum der beschränkten Funktionen.)

(2) Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ heißt *gleichmäßig konvergent (gegen die Grenzfunktion f)*, wenn die Zahlenfolge $(\|f_n - f\|)_{n\in\mathbb{N}}$ eine Nullfolge ist:

$$\lim_{n\to\infty} \|f_n - f\| = 0.$$

12.1.4 Bemerkungen

1. Man kann die Definition der gleichmäßigen Konvergenz auch ohne Bezugnahme auf die Supremumsnorm formulieren:

   Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ heißt gleichmäßig konvergent gegen die Grenzfunktion f, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n \geq N$ gilt:

   $$|f_n(x) - f(x)| \leq \varepsilon \qquad \text{für alle } x \in M.$$

   Wesentlich ist, dass das N unabhängig von $x \in M$ gefunden werden muss.

2. Für den Fall $M \subseteq \mathbb{R}$ und $\mathbb{K} = \mathbb{R}$ kann man diese Definition anhand der Graphen veranschaulichen: Zu jedem ε–Streifen um den Graphen von f muss es ein $N \in \mathbb{N}$ geben, so dass für alle $n \geq N$ die Graphen von f_n in diesem Streifen drin liegen müssen.

3. Aus der Definition ist sofort ersichtlich, dass eine gleichmäßig konvergente Funktionenfolge auch punktweise gegen die gleiche Grenzfunktion konvergiert.

12.2 Eigenschaften der Grenzfunktion

12.2.1 Satz: Stetigkeit der Grenzfunktion

Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger Funktionen $M \to \mathbb{K}$, die gleichmäßig gegen die Grenzfunktion f konvergiert. Dann ist f stetig.

12.2.2 Beweis Es sei $x \in M$ und $\varepsilon > 0$ vorgegeben. Wir wollen zeigen, dass es ein $\delta > 0$ gibt, so dass

$$|f(x) - f(a)| \leq \varepsilon, \qquad \text{falls } |x-a| \leq \delta.$$

Da (f_n) gleichmäßig konvergiert, gibt es zu $\frac{\varepsilon}{3}$ ein $N \in \mathbb{N}$, so dass

$$|f_n(z) - f(z)| \leq \frac{\varepsilon}{3}, \qquad \text{für alle } z \in M, \qquad \text{falls } n \geq N.$$

Insbesondere ist

$$|f_N(z) - f(z)| \leq \frac{\varepsilon}{3} \qquad \text{für alle } z \in M. \qquad (*)$$

Da f_N stetig ist, gibt es zu $\frac{\varepsilon}{3}$ ein $\delta > 0$, so dass

$$|f_N(x) - f_N(a)| \leq \frac{\varepsilon}{3}, \qquad \text{falls } |x-a| \leq \delta. \qquad (**)$$

Insgesamt gilt für alle $x \in M$ mit $|x-a| \leq \delta$

$$|f(x) - f(a)| \leq \underbrace{|f(x) - f_N(x)|}_{(*)\,\leq\frac{\varepsilon}{3}} + \underbrace{|f_N(x) - f_N(a)|}_{(**)\,\leq\frac{\varepsilon}{3}} + \underbrace{|f_N(a) - f(a)|}_{(*)\,\leq\frac{\varepsilon}{3}} \leq \varepsilon.$$

12.2.3 Satz: Integration und Differentiation der Grenzfunktion

Es sei M ein Intervall von $\mathbb{R}$ und (f_n) eine Folge von stetigen Funktionen $M \to \mathbb{R}$, die gleichmäßig gegen die Grenzfunktion f konvergiert.

(i) Sind $c, d \in M$, so gilt für die zugehörigen Integrale:

$$\int_c^d f(x)\,dx = \lim_{n\to\infty} \int_c^d f_n(x)\,dx.$$

(Integration und Gleichmäßig-Limes-Bildung sind vertauschbar.)

(ii) Für jedes $n \in \mathbb{N}$ sei F_n eine Stammfunktionen von f_n und F eine Stammfunktion von f. Konvergiert die Folge der Stammfunktionen an (mindestens) einer Stelle $a \in [c,d]$ gegen den zugehörigen Wert der Stammfunktion der Grenzfunktion, d.h.

$$\lim_{n\to\infty} F_n(a) = F(a),$$

so konvergiert die Folge der Stammfunktionen F_n gleichmäßig gegen die Stammfunktion F.

(iii) Aus der Differenzierbarkeit der Funktionen f_n kann — auch bei gleichmäßiger Konvergenz von (f_n) — **nicht** auf die Differenzierbarkeit der Grenzfunktion f geschlossen werden.

12.2.4 Beweis (i) Es sei $\varepsilon > 0$ vorgegeben. Für $n \in \mathbb{N}$ gilt gemäß Satz 10.6.2 und Satz 10.7.1(ii)

$$\begin{aligned}\Big|\int_c^d f(x)\,dx - \int_c^d f_n(x)\,dx\Big| &= \Big|\int_c^d [f(x)-f_n(x)]\,dx\Big| \le \int_c^d |f(x)-f_n(x)|\,dx \\ &\le (d-c)\cdot \sup_{x\in[c,d]}\{|f(x)-f_n(x)|\}.\end{aligned}$$

Es sei dann $\varepsilon > 0$ gegeben. Da aufgrund der gleichmäßigen Konvergenz das Supremum kleiner als $\frac{\varepsilon}{(d-c)}$ gemacht werden kann, kann auch die Differenz der Integrale kleiner als ε gemacht werden.

(ii) Für die Stammfunktionen F_n und F gilt aufgrund des HDI:

$$F_n(x) = F_n(a) + \int_a^x f_n(t)\,dt, \qquad F(x) = F(a) + \int_a^x f(t)\,dt.$$

Es ist dann

$$\begin{aligned}\Big|F(x) - F_n(x)\Big| &= \Big|F(a) + \int_a^x f(t)\,dt - F_n(a) - \int_a^x f_n(t)\,dt\Big| = \\ &\le \Big|F(a) - F_n(a)\Big| + \Big|\int_a^x [f(t) - f_n(t)]\,dt\Big|.\end{aligned}$$

Zu einem vorgegebenen $\varepsilon > 0$ kann — gemäß Voraussetzung des Satzes und Aussage (i) — ein $N \in \mathbb{N}_0$ gefunden werden kann, so dass für alle $n \ge N$ gilt

$$\Big|\int_a^x [f(t) - f_n(t)]\,dt\Big| \le \frac{\varepsilon}{2}, \qquad \Big|F(a) - F_n(a)\Big| \le \frac{\varepsilon}{2}$$

gilt und damit für alle $x \in [c, d]$

$$\Big|F(x) - F_n(x)\Big| \le \varepsilon.$$

(iii) Wir zeigen, dass die Funktionenfolge $f_n(x) = |x|^{1+\frac{1}{n}}$ aus Beispiel (4) sogar gleichmäßig gegen die Betragsfunktion konvergiert. Zunächst ist

$$\lim_{x\to\infty}\frac{1}{x} = 0 \quad\Longrightarrow\quad \lim_{x\to\infty}\frac{1}{x}\cdot\ln(\frac{1}{x}) = -\lim_{x\to\infty}\frac{\ln x}{x} \overset{\text{l'H}}{=} 0.$$

Daraus folgt für die Funktion

$$\lim_{x\searrow 0}(x\cdot\ln x) = \lim_{x\to\infty}\frac{1}{x}\cdot\ln(\frac{1}{x}) = 0,$$

so dass die Funktion $f(x) = x \cdot \ln x$ auf dem Intervall $]0, 1]$ ein Betragsmaximum hat:

$$M := \max_{x\in]0,1]}\{|x\cdot\ln x|\}.$$

Für ein festes $x \in\ ]0, 1]$ und $\alpha \in [1, 2]$ gibt es aufgrund des Mittelwertsatzes ein $\xi \in [1, \alpha] \subseteq [1, 2]$, so dass

$$\Big|x - x^\alpha\Big| \le \frac{dx^\alpha}{d\alpha}\Big|_{\alpha=\xi}\cdot|1-\alpha| = |\ln x\cdot x^\xi|\cdot|1-\alpha| \le |\ln x\cdot x|\cdot|1-\alpha| \le M\cdot|1-\alpha|.$$

Daraus folgt, dass für alle $x \in [-1, +1]$ gilt:

$$\Big||x| - |x|^\alpha\Big| \le M\cdot|1-\alpha|.$$

Das aber bedeutet, dass für $\alpha = 1 + \frac{1}{n} \to 1$ die Funktionenfolge gleichmäßig konvergiert.

12.3 Gleichmäßige Konvergenz einer Funktionenreihe

12.3.1 Satz: Weierstrass–Kriterium

Es seien $g_k : M \to \mathbb{K}$ die Glieder einer Folge $(g_k)_{k\in\mathbb{N}_0}$ von Funktionen und

$$f_n \ := \ \sum_{k=0}^{n} g_k$$

die zugehörige Partialsummenfolge. Wenn die Zahlenreihe

$$\sum_{k=0}^{\infty} \|g_k\|$$

konvergiert, dann konvergiert die Partialsummenfolge (f_n) gleichmäßig (und absolut). Man sagt, die Funktionenreihe *konvergiert gleichmäßig absolut*.

12.3.2 Beweis Für jedes feste $x \in M$ gilt

$$|g_k(x)| \ \leq \ \sup_{t\in M}\{|g_k(t)|\} \ = \ \|g_k\|.$$

Daher stellt die Zahlenreihe $\sum_{k=0}^{\infty} \|g_k\|$ eine konvergente Majorante für die Reihe

$$\sum_{k=0}^{\infty} g_k(x)$$

dar, die also nach dem Majorantenkriterium 6.3.4 absolut konvergiert. Wir definieren die Funktion f durch diese Grenzwerte

$$f(x) \ := \ \lim_{n\to\infty} f_n(x) \ = \ \sum_{k=0}^{\infty} g_k(x)$$

und zeigen, dass die Funktionenfolge f_n gleichmäßig gegen f konvergiert. Zu einem vorgegebenen $\varepsilon > 0$ gibt es nach Voraussetzung des Satzes ein $N \in \mathbb{N}$, so dass für alle $n \geq N$

$$\sum_{k=n}^{\infty} \|g_k\| \ \leq \ \sum_{k=N}^{\infty} \|g_k\| < \varepsilon.$$

Für alle $n \geq N$ und alle $x \in M$ folgern wir weiter (mit der „∞–Ecks–Ungleichung" Satz 6.3.2)

$$\Big|f(x) - f_n(x)\Big| = \Big|\sum_{k=n+1}^{\infty} g_k(x)\Big| \leq \sum_{k=n+1}^{\infty} |g_k(x)| \leq \sum_{k=n+1}^{\infty} \|g_k\| < \varepsilon.$$

13 Potenzreihen

13.1 Definition und erste Beispiele

In diesem Abschnitt sei $\mathbb{K}$ wieder einer der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$.

13.1.1 Definition: Ist $(p_k)_{k\in\mathbb{N}_0}$ eine Folge von Zahlen aus $\mathbb{K}$ und $a \in \mathbb{K}$, so heißt der Ausdruck

$$\sum_{k=0}^{\infty} p_k(z-a)^k$$

(formale) Potenzreihe mit der Koeffizientenfolge $(p_k)_{k\in\mathbb{N}_0}$ *und Entwicklungsstelle* a. Man spricht von reellen oder komplexen Potenzreihen.

13.1.2 Bemerkungen

1. Das Zusatzattribut ,,formal" soll zum Ausdruck bringen, dass die Frage der Konvergenz noch nicht geklärt ist und deshalb noch nicht von einer Funktion (mit Definitionsvariabler z) gesprochen werden kann.

   Erst, wenn für z eine konkrete reelle oder komplexe Zahl eingesetzt wird, stellt sich die Frage nach der Konvergenz.

2. Oft wird in der Theorie der Potenzreihen nur die Entwicklungsstelle $a = 0$ betrachtet, die Potenzreihen haben dann die Form

$$\sum_{k=0}^{\infty} p_k\, z^k.$$

   Dies bedeutet keine wesentliche Einschränkung, da man mittels der Substitution $w = z - a$ immer zwischen den beiden Fällen wechseln kann.

3. Potenzreihen können addiert und mit einer Zahl aus $\mathbb{K}$ (Skalar) vervielfacht werden, indem man diese Operation mit den Koeffizientenfolgen durchführt.

$$\sum_{k=0}^{\infty} p_k\, z^k + \sum_{k=0}^{\infty} q_k\, z^k = \sum_{k=0}^{\infty}(p_k+q_k)\, z^k, \qquad \alpha\cdot\sum_{k=0}^{\infty} p_k\, z^k = \sum_{k=0}^{\infty}(\alpha\cdot p_k)\, z^k.$$

4. Durch Cauchy–Produktbildung (vgl. Satz 6.7.2) entsteht aus zwei Potenzreihen $\sum_{k=0}^{\infty} p_k\, z^k$ und $\sum_{k=0}^{\infty} q_k\, z^k$ die Potenzreihe

$$\sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{k} p_\ell\, z^\ell\cdot q_{k-\ell}\, z^{k-\ell}\right) = \sum_{k=0}^{\infty}\left(\sum_{\ell=0}^{k} p_\ell\cdot q_{k-\ell}\right) z^k.$$

   Den beiden Koeffizientenfolgen (p_k) und (q_k) wird dabei die Koeffizientenfolge $(\sum_{\ell=0}^{k} p_\ell\cdot q_{k-\ell})$ zugeordnet. Man spricht hier auch von *Faltung* der beiden Folgen.

13.1.3 Beispiele

(1) Jedes Polynom

$$p_0 + p_1 z + p_2 z^2 + \ldots + p_n z^n$$

stellt eine Potenzreihe (mit $p_k = 0$ für $k \geq n+1$) dar. Offenbar ,,konvergiert” sie für jedes $z \in \mathbb{R}$.

(2) Die Exponentialreihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} z^k$$

ist eine Potenzreihe mit $p_k = \frac{1}{k!}$. Wir wissen bereits, dass sie für jedes $z \in \mathbb{C}$ konvergiert.

(3) Kosinus und Sinus sind durch die Potenzreihen

$$\sum_{k=0}^{\infty} p_k x^k \qquad \text{mit} \qquad p_k = \begin{cases} \frac{(-1)^\ell}{(2\ell)!}, & \text{falls } k = 2\ell \text{ gerade,} \\ 0, & \text{sonst,} \end{cases}$$

bzw.

$$\sum_{k=0}^{\infty} p_k x^k \qquad \text{mit} \qquad p_k = \begin{cases} \frac{(-1)^\ell}{(2\ell+1)!}, & \text{falls } k = 2\ell+1 \text{ ungerade,} \\ 0, & \text{sonst,} \end{cases}$$

gegeben. Sie konvergieren für jedes $x \in \mathbb{R}$.

(4) Setzt man $p_k = 1$ für alle $k \in \mathbb{N}_0$, so entsteht die geometrische Reihe

$$\sum_{k=0}^{\infty} z^k.$$

Wir wissen bereits, dass sie für jedes $z \in \mathbb{C}$ mit $|z| < 1$ konvergiert und für $|z| \geq 1$ divergiert.

(5) Die Potenzreihe

$$\sum_{k=0}^{\infty} k^k z^k$$

konvergiert nur für $z = 0$. Für jedes $z \neq 0$ gibt es nämlich ein $N \in \mathbb{N}$, so dass $N \cdot |z| \geq 2$. Dann ist für $k \geq N$

$$|k^k z^k| = |k \cdot z|^k \geq |N \cdot z|^k \geq 2^k.$$

Also ist die Reihe $\sum_{k=0}^{\infty} 2^k$ eine ,,divergente Minorante”.

13.2 Der Konvergenzradius

Eine erste wesentliche Beobachtung über die Konvergenz von Potenzreihen ist in dem folgenden Satz enthalten.

13.2.1 Satz: Konvergenzradius

(i) Jede Potenzreihe konvergiert an ihrer Entwicklungsstelle.

(ii) Konvergiert die Potenzreihe $\sum_{k=0}^{\infty} p_k z^k$ für ein festes $w \in \mathbb{C}$, so konvergiert sie auch (absolut) für alle $z \in \mathbb{C}$ mit $|z| < |w|$.

(iii) Konvergiert die Potenzreihe $\sum_{k=0}^{\infty} p_k (z-a)^k$ für ein festes $w \in \mathbb{C}$, so konvergiert sie auch (absolut) für alle $z \in \mathbb{C}$ mit $|z-a| < |w-a|$.

13.2.2 Beweis (i) muss nicht bewiesen werden.

(ii) O.B.d.A. ist $w \neq 0$. Da die Reihe $\sum_{k=0}^{\infty} p_k\, w^k$ konvergiert, ist $(p_k\, w^k)_{k\in\mathbb{N}_0}$ eine Nullfolge und daher beschränkt:

$$|p_k\, w^k| \leq M \ \text{ für alle } \ k \in \mathbb{N}_0.$$

Für alle z mit $|z| < |w|$ gilt dann

$$|p_k\, z^k| = |p_k\, w^k| \cdot \left|\frac{z}{w}\right|^k \leq M \cdot q^k, \qquad \text{wobei } q := \left|\frac{z}{w}\right| < 1.$$

Die Reihe $\sum_{k=0}^{\infty} M \cdot q^k$ ist eine geometrische (absolut konvergente) Reihe und stellt daher eine Majorante für die Reihe $\sum_{k=0}^{\infty} p_k\, z^k$ dar.

Für den Beweis von (iii) muss man in (ii) nur z durch $z-a$ und w durch $w-a$ ersetzen.

13.2.3 Bemerkung: Kreisscheibe Aus dem Lemma folgt unmittelbar, dass die Menge der $z \in \mathbb{C}$, für die eine Potenzreihe $\sum_{k=0}^{\infty} p_k\, (z-a)^k$ konvergiert, eine in a zentrierte Kreisfläche oder die gesamte komplexe Ebene ist. Das Lemma macht aber keinerlei Aussage darüber, in welchen Punkten der begrenzenden Kreislinie die Potenzreihe konvergiert.

13.2.4 Definitionen

(1) Für eine gegebene Potenzreihe $\sum_{k=0}^{\infty} p_k(z-a)^k$ heißt die Menge

$$\mathcal{K} := \left\{ z \in \mathbb{C} \;\middle|\; \sum_{k=0}^{\infty} p_k\, (z-a)^k \text{ konvergiert } \right\}$$

die *Konvergenzmenge* oder *Konvergenzkreisscheibe (der Potenzreihe)*.

(2) Die Zahl ϱ mit $0 \leq \varrho \leq +\infty$, definiert durch

$$\sup \left\{ |z-a| \;\middle|\; z \in \mathcal{K} \right\}$$

heißt *Konvergenzradius (der Potenzreihe)*.

(3) Bei reellen Potenzreihen ist die Konvergenzkreisscheibe ein Intervall. Man spricht aber auch hier vom Konvergenzradius.

13.2.5 Beispiele

(1) Polynome, die Exponentialfunktion und die beiden Funktionen sin und cos haben den Konvergenzradius $\varrho = +\infty$.

(2) Ist r eine beliebige positive Zahl, so konvergiert die Potenzreihe

$$\sum_{k=0}^{\infty} \left(\frac{z}{r}\right)^k$$

genau dann, wenn

$$\left|\frac{z}{r}\right| < 1 \quad \Longleftrightarrow \quad |z| < r.$$

Der Konvergenzradius ist also $\varrho = r$.

(3) Wir betrachten die Potenzreihe

$$\sum_{k=1}^{\infty} \frac{1}{k} z^k$$

Sie wird für $|z| < 1$ von der geometrischen Reihe majorisiert. Wir wissen, dass sie für $z = 1$ divergiert (harmonische Reihe). Es ist also $\varrho = 1$.

And der Stelle $z = 1$ divergiert sie, bei $z = -1$ konvergiert sie (Leibniz–Kriterium).

13.2.6 Satz: Algebra der Potenzreihen Haben die beiden Potenzreihen $\sum_{k=0}^{\infty} p_k z^k$ und $\sum_{k=0}^{\infty} q_k z^k$ Konvergenzradien größer-gleich R, so haben auch die Potenzreihen ($\alpha \in \mathbb{C}$)

$$\sum_{k=0}^{\infty}(p_k + q_k) z^k, \qquad \sum_{k=0}^{\infty}(\alpha \cdot p_k) z^k, \qquad \sum_{k=0}^{\infty} \left(\sum_{\ell=0}^{k} p_\ell \cdot q_{k-\ell}\right) z^k$$

Konvergenzradien größer-gleich R.

13.2.7 Bemerkung Der Satz sagt, dass die Menge der Potenzreihen mit Konvergenzradius größer–gleich R eine $\mathbb{K}$–Algebra bilden. Eine $\mathbb{K}$–*Algebra* ist ein $\mathbb{K}$–Vektorraum V, auf dem zusätzlich eine (bilineare, assoziative) Multiplikation $V \times V \to V$ zweier Vektoren erklärt ist.

13.2.8 Beweis Ist $r < R$, so konvergieren die Zahlenreihen

$$\sum_{k=0}^{\infty} p_k r^k, \qquad \sum_{k=0}^{\infty} q_k r^k$$

absolut. Dann konvergieren aber auch die Zahlenreihen

$$\sum_{k=0}^{\infty}(p_k + q_k) r^k, \qquad \sum_{k=0}^{\infty}(\alpha \cdot p_k) r^k,$$

$$\sum_{k=0}^{\infty} \left(\sum_{\ell=0}^{k} p_\ell r^\ell \cdot q_{k-\ell} r^{k-\ell}\right) = \sum_{k=0}^{\infty} \left(\sum_{\ell=0}^{k} p_\ell \cdot q_{k-\ell}\right) r^k.$$

absolut für alle $r < R$, vgl. auch den Satz 6.9.2 über das Cauchy-Produkt. Deshalb muss der Konvergenzradius der im Satz genannten Reihen jeweils größer-gleich R sein.

13.3 Eigenschaften der Grenzfunktion einer Potenzreihe

13.3.1 Präposition Die beiden Potenzreihen

$$\sum_{k=0}^{\infty} p_k \, (z-a)^k \qquad \text{und} \qquad \sum_{k=0}^{\infty} k p_k \, (z-a)^k$$

haben den gleichen Konvergenzradius ϱ.

13.3.2 Beweis

(1) Wir bemerken zunächst, dass für festes $n \in \mathbb{N}_0$ und $b \in]0,1[$ gilt:

$$\lim_{x\to\infty} x^n \cdot b^x = |\ln b|^{-n} \cdot \lim_{x\to\infty} (|\ln b| x)^n e^{-x\cdot|\ln b|} = |\ln b|^{-n} \cdot \lim_{y\to\infty} y^n e^{-y} \overset{9.11.4}{=} 0.$$

(2) Es genügt, den Fall $a = 0$ zu betrachten. Für ein beliebiges z ist

$$|p_k z^k| \le |k p_k \, z^k|,$$

so dass aus der absoluten Konvergenz der rechten Reihe in 13.3.1 mit dem Majorantenkriterium auf die Konvergenz der linken Reihe geschlossen werden kann. Deshalb ist der Konvergenzradius ϱ_r der rechten Potenzreihe kleiner oder gleich dem Konvergenzradius ϱ_ℓ der linken Potenzreihe.

(3) Es bleibt $\varrho_r \ge \varrho_\ell$ zu zeigen. Dazu sei w mit $0 < |w| < \varrho_\ell$ vorgegeben. Es existiert ein r mit $|w| < r < \varrho_\ell$, so dass die Reihe $\sum_{k=0}^{\infty} p_k \, r^k$ absolut konvergiert. Nach (1) (mit $n = 1$) ist die Folge

$$k \cdot \left|\frac{w}{r}\right|^k$$

eine Nullfolge, also durch M beschränkt. Dann gilt weiter

$$\Big|k p_k w^k\Big| = \Big|k \cdot (\frac{w}{r})^k \cdot p_k \, r^k\Big| \le M \Big|p_k \, r^k\Big|.$$

Damit haben wir eine absolut konvergente Majorante für die Reihe

$$\sum_{k=1}^{\infty} k p_k \, w^k$$

gefunden, sie konvergiert also absolut.

(4) Insgesamt haben wir bewiesen, dass

$$\varrho_r \ge |w| \qquad \text{für alle } w \text{ mit } |w| < \varrho_\ell$$

gilt. Das heißt aber gerade $\varrho_r \ge \varrho_\ell$.

13.3.3 Satz: Gleichmäßige Konvergenz einer Potenzreihe

Ist $r < \varrho$, so konvergiert die Potenzreihe $\sum_{k=0}^{\infty} p_k (z-a)^k$ gleichmäßig absolut auf der Kreisscheibe (bzw. dem Intervall)

$$B_r(a) = \left\{ z \in \mathbb{K} \;\middle|\; |z-a| \leq r \right\}.$$

13.3.4 Beweis (Nur für $a = 0$) Für alle z mit $|z| \leq r$ gilt

$$|p_k\, z^k| \leq |p_k r^k|$$

und daher

$$\|p_k\, z^k\| \;:=\; \sup_{z \in B_r(a)} \{|p_k\, z^k|\} \leq |p_k r^k|.$$

Damit stellt die Reihe $\sum_{k=0}^{\infty} |p_k r^k|$ eine konvergente Majorante für die Reihe

$$\sum_{k=0}^{\infty} \|p_k\, z^k\|$$

dar. Nach dem Weierstrass–Kriterium Satz 12.3.1 konvergiert die Potenzreihe $\sum_{k=0}^{\infty} p_k z^k$ gleichmäßig auf $B_r(a)$.

13.3.5 Bemerkung Beachten Sie den feinen Unterschied: Die Potenzreihe konvergiert für jedes z mit $|z-a| < \varrho$. Gleichmäßige Konvergenz liegt aber nur auf den durch $|z-a| \leq r$ gegebenen Mengen mit $r < \varrho$ vor.

13.3.6 Satz: Eigenschaften der Grenzfunktion einer Potenzreihe

Es sei $\sum_{k=0}^{\infty} p_k(z-a)^k$ eine Potenzreihe mit Konvergenzradius ϱ. Wir definieren durch

$$U_\varrho(a) \quad := \quad \{z \in \mathbb{K} | |z-a| < \varrho\}$$

das Innere der Konvergenzkreisscheibe (bzw. des Konvergenzintervalls) und dann die Grenzfunktion

$$f : \begin{cases} U_\varrho(a) & \to \quad \mathbb{K} \\ z & \mapsto \quad \sum_{k=0}^{\infty} p_k(z-a)^k \end{cases} .$$

Dann gilt:

(i) f ist stetig — auf $U_\varrho(a)$.

Ab jetzt sei $\mathbb{K} = \mathbb{R}$. Wir betrachten reelle Potenzreihen.

(ii) f ist auf jedem kompakten Intervall $[c,d] \subseteq]a-\varrho, a+\varrho[$ integrierbar. Das Integral darf gliedweise berechnet werden.

$$\begin{aligned} \int_c^d f(x)\,dx &= \int_c^d \Big(\sum_{k=0}^{\infty} p_k(x-a)^k\Big)\,dx = \sum_{k=0}^{\infty} \int_c^d p_k(x-a)^k\,dx \\ &= \sum_{k=0}^{\infty} \Big[\frac{p_k}{k+1}\,(x-a)^{k+1}\Big]_{x=c}^{x=d}. \end{aligned}$$

(iii) f besitzt eine Stammfunktion F auf $]a-\varrho, a+\varrho[$. Die durch $F(a) = 0$ festgelegte Stammfunktion ist gegeben durch die Potenzreihe

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

(„Gliedweise Stammfunktion"). Ihr Konvergenzradius ist ϱ.

(iv) f ist auf $]a-\varrho, a+\varrho[$ unendlich oft differenzierbar. Dabei gilt

$$\begin{aligned} f'(x) &= \sum_{k=1}^{\infty} kp_k\,(x-a)^{k-1} = \sum_{k=0}^{\infty}(k+1)p_{k+1}\,(x-a)^k \\ f^{(n)}(x) &= \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} p_k\,(x-a)^{k-n}, \qquad \text{insbesondere } f^{(n)}(a) = n!\,p_n. \end{aligned}$$

(„Gliedweise Ableitung"). Die Potenzreihen haben den Konvergenzradius ϱ.

13.3.7 Beweis (Generell nur für $a = 0$).

Für $r < \varrho$ ist die Funktionen–Partialsummenfolge

$$f_n(z) := \sum_{k=0}^{n} p_k \, z^k$$

gleichmäßig konvergent auf $[-r, +r]$. Damit sind die Aussagen der Sätze 12.2.1 und 12.2.3(i),(ii) anwendbar. Die Grenzfunktion hat also die in (i) – (iii) dieses Satzes 13.3.6 aufgezählten Eigenschaften auf $[-r, +r]$.

Da dies für alle $r < \varrho$ so ist, sind die Eigenschaften (i) – (iii) auch auf

$$]-\varrho, +\varrho[\; = \bigcup_{n \geq N} [-\varrho + \frac{1}{n}, +\varrho - \frac{1}{n}], \qquad (N \text{ groß genug})$$

erfüllt.

(iv) Die Präposition 13.3.1 zeigt, dass die in der ersten Zeile von (iv) angegebene Potenzreihe den gleichen Konvergenzradius ϱ hat wie die ursprüngliche Potenzreihe und daher für festes $r < \varrho$ gleichmäßig auf $[-r, +r]$ gegen eine Funktion g konvergiert. Das aber bedeutet gemäß Satz 12.2.3 (ii), dass die Grenzfunktion f der ursprünglichen Potenzreihe auf $[-r, +r]$ eine Stammfunktion von g ist. Demzufolge ist $g = f'$.

Es ist also auf jedem Intervall $[-r, +r]$ mit $r < \varrho$ die Funktion g eine Ableitungsfunktion von f. Das bedeutet, dass die Ableitungsfunktion auf $]-\varrho, +\varrho[$ existiert.

Die Aussagen über die höheren Ableitungen folgen daraus einfach mit Induktion.

13.4 Zwei Aussagen über den Konvergenzradius

Ist eine Folge $(a_k)_{k\in\mathbb{N}}$ gegeben, so können wir weitere Folgen bilden:

- $\sigma_\ell := \sup\limits_{k\geq\ell}\{a_k\}$ (Folge der „Rest–Suprema")
- $\iota_\ell := \inf\limits_{k\geq\ell}\{a_k\}$ (Folge der „Rest–Infima")

Es ist klar, dass $(\sigma_\ell)_{\ell\in\mathbb{N}}$ eine monoton fallende, $(\iota_\ell)_{\ell\in\mathbb{N}}$ eine monoton wachsende Folge mit Werten in $\overline{\mathbb{R}} = [-\infty, +\infty]$ ist.

Daraus folgt, dass jede der beiden Folgen einen (eigentlichen oder uneigentlichen) Grenzwert hat. Für sie gibt es eine besondere Bezeichnung.

13.4.1 Definition: Die beiden Grenzwerte heißen *Limes superior* bzw. *Limes inferior.*

$$\begin{aligned} \limsup_{k\to\infty} a_k &:= \lim_{\ell\to\infty} \sigma_\ell = \lim_{\ell\to\infty} \sup_{k\geq\ell}\{a_k\} \in [-\infty, +\infty], \\ \liminf_{k\to\infty} a_k &:= \lim_{\ell\to\infty} \iota_\ell = \lim_{\ell\to\infty} \inf_{k\geq\ell}\{a_k\} \in [-\infty, +\infty]. \end{aligned}$$

Wir halten einige Eigenschaften dieser Grenzwerte in einem Satz fest:

13.4.2 Satz: Eigenschaften des Limes superior und Limes inferior

(i) Der Limes superior S einer Folge ist durch die folgende Eigenschaft eindeutig festgelegt: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass

$$\begin{aligned} \alpha_k &\leq S + \varepsilon && \text{für alle } k \geq N, \\ \alpha_k &\geq S - \varepsilon && \text{für unendlich viele } k \geq N. \end{aligned}$$

(ii) Der Limes inferior I einer Folge ist durch die folgende Eigenschaft eindeutig festgelegt: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass

$$\begin{aligned} \alpha_k &\geq I - \varepsilon && \text{für alle } k \geq N, \\ \alpha_k &\leq I + \varepsilon && \text{für unendlich viele } k \geq N. \end{aligned}$$

(iii) Hat die Folge $(a_k)_{k\in\mathbb{N}}$ selbst einen eigentlichen oder uneigentlichen Grenzwert, so gilt

$$\liminf_{k\to\infty} a_k = \lim_{k\to\infty} a_k = \limsup_{k\to\infty} a_k.$$

13.4.3 Beispiel Es sei $(a_\ell)_{\ell\in\mathbb{N}}$ eine konvergente Folge mit Grenzwert a. Wir schieben Nullen dazwischen:

$$\widetilde{a}_k := \begin{cases} a_\ell, & \text{falls } k = 2\,\ell \text{ gerade }, \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt

$$\limsup_{k\to\infty} \widetilde{a}_k = \begin{cases} a, & \text{falls } a \geq 0, \\ 0, & \text{falls } a < 0, \end{cases} \qquad \liminf_{k\to\infty} \widetilde{a}_k = \begin{cases} a, & \text{falls } a \leq 0, \\ 0, & \text{falls } a > 0. \end{cases}$$

13.4.4 Satz: Formeln für den Konvergenzradius

Es sei $\sum_{k=0}^{\infty} p_k(z-a)^k$ eine Potenzreihe mit Konvergenzradius ϱ.

(i) Es gilt die Formel von Cauchy–Hadamard

$$\varrho \quad = \quad \frac{1}{\limsup\limits_{k\to\infty} \sqrt[k]{|p_k|}}$$

(ii) Es gelten die Abschätzungen

$$\liminf_{k\to\infty} \left|\frac{p_k}{p_{k+1}}\right| \quad \leq \quad \varrho \quad \leq \quad \limsup_{k\to\infty} \left|\frac{p_k}{p_{k+1}}\right|$$

Dabei sind die Rechenregeln $\frac{1}{0} = \infty$, $\frac{1}{\infty} = 0$ anwendbar.

(iii) Existiert der (eigentliche oder uneigentliche) Grenzwert auf der rechten Seite, so gilt

$$\varrho = \lim_{k\to\infty} \left|\frac{p_k}{p_{k+1}}\right|$$

13.4.5 Beweis Die beiden Aussagen sind Übertragungen von Wurzel– bzw. Quotientenkriterium auf Potenzreihen. Wir beweisen nur (i), können dabei wieder $a = 0$ annehmen.

(1) Wir bezeichnen den Cauchy–Hadamard–Ausdruck mit

$$R \;:=\; \frac{1}{\limsup\limits_{k\to\infty} \sqrt[k]{|p_k|}} \;=\; \liminf_{k\to\infty} \underbrace{\frac{1}{\sqrt[k]{|p_k|}}}_{=:\alpha_k}.$$

Gemäß der Charakterisierung (ii) des Limes inferior in Satz 13.4.2 gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass

$$\frac{1}{\sqrt[k]{|p_k|}} \;\geq\; R - \varepsilon \quad \text{für alle } k \geq N, \qquad (*)$$

$$\frac{1}{\sqrt[k]{|p_k|}} \;\leq\; R + \varepsilon \quad \text{für unendlich viele } k \geq N. \qquad (**)$$

(2) Es sei jetzt (O.B.d.A.) $R > 0$ und $y \in \mathbb{R}$ beliebig, so dass $0 < |y| < R$. Dann gibt es ein $\varepsilon > 0$, so dass $|y| < R - \varepsilon$. Aufgrund von $(*)$ gibt es zu diesem ε ein $N \in \mathbb{N}$, so dass

$$\sqrt[k]{|p_k \cdot y^k|} = \sqrt[k]{|p_k|} \cdot |y| \leq \frac{|y|}{R-\varepsilon} = q < 1 \quad \text{für alle } k \geq N.$$

Gemäß Wurzel–Kriterium 6.3.9 konvergiert die Reihe $\sum_{k=0}^{\infty} p_k y^k$ absolut. Da y eine beliebige Zahl mit $|y| < R$ war, können wir $R \leq \varrho$ folgern.

(3) Wir betrachten jetzt ein $y \in \mathbb{R}$ mit $|y| > R$ und setzen $\varepsilon = |y| - R$. Aufgrund von $(**)$ gibt es $N \in \mathbb{N}$, so dass

$$\frac{1}{\sqrt[k]{|p_k|}} \leq R + \varepsilon = |y| \quad \text{für unendlich viele } k \geq N,$$

woraus folgt, dass

$$1 \leq |p_k y^k| \qquad \text{für unendlich viele } k \geq N.$$

Das aber bedeutet, dass die Reihe $\sum_{k=0}^{\infty} p_k y^k$ nicht konvergieren kann, es folgt:

$$R \geq \varrho.$$

(4) Die zweite Formel (ii) wird hier nicht bewiesen.

13.4.6 Definition: Verallgemeinerter Binomialkoeffizient

Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}_0$ definieren wir zunächst die verallgemeinerten Binomialkoeffizienten

$$\binom{\alpha}{k} := \begin{cases} \prod_{\ell=1}^{k} \frac{\alpha-\ell+1}{\ell} = \frac{\alpha\cdot(\alpha-1)\cdot(\alpha-2)\cdot\ldots\cdot(\alpha-k+1)}{1\cdot 2\cdot\ldots\cdot k}, & \text{falls } k \geq 1, \\ 1, & \text{falls } k = 0. \end{cases}$$

Man kann sich ganz leicht davon überzeugen, dass für $\alpha \in \mathbb{N}_0$ diese Definition mit der ursprünglichen in 2.10.9 gegebenen übereinstimmt.

13.4.7 Beispiel: Binomialreihe

Wir betrachten die Binomialreihe (= binomische Reihe)

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} z^k$$

mit Entwicklungsstelle 0. Ihr Konvergenzradius ist wegen

$$\left|\frac{p_k}{p_{k+1}}\right| = \left|\frac{k+1}{\alpha-k}\right| = \left|\frac{k+1}{k-\alpha}\right| \longrightarrow 1$$

gleich 1, vgl. Satz 13.4.4 (iii).

13.5 Der Abel'sche Grenzwertsatz

13.5.1 Satz: Abel'scher Grenzwertsatz

Es sei $\sum_{k=0}^{\infty} p_k x^k$ eine **reelle** Potenzreihe mit Konvergenzradius ϱ, so dass sie eine stetige Funktion $f :]-\varrho, +\varrho[\to \mathbb{R}$ darstellt.

(i) Konvergiert die Potenzreihe in dem Randpunkt $x = -\varrho$ des Konvergenzintervalls, so existiert

$$\lim_{x \searrow -\varrho} f(x) = \sum_{k=0}^{\infty} p_k (-\varrho)^k =: f(-\varrho).$$

(ii) Konvergiert die Potenzreihe in dem Randpunkt $x = \varrho$ des Konvergenzintervalls, so existiert

$$\lim_{x \nearrow +\varrho} f(x) = \sum_{k=0}^{\infty} p_k \varrho^k =: f(\varrho).$$

13.5.2 Beweis

(1) Die Aussage braucht nur für $\varrho = 1$ und $x = +1$ gezeigt zu werden.

(2) Mit $r_{n+1} := \sum_{\ell=n+1}^{\infty} p_\ell$ gilt

$$\begin{aligned}
|f(x) - f(1)| &= \left| \underbrace{(1-x) \cdot \sum_{k=0}^{\infty} x^k}_{=1} \cdot \left(\sum_{k=0}^{\infty} p_k x^k - \sum_{k=0}^{\infty} p_k\right) \right| \qquad \text{(Cauchy-Produkt)} \\
&= \left| (1-x) \left[\sum_{k=0}^{\infty} \sum_{\ell=0}^{k} p_\ell x^\ell x^{k-\ell} - \sum_{k=0}^{\infty} x^k \sum_{\ell=0}^{\infty} p_\ell \right] \right| \\
&= \left| (1-x) \left[\sum_{k=0}^{\infty} \sum_{\ell=0}^{k} p_\ell x^k - \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} p_\ell x^k \right] \right| \\
&= \left| (1-x) \left[\sum_{k=0}^{\infty} r_{k+1} \cdot x^k \right] \right| \\
&\le |1-x| \cdot \sum_{k=0}^{\infty} |r_{k+1}| \cdot |x|^k
\end{aligned}$$

(3) Wegen $\lim_{k\to\infty} r_{k+1} = 0$ gibt es $N \in \mathbb{N}$, so dass $|r_{k+1}| < \frac{\varepsilon}{2}$ für $k \ge N$, dann für $|x| < 1$

$$\begin{aligned}
&= |1-x| \cdot \left[\sum_{k=0}^{N} |r_{k+1}| \cdot |x|^k + \sum_{k=N+1}^{\infty} |r_{k+1}| \cdot |x|^k \right] \\
&\le |1-x| \cdot \left[\sum_{k=0}^{N} |r_{k+1}| + \frac{\varepsilon}{2} \sum_{k=N+1}^{\infty} |x|^k \right]
\end{aligned}$$

$$
\begin{aligned}
&= |1-x| \cdot \left[\sum_{k=0}^{N} |r_{k+1}| + \frac{\varepsilon}{2} \, \frac{|x|^{N+1}}{1-|x|} \right] \\
&\le |1-x| \cdot \sum_{k=0}^{N} |r_{k+1}| + \frac{\varepsilon}{2} \, \frac{|1-x|}{1-|x|} \\
&= |1-x| \cdot \sum_{k=0}^{N} |r_{k+1}| + \frac{\varepsilon}{2}.
\end{aligned}
$$

(4) Insgesamt haben wir bis jetzt zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gefunden, so dass

$$
|f(x) - f(1)| \le |1-x| \cdot \sum_{k=0}^{N} |r_{k+1}| + \frac{\varepsilon}{2}
$$

Beim Grenzübergang $\lim_{x \nearrow 1}$ geht diese Ungleichung über in

$$
\lim_{x \nearrow 1} |f(x) - f(1)| \le \frac{\varepsilon}{2}.
$$

Da dies für alle $\varepsilon > 0$ gilt, muss

$$
\lim_{x \nearrow 1} |f(x) - f(1)| = 0
$$

sein.

13.5.3 Die Potenzreihe des natürliche Logarithmus

Wir betrachten die geometrische Reihe in der Form

$$\sum_{k=0}^{\infty}(-1)^k x^k = \sum_{k=0}^{\infty}(-x)^k = \frac{1}{1-(-x)} = \frac{1}{1+x}, \qquad |x|<1$$

und bilden gemäß Satz 13.3.6 (iii) auf dem Intervall $]-1,+1[$ die Stammfunktion mit Wert 0 bei $x=0$:

$$\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k = \sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}x^{k+1} = \int_0^x \frac{1}{1+y}\,dy \;=\; \ln(x+1), \qquad |x|<1.$$

Damit haben wir eine Potenzreihendarstellung für die Logarithmus–Funktion gefunden.

Da die Reihe gemäß Leibniz–Kriterium auch für $x=1$ konvergiert, folgt mit dem Abel'schen Grenzwertsatz

$$\ln 2 = \lim_{x\nearrow 1}\ln(x+1) = \lim_{x\nearrow 1}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k \;=\; \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}.$$

Eine Transformation $x+1 \rightsquigarrow x$ ergibt

$$\ln x = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}(x-1)^k, \qquad x\in\,]0,2].$$

Ist $a>0$, so kann man mit Hilfe der Funktionalgleichung des Logarithmus schreiben:

$$\begin{aligned}\ln x = \ln a + \ln(\frac{x}{a}) \;&=\; \ln a + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}(\frac{x}{a}-1)^k \\ &=\; \ln a + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{ka^k}(x-a)^k, \qquad x\in\,]0,2a].\end{aligned}$$

So haben wir eine Potenzreihendarstellung mit Entwicklungsstelle a und Konvergenzradius a ermittelt.

13.5.4 Potenzreihendarstellung für den Arcustangens

In ähnlicher Weise kann man eine Potenzreihendarstellung für den arctan über die Ableitung

$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{k=0}^{\infty}(-x^2)^k, \qquad |x|<1,$$

ermitteln.

13.5.5 Satz: Binomialreihe Für die in Abschnitt 13.4.7 für beliebiges $\alpha \in \mathbb{R}$ definierte Binomialreihe gilt

$$(x+1)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \qquad \text{falls} \quad |x| < 1.$$

13.5.6 Beweis

(1) Man rechnet sofort nach, dass für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}_0$ gilt:

$$(k+1)\binom{\alpha}{k+1} = (\alpha - k)\binom{\alpha}{k}$$

(2) Die im Satz angegebene Reihe hat den Konvergenzradius 1, da

$$\left| \frac{\binom{\alpha}{k}}{\binom{\alpha}{k+1}} \right| = \left| \frac{k+1}{\alpha - k} \right| \stackrel{k\to\infty}{\longrightarrow} 1.$$

(3) Wir bilden die Hilfsfunktion

$$h : \left\{ \begin{array}{rcl}]-1,+1[& \to & \mathbb{R} \\ x & \mapsto & (1+x)^{-\alpha} \cdot \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \end{array} \right.$$

und zeigen, dass sie die Konstant–1–Funktion ist.

(4) Zunächst gilt $h(0) = 1$, dann erhalten wir mit der Leibniz–Produktregel und (erlaubter) gliedweiser Differentiation

$$\begin{aligned}
h'(x) &= (-\alpha)(1+x)^{-\alpha-1} \cdot \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k + (1+x)^{-\alpha} \cdot \sum_{k=1}^{\infty} \binom{\alpha}{k} k x^{k-1} \\
&= (1+x)^{-\alpha-1} \cdot \left[\sum_{k=0}^{\infty} (-\alpha) \binom{\alpha}{k} x^k + \sum_{k=1}^{\infty} \binom{\alpha}{k} k x^{k-1} + \sum_{k=1}^{\infty} \binom{\alpha}{k} k x^k \right] \\
&= (1+x)^{-\alpha-1} \cdot \left[\sum_{k=0}^{\infty} (k-\alpha) \binom{\alpha}{k} x^k + \sum_{k=0}^{\infty} \binom{\alpha}{k+1} (k+1) x^k \right] \\
&= (1+x)^{-\alpha-1} \cdot \sum_{k=0}^{\infty} \underbrace{\left[(k-\alpha)\binom{\alpha}{k} + \binom{\alpha}{k+1}(k+1) \right]}_{\stackrel{(1)}{=} 0} x^k = 0.
\end{aligned}$$

14 Taylor–Approximation, Analytische Funktionen

14.1 Taylor–Approximation

14.1.1 Definition: Es sei $a \in J \subseteq \mathbb{R}$ eine Stelle im (echten) Intervall J.

(1) Existiert für eine Funktion $g : J \to \mathbb{R}$ die n–te Ableitung an der Stelle a, so kann man ihr das *Taylorpolynom n–ten Grades mit Entwicklungsstelle a*

$$\sum_{k=0}^{n} g^{(k)}(a)\,\frac{(x-a)^k}{k!} \quad = \quad g(a) + g'(a)\,\frac{(x-a)}{1} + g''(a)\,\frac{(x-a)^2}{2} + \ldots + g^{(n)}(a)\,\frac{(x-a)^n}{n!}$$

zuordnen.

(2) Die Differenz $r_{n+1}(x)$ zwischen Funktion und zugeordnetem Taylorpolynom n–ten Grades wird als *(Taylor–)Restglied $(n+1)$–ter Ordnung* bezeichnet:

$$g(x) = \sum_{k=0}^{n} g^{(k)}(a)\,\frac{(x-a)^k}{k!} + r_{n+1}(x).$$

Vorsicht: Manchmal (Blatter, Erwe) ist die Indizierung dieses Restglieds mit n statt $n+1$ üblich.

Es gilt offenbar:

$$r_n(x) - r_{n+1}(x) = g^{(n)}(a)\,\frac{(x-a)^n}{n!}.$$

14.1.2 Beispiel Ist g ein Polynom m–ten Grades

$$g(x) = p_0 + p_1 x + p_2 x^2 + \ldots + p_m x^m,$$

so gilt mit Satz 13.3.6 (iv)

$$g^{(k)}(0) \quad = \quad \begin{cases} k!\,p_k, & \text{falls } k \le m \\ 0, & \text{sonst.} \end{cases}$$

Deshalb hat für $n \le m$ das Taylor–Polynom n–ten Grades die Form

$$p_0 + p_1 x + p_2 x^2 + \ldots + p_n x^n,$$

insbesondere stimmt für $n = m$ das Taylor-Polynom mit dem ursprünglichen Polynom überein.

Dies führt ganz allgemein auf die Idee eine gegebene n-fach differenzierbare Funktion mit ihrem Taylor-Polynom zu vergleichen.

14.1.3 Satz: Taylor–Formel

Es sei $g : J \to \mathbb{R}$ eine $\mathcal{C}^{n+1}$–Funktion und $a \in J$ eine Entwicklungsstelle.

(i) Für das Restglied r_{n+1} gilt:

$$r_{n+1}(x) = \int_a^x g^{(n+1)}(t)\,\frac{(x-t)^n}{n!}\,dt.$$

(ii) Zu jedem $x \in J$ existiert ein ξ zwischen a und x, so dass

$$r_{n+1}(x) = g^{(n+1)}(\xi)\,\frac{(x-a)^{n+1}}{(n+1)!}.$$

Man nennt diesen Ausdruck das *Lagrange–Restglied*.

14.1.4 Beweis

(i) Für $n = 0$ ist dies einfach der HDI. Der Induktionsschritt wird im wesentlichen mit Hilfe von Partieller Integration Satz 11.3.1 ausgeführt:

$$\begin{aligned}
r_{n+2}(x) &= r_{n+1}(x) - g^{(n+1)}(a) \cdot \frac{(x-a)^{n+1}}{(n+1)!} \\
&\overset{\text{IndV}}{=} \int_a^x \underbrace{g^{(n+1)}(t)}_{F(t)} \underbrace{\frac{(x-t)^n}{n!}}_{G'(t)} \, dt - \Big[\underbrace{g^{(n+1)}(t)}_{F(t)} \cdot \underbrace{\frac{-(x-t)^{n+1}}{(n+1)!}}_{G(t)} \Big]_{t=a}^{t=x} \\
&\overset{\text{Part Int}}{=} - \int_a^x \underbrace{g^{(n+2)}(t)}_{F'(t)} \underbrace{\frac{-(x-t)^{n+1}}{(n+1)!}}_{G(t)} \, dt.
\end{aligned}$$

Ist also der Ausdruck für das $(n+1)$–te Restglied richtig, so ist er auch für das $(n+2)$–te richtig.

(ii) Gemäß Mittelwertsatz 10.7.1(iii) der Integralrechnung existiert zwischen a und x ein ξ, so dass

$$\begin{aligned}
r_{n+1}(x) &= \int_a^x g^{(n+1)}(t) \, \frac{(x-t)^n}{n!} \, dt = g^{(n+1)}(\xi) \int_a^x \frac{(x-t)^n}{n!} \, dt \\
&= g^{(n+1)}(\xi) \left[\frac{-(x-t)^{n+1}}{(n+1)!} \right]_{t=a}^{t=x} = g^{(n+1)}(\xi) \, \frac{(x-a)^{n+1}}{(n+1)!}.
\end{aligned}$$

Beachte dabei, dass die Funktion $t \mapsto \frac{(x-t)^n}{n!}$ konstantes Vorzeichen im Integrationsintervall hat.

14.1.5 Satz: Taylor Approximationssatz

Es sei $g : J \to \mathbb{R}$ eine $\mathcal{C}^n$–Funktion und $a \in J$ eine Entwicklungsstelle.

Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\underbrace{\Big| g(x) - \sum_{k=0}^{n} g^{(k)}(a) \, \frac{(x-a)^k}{k!} \Big|}_{=r_{n+1}(x)} \le \varepsilon \, |x-a|^n, \qquad \text{falls } |x-a| \le \delta.$$

14.1.6 Beweis Wir verwenden das Lagrange–Restglied. Es ist dann für ein geeignetes ξ zwischen a und x

$$|r_{n+1}(x)| = |r_n(x) - g^{(n)}(a) \, \frac{(x-a)^n}{n!}| = |g^{(n)}(\xi) - g^{(n)}(a)| \cdot \Big| \frac{(x-a)^n}{n!} \Big|.$$

Wegen der Stetigkeit von $g^{(n)}$ kann man den ersten Faktor rechts kleiner als ein vorgegebenes $\varepsilon > 0$ machen, wenn nur $|x-a|$, und damit $|\xi - a|$, kleiner als ein geeignetes $\delta > 0$ gemacht wird.

14.2 Ausblick: Analytische Funktionen (Nicht klausurrelevant)

14.2.1 Definitionen (und Bezeichnungen) Es sei a eine Stelle in $\mathbb{R}$.

(1) Für $k \in \mathbb{N}_0$ sei $\mathcal{C}_a^k$ die Menge der reellen Funktionen g, die auf einem ,,Intervall" $]a-r, a+r[$ mit $r > 0$ definiert und dort k–mal differenzierbar sind.

(2) Es sei weiter

$$\mathcal{C}_a^\infty \quad := \quad \bigcap_{k \in \mathbb{N}} \mathcal{C}_a^k$$

die Menge aller Funktionen, die ,,an der Stelle a unendlich oft differenzierbar" sind.

Beachte, dass eine Funktion g in $\mathcal{C}_a^\infty$ enthalten sein kann, ohne dass ein Intervall $]a-r, a+r[$ existieren muss, auf dem g unendlich differenzierbar sein muss. Es kann passieren, dass die Intervalle $]a-r_k, a+r_k[$, auf denen die Ableitungen $g^{(k)}$ existieren, nur die Stelle a gemeinsam haben.

(3) $\mathcal{P}_a^{\text{formal}}$ sei die Menge der (formalen) Potenzreihen mit Entwicklungsstelle a. Genau genommen handelt es sich dabei um die Menge $\{(p_k)_{k \in \mathbb{N}_0}\}$ aller Koeffizientenfolgen.

(4) $\mathcal{P}_a^{\text{konv}}$ sei die Menge der Potenzreihen mit Entwicklungsstelle a, die einen positiven Konvergenzradius ϱ haben, also in einem Intervall $]a-\varrho, a+\varrho[$ konvergieren.

Zu jeder Funktion in $g \in \mathcal{C}_a^\infty$ können wir die *Taylorreihe*

$$\sum_{k=0}^{\infty} g^{(k)}(a)\, \frac{(x-a)^k}{k!}$$

bilden. Dies ist eine formale Potenzreihe mit Entwicklungsstelle a. Über ihre Konvergenz wissen wir a priori NIX.

14.2.2 Operatoren-Sichtweise

Wir kennen inzwischen zwei ,,Operatoren", die eine Beziehung zwischen den in a ∞–oft differenzierbaren Funktionen und Potenzreihen mit Entwicklungsstelle a aufbauen.

$$\mathcal{C}_a^\infty \xrightarrow{\text{Taylorreihe}} \mathcal{P}_a^{\text{formal}}$$

$$\mathcal{C}_a^\infty \xleftarrow[\text{Grenzfunktion}]{} \mathcal{P}_a^{\text{konv}}$$

Dazu einige Erläuterungen:

- Der Operator ,,Taylorreihe" der oberen Zeile des Diagramms ist nicht injektiv. Dies zeigt das Beispiel der Funktion

$$\eta(x) \;:=\; \begin{cases} \exp(-\frac{1}{x}), & \text{falls } x > 0, \\ 0, & \text{falls } x \le 0. \end{cases}$$

aus dem Beweis von Satz 9.11.6. Sie ist in $a = 0$ ∞–oft differenzierbar mit

$$\eta^{(k)}(0) = 0 \qquad \text{für alle } k \in \mathbb{N}_0.$$

Insbesondere ist ihre Taylorreihe mit Entwicklungsstelle 0 gleich Null. Die Funktion η besitzt also die gleiche Taylorreihe wie die Nullfunktion, obwohl sie in jeder (beliebig kleinen) ε–Umgebung von 0 nicht mit der Nullfunktion übereinstimmt.

- Ein Satz von Emile Borel besagt, dass es zu jeder (formalen) Potenzreihe eine ∞–oft differenzierbare Funktion gibt, die diese Potenzreihe als Taylorreihe besitzt. Mit anderen Worten, der Taylorreihe–Operator ist surjektiv. Das heißt insbesondere, dass die Taylorreihe einer $\mathcal{C}^\infty$–Funktion den Konvergenzradius $\varrho = 0$ haben kann, also nur in der Entwicklungsstelle konvergiert. Beim Beweis des Satzes von Borel muss zu jeder gegebenen Potenzreihe eine passende $\mathcal{C}^\infty$ Funktion konstruiert werden. Dabei wird gerade die Funktion η von weiter oben benutzt.

14.2.3 Satz und Definition: analytische Funktionen

Es sei $g : D \to \mathbb{R}$ eine Funktion, $a \in D$ sei eine innere Stelle von D.

Die Funktion heißt *analytisch in a*, wenn eine der folgenden äquivalenten Eigenschaften erfüllt ist:

(A) Es gibt eine Potenzreihe mit Entwicklungsstelle a und positivem Konvergenzradius, so dass die Grenzfunktion f auf einer ε–Umgebung von a mit g übereinstimmt.

(B) Die Funktion g erfüllt die drei Bedingungen

- g ist in a unendlich oft differenzierbar
- die zugehörige Taylorreihe hat positiven Konvergenzradius
- die Grenzfunktion f der Taylorreihe stimmt in einer ε–Umgebung mit g überein.

(C) Die Funktion g ist in a unendlich oft differenzierbar, die Folge der Restglieder

$$r_{n+1}(x) \quad = \quad g(x) - \sum_{k=0}^{n} g^{(k)}(a)\, \frac{(x-a)^k}{k!}$$

konvergiert in einer ε–Umgebung von a punktweise gegen die Nullfunktion.

14.2.4 Beachte nochmals, dass in der Aussage (B) die erste Bedingung nicht die zweite und die zweite nicht die dritte impliziert.

14.2.5 Diagramm für analytische Funktionen Bezeichnet man nun die Menge der in a analytischen Funktionen mit $\mathcal{C}_a^\omega$, so können wir das Diagramm in 14.2.2 so abändern:

$$\begin{array}{ccc} \mathcal{C}_a^\omega & \xrightarrow{\text{Taylorreihe}} & \mathcal{P}_a^{\text{konv}} \\ \mathcal{C}_a^\infty & \xleftarrow[\text{Grenzfunktion}]{} & \mathcal{P}_a^{\text{konv}} \end{array}$$

Die beiden Operatoren sind jetzt invers zueinander, wenn man Mehrdeutigkeiten bezüglich der Definitionsmenge der analytischen Funktionen um a herum — durch Bildung von Äquivalenzklassen, die hier Funktionskeime heißen —, beseitigt.

14.2.6 Satz: Grenzfunktion einer Potenzreihe ist analytisch

Ist

$$\sum_{k=0}^{\infty} p_k(x-a)^k$$

eine Potenzreihe mit positivem Konvergenzradius ϱ, so ist die Grenzfunktion f analytisch auf $]a-\varrho, a+\varrho[$.

Es folgt, dass eine Funktion g, die in einem $a \in D$ analytisch ist, immer auch in einer ε–Umgebung von a analytisch ist.

14.2.7 Beweis Man muss beweisen, dass eine Potenzreihe mit Entwicklungsstelle a auch bzgl. eines anderen $b \in]a-\varrho, a+\varrho[$ als Potenzreihe entwickelbar ist. Das ist ein bisschen aufwändiger als es scheint und wird hier weggelassen.

15 Topologie metrischer Räume

15.1 Der euklidische Raum $\mathbb{R}^d$

Es sei

$$\mathbb{R}^d := \underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{d\text{–mal}} = \{(x_1, x_2, \ldots, x_d) | x_i \in \mathbb{R} \text{ für alle } i = 1, \ldots, d\}.$$

In der Linearen Algebra steht im Vordergrund, dass man die Elemente des $\mathbb{R}^d$ addieren und skalar multiplizieren kann: Der $\mathbb{R}^d$ ist ein (reeller) Vektorraum. Der Nullvektor $\vec{0} = (0, \ldots, 0)$ heißt auch *Ursprung*.

Oft — je nach Kontext — werden für die Elemente des $\mathbb{R}^d$ auch die Schreibweisen

$$x = (x_1, x_2, \ldots, x_d) \qquad \text{oder} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

benutzt.

Für $d = 2$ wird der $\mathbb{R}^2$ oft als Zeichenebene, für $d = 3$ als der uns umgebende 3–dimensionale Raum veranschaulicht.

Im Zusammenhang mit einer Folge $(a_n)_{n \in \mathbb{N}}$ in $\mathbb{R}^d$ treten zwei Indizierungen auf: Der Koordinatenindex und der Folgenindex. Wir vereinbaren, dass $a_{i|n}$ die i-te Koordinate des n-ten Folgengliedes bedeuten soll. Beachte, dass dann

- für festes $i \in \{1, \ldots, d\}$ $\quad (a_{i|n})_n$ die aus den i-ten Koordinaten gebildete Zahlenfolge ist,
- für festes $n \in \mathbb{N}$ $\quad a_n = a_{|n}$ der Vektor an der n-ten Stelle der Folge ist.

15.1.1 Definition: Euklidische Norm

(1) Die *euklidische Norm* im $\mathbb{R}^d$ ist definiert als die Abbildung

$$\|\cdot\| : \begin{cases} \mathbb{R}^d & \to & \mathbb{R}_0^+ \\ x & \mapsto & \|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_d^2}. \end{cases}$$

(2) Wird bei mathematischen Überlegungen über den $\mathbb{R}^d$ die euklidische Norm benutzt, so spricht man auch vom *euklidischen Vektorraum* $\mathbb{R}^d$.

15.1.2 Bemerkung „Identifiziert" man die komplexe Ebene durch die bijektive Zuordnung

$$\begin{cases} \mathbb{C} & \to & \mathbb{R}^2 \\ z & \mapsto & (\operatorname{Re} z, \operatorname{Im} z) \end{cases}$$

mit der euklidischen Ebene, so stimmen die Definitionen des Betrags einer komplexen Zahl und der Norm eines $\mathbb{R}^2$–Vektors überein.

15.1.3 Definitionen: Norm

(1) Ist V ein $\mathbb{K}$-Vektorraum, so heißt eine Abbildung

$$\|\cdot\| : \begin{cases} V & \to & \mathbb{R}_0^+ \\ x & \mapsto & \|x\|. \end{cases}$$

eine *Norm*, wenn für $x, \widetilde{x} \in V$ und $\alpha \in \mathbb{K}$ die folgenden drei Eigenschaften erfüllt sind:

(N1) $\|x\| = 0 \quad \Longrightarrow \quad x = 0$

(N2) $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$

(N3) $\|x + \widetilde{x}\| \leq \|x\| + \|\widetilde{x}\|$ (Dreiecksungleichung)

Die dritte Eigenschaft impliziert außerdem

$$\Big| \|x\| - \|\widetilde{x}\| \Big| \leq \|x - \widetilde{x}\|.$$

(2) Ein Vektorraum zusammen mit einer Norm heißt *normierter Raum*.

(3) Durch

$$d(x, \widetilde{x}) := \|x - \widetilde{x}\|$$

induziert eine Norm eine Metrik. Die Dreiecksungleichung zeigen wir auf:

$$d(x, \widetilde{x}) = \|x - \widetilde{x}\| = \|(x - \widetilde{\widetilde{x}}) + (\widetilde{\widetilde{x}} - \widetilde{x})\| \leq \|x - \widetilde{\widetilde{x}}\| + \|\widetilde{\widetilde{x}} - \widetilde{x}\| = d(x, \widetilde{\widetilde{x}}) + d(\widetilde{\widetilde{x}}, \widetilde{x}).$$

15.1.4 Satz: Eigenschaften der euklidischen Norm

(i) Die euklidische Norm ist eine Norm.

(ii) Der euklidische Raum $\mathbb{R}^d$ wird durch die Metrik

$$d_{\text{eu}}(x, \widetilde{x}) := \|x - \widetilde{x}\|$$

ein vollständiger metrischer Raum.

15.1.5 Beweis

(i) Die Eigenschaften (N1) und (N2) sind einfach nachzurechnen. Die Dreiecksungleichung (N3) wird in der Linearen Algebra mit Hilfe des Begriffs des Skalarprodukts und der ,,Cauchy–Schwarz–Ungleichung" bewiesen.

(ii) Dazu sei (a_n) eine Cauchy–Folge in $\mathbb{R}^d$. Für festes $i \in \{1, \ldots, d\}$ ist dann wegen

$$|a_{i|n} - a_{i|m}| \leq \|a_n - a_m\|$$

auch die Zahlenfolge $(a_{i|n})_n$ eine Cauchy–Folge in $\mathbb{R}$. Aufgrund der Vollständigkeit von $\mathbb{R}$ konvergiert sie gegen eine Zahl b_i. Wir bilden den Vektor $b = (b_1, \ldots, b_d)$. Aufgrund von

$$\|a_n - b\| = \sqrt{(a_{1|n} - b_1)^2 + (a_{2|n} - b_2)^2 + \ldots + (a_{d|n} - b_d)^2}$$

gilt dann auch $\lim_{n\to\infty} a_n = b$.

15.1.6 Definition: Schur-Norm einer Matrix

Lineare Abbildungen $\mathbb{K}^d \to \mathbb{K}^w$ werden bzgl. der kanonischen Basen beschrieben durch $w \times d$–Matrizen mit w Zeilen und d Spalten.

Die Menge der $w \times d$–Matrizen bildet einen Vektorraum $\mathbb{K}^{w\times d} \cong \mathbb{K}^{wd}$, auf dem wir auch die obige Norm erklären können. Für eine $w \times d$–Matrix A sei die *Schur–Norm* definiert durch

$$\|A\| := \sqrt{\sum_{j=1}^{w}\sum_{k=1}^{d}(A_{jk})^2}.$$

Es gilt dann

15.1.7 Satz: Submultiplikativität Für zwei Matrizen $A \in \mathbb{K}^{v\times w}$ und $B \in \mathbb{K}^{w\times d}$ in dem Diagramm gilt

$$\mathbb{K}^d \xrightarrow{B} \mathbb{K}^w \xrightarrow{A} \mathbb{K}^v.$$

gilt:

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

Man nennt diese Eigenschaft *Submultiplikativität*.

15.1.8 Beweis Es sei $j \in \{1,\dots,v\}$ die Nummer einer Zeile von A und es sei $k \in \{1,\dots,d\}$ die Nummer einer Spalte von B.

Für das Produkt dieser beiden Vektoren gilt aufgrund der Cauchy–Schwarz–Ungleichung

$$(\sum_{\ell=1}^{w} A_{j\ell}B_{\ell k})^2 \leq \sum_{\ell=1}^{w}(A_{j\ell})^2 \cdot \sum_{\ell=1}^{w}(B_{\ell k})^2$$

Wir addieren über alle Zeilen j von A

$$\sum_{j=1}^{v}(\sum_{\ell=1}^{w} A_{j\ell}B_{\ell k})^2 \leq \sum_{j=1}^{v}\sum_{\ell=1}^{w}(A_{j\ell})^2 \cdot \sum_{\ell=1}^{w}(B_{\ell k})^2$$

und dann über alle Spalten k von B

$$\sum_{k=1}^{d}\sum_{j=1}^{v}(\sum_{\ell=1}^{w} A_{j\ell}B_{\ell k})^2 \leq \sum_{j=1}^{v}\sum_{\ell=1}^{w}(A_{j\ell})^2 \cdot \sum_{k=1}^{d}\sum_{\ell=1}^{w}(B_{\ell k})^2,$$

das aber ist gerade

$$\|A \cdot B\|^2 \leq \|A\|^2 \cdot \|B\|^2.$$

15.2 Weitere Begriffe für metrische Räume

In der Bemerkung 5.1.3 ist bereits darauf hingewiesen worden, dass Begriffsbildungen wie

- Cauchy–Folge
- Konvergenz einer Folge
- Vollständigkeit
- Stetigkeit einer Funktion

allein auf dem Begriff der Metrik basieren. Lediglich bei der Schreibweise $|x - \widetilde{x}|$ für den Abstand $d(x, \widetilde{x})$ hatten wir Bezug auf die konkreten Beispiele $\mathbb{R}$ oder $\mathbb{C}$ genommen. Innerhalb dieses Kapitels 15 wollen wir die allgemeine Symbolik $d(x, \widetilde{x})$ verwenden.

15.2.1 Diskrete Metrik Ein interessantes Beispiel einer Metrik auf einer beliebigen Menge M ist

$$d_{\text{dis}}(x, \widetilde{x}) := \begin{cases} 1, & \text{falls } x \neq \widetilde{x}, \\ 0, & \text{falls } x = \widetilde{x}. \end{cases}$$

Man überzeugt sich leicht davon, dass (M, d_{dis}) ein metrischer Raum ist. d heißt die *diskrete Metrik* auf M.

15.2.2 Definition: Es sei (M, d) ein metrischer Raum und $a \in M$.

(1) Für $\varepsilon > 0$ heißen die Mengen

$$U_\varepsilon(a) := \left\{ x \in M \middle| d(x, a) < \varepsilon \right\}$$

$$B_\varepsilon(a) := \left\{ x \in M \middle| d(x, a) \leq \varepsilon \right\}$$

offene bzw. *abgeschlossene ε–Umgebung* von a.

(2) Eine Teilmenge $U \subseteq M$ heißt *Umgebung von a*, wenn es ein $\varepsilon > 0$ gibt, so dass

$$a \in U_\varepsilon(a) \subseteq U.$$

Offene und abgeschlossene ε–Umgebungen sind auch Umgebungen, sprachliche und mathematische Logik stimmen also überein.

(3) Für eine beliebige Teilmenge X eines metrischen Raumes heißt die Zahl

$$\operatorname{diam} X := \sup \left\{ d(x, \widetilde{x}) \,\middle|\, x, \widetilde{x} \in X \right\} \in [0, \infty]$$

Durchmesser (diameter) dieser Teilmenge.

(4) Eine Teilmenge $X \subseteq M$ heißt *beschränkt*, wenn ihr Durchmesser endlich ist. Wegen der Dreiecksungleichung ist dies gleichbedeutend damit, dass für festes $a \in M$ die Teilmenge

$$\left\{ d(x, a) \,\middle|\, x \in X \right\}$$

von $\mathbb{R}_0^+$ beschränkt ist.

(5) Eine Funktion $f : X \to M$ zwischen einer beliebigen Menge X und einem metrischen Raum M heißt *beschränkt*, wenn ihre Bildmenge $f(X)$ in M beschränkt ist.

Für Folgen und reellwertige Funktionen erkennen wir darin die bereits benutzte Definition von Beschränktheit wieder.

15.3 Weitere Beispiele für metrische Räume

15.3.1 Die erweiterte reelle Zahlengerade Wir *adjungieren* an $\mathbb{R}$ zwei Punkte $-\infty$ und $+\infty$, erhalten so die Menge

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}.$$

Auf $\overline{\mathbb{R}}$ ist eine lineare Ordnung erklärt durch

$$x < y \qquad \Longleftrightarrow \qquad \begin{cases} x < y, & \text{falls } x, y \in \mathbb{R}, \\ x < +\infty, & \text{falls } x \in \mathbb{R} \text{ oder } x = -\infty, \\ -\infty < y, & \text{falls } y \in \mathbb{R} \text{ oder } y = +\infty. \end{cases}$$

Man spricht dann von der *erweiterten reellen Zahlengerade*. Man beachte, dass die auf $\mathbb{R}$ gegebenen Körperstrukturen nicht auf $\overline{\mathbb{R}}$ fortgesetzt werden können.

15.3.2 Metrik auf $\overline{\mathbb{R}}$ Wir definieren weiter die zueinander inversen streng monoton steigenden Funktionen

$$h : \begin{cases} \overline{\mathbb{R}} & \to & [-1, +1] \\ x & \mapsto & \begin{cases} \frac{x}{1+|x|}, & \text{falls } x \in \mathbb{R}, \\ -1, & \text{falls } x = -\infty, \\ +1, & \text{falls } x = +\infty, \end{cases} \end{cases}$$

$$h^{-1} : \begin{cases} [-1, +1] & \to & \overline{\mathbb{R}} \\ y & \mapsto & \begin{cases} \frac{y}{1-|y|}, & \text{falls } |y| < 1, \\ -\infty, & \text{falls } y = -1, \\ +\infty, & \text{falls } y = +1. \end{cases} \end{cases}$$

Dann ist, wie man überprüfen kann,

$$\overline{d}(x_1, x_2) := |h(x_1) - h(x_2)|$$

eine Metrik auf $\overline{\mathbb{R}}$.

15.3.3 Satz

(i) Die Funktionen h und h^{-1} sind stetige Abbildungen zwischen den metrischen Räumen $(\overline{\mathbb{R}}, \overline{d})$ und $([-1,+1], d)$, wobei $d(x_1, x_2) = |x_1 - x_2|$ die Standardmetrik ist.

(ii) Für eine Folge (a_n) in $\mathbb{R}$ sind die folgenden Aussagen äquivalent:

- Die Folge (a_n) konvergiert uneigentlich gegen $+\infty$ bzw. $-\infty$ im Sinne von Definition 5.8.1 bzgl. der Standardmetrik auf $\mathbb{R}$.
- Die Folge $(h(a_n))$ konvergiert gegen $+1$ bzw. -1.
- Die Folge (a_n) konvergiert als Folge in $\overline{\mathbb{R}}$ gegen $+\infty$ bzw. $-\infty$ bzgl. der Metrik $\overline{d}$.

(iii) $(\overline{\mathbb{R}}, \overline{d})$ ist ein vollständig metrischer Raum.

15.3.4 Beweis wird weggelassen.

15.3.5 Räume beschränkter Funktionen Für eine beliebige Menge X ist

$$\mathcal{B}(X,\mathbb{K}) := \{f : X \to \mathbb{K} | f \text{ beschränkt } \}$$

der (∞–dimensionale) Vektorraum der beschränkten Funktionen. Die *Supremumsnorm*

$$\|\cdot\| : \begin{cases} \mathcal{B}(X,\mathbb{K}) & \to & \mathbb{R}_0^+ \\ f & \mapsto & \sup_{x\in X}\{|f(x)|\} \end{cases}$$

ist eine Norm auf $\mathcal{B}(X,\mathbb{K})$. Dabei zeigt man die Dreiecksungleichung (N3) für zwei beschränkte Funktionen f und g wie folgt: Für $x \in X$ ist

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|.$$

Bilden wir das Supremum über alle $x \in X$ auf der linken Seite, so folgt:

$$\|f + g\| \leq \|f\| + \|g\|.$$

Eine Funktionenfolge $(f_n) \subseteq \mathcal{B}(X, \mathcal{K})$ konvergiert genau dann bzgl. der Metrik

$$d(f, g) := \|f - g\|,$$

wenn sie gleichmäßig konvergiert.

15.3.6 Satz: Der Vektorraum $\mathcal{B}(X,\mathbb{K})$ ist bzgl. der Supremumsnorm und der zugehörigen Metrik ein vollständiger Raum.

15.3.7 Beweis

(1) Ermittlung der Grenzfunktion g: Es sei (f_n) eine Cauchy–Folge in $\mathcal{B}(X,\mathbb{K})$. Für festes $x \in X$ ist dann wegen

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|$$

auch die Zahlenfolge $(f_n(x))$ in $\mathbb{K}$ eine Cauchy–Folge. Aufgrund der Vollständigkeit von $\mathbb{K}$ konvergiert sie gegen eine Zahl, die wir mit $g(x)$ bezeichnen. Dadurch wird eine Funktion $g : X \to \mathbb{K}$ definiert.

(2) Es sei jetzt ein $\varepsilon > 0$ vorgegeben. Dazu existiert ein $N \in \mathbb{N}$, so dass für alle $n, m \geq N$ gilt:

$$\|f_n - f_m\| \leq \varepsilon.$$

Für festes $x \in X$ und $n, m \geq N$ gilt dann auch

$$|f_n(x) - f_m(x)| \leq \varepsilon.$$

In dieser Ungleichung lassen wir m gegen ∞ gehen, wegen der Stetigkeit des Betrags folgt

$$|f_n(x) - g(x)| \leq \varepsilon.$$

Da dies für alle $x \in X$ gilt, folgt

$$\|f_n - g\| \leq \varepsilon,$$

was bedeutet, dass die Folge f_n gleichmäßig gegen die Funktion g konvergiert.

(3) Wegen

$$\|g\| \leq \|g - f_N\| + \|f_N\| \leq \varepsilon + \|f_N\|$$

sehen wir auch noch, dass die Funktion g in dem Raum $\mathcal{B}(X,\mathbb{K})$ enthalten ist.

15.4 Topologie bei metrischen Räumen

15.4.1 Definitionen: Das System der offenen Mengen
(1) Eine Teilmenge X eines metrischen Raumes (M, d) heißt *offen (in M)*, wenn es zu jedem $x \in X$ ...

- eine Umgebung U von x gibt mit $U \subseteq X$ oder (äquivalent dazu)
- ein $\varepsilon > 0$ gibt, so dass $U_\varepsilon(x) \subseteq X$.

(2) Ist M ein metrischer Raum, so bezeichnen wir das *System aller offenen Mengen* $X \subseteq M$ mit

$$\mathcal{O}_M := \{X \subseteq M | X \text{ offen } \} \subseteq \mathcal{P}(M).$$

15.4.2 Beispiele
(1) Die leere Menge und die gesamte Menge M sind immer offen.

(2) ,,Intervalle ohne Randpunkte" von $\mathbb{R}$ sind offen

$$]a, b[, \qquad]-\infty, b[, \qquad]a, +\infty[, \qquad]-\infty, +\infty[.$$

(3) Die Teilmengen von $\mathbb{R}$

$$[a, b[, \qquad]a, b], \qquad \mathbb{Q}, \qquad \mathbb{R} \setminus \mathbb{Q},$$

sind nicht offen.

(4) Offene Quader im $\mathbb{R}^d$, das sind kartesische Produkte

$$]a_1, b_1[\times]a_2, b_2[\times \ldots \times]a_d, b_d[\subseteq \mathbb{R}^d$$

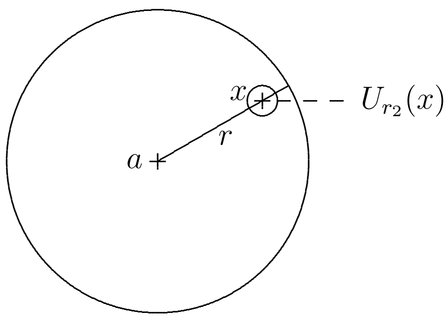
aus offenen Intervallen von $\mathbb{R}$, sind offen im $\mathbb{R}^d$.

(5) Ist eine Menge M mit der diskreten Metrik versehen, so sind alle Teilmengen offen.

(6) Offene Kugeln $U_r(a)$ sind immer offen. Ist nämlich $x \in U_r(a)$ und dabei $r_1 := d(x, a) < r$, so gilt für alle $y \in U_{r_2}(x)$ mit $r_2 := \frac{r-r_1}{2}$:

$$d(y, a) \leq d(y, x) + d(x, a) \leq r_2 + r_1 = \frac{r + r_1}{2} = \frac{r}{2} + \frac{r_1}{2} < r,$$

also $y \in U_r(a)$. Insgesamt bedeutet dies $U_{r_2}(x) \subseteq U_r(a)$.



15.4.3 Satz: Eigenschaften von $\mathcal{O}_M$

(O_i) Sind X_i, $i \in I$ (Indexmenge beliebig) beliebig viele offene Teilmengen, so ist auch die Vereinigung offen:

$$X_i \in \mathcal{O}_M \text{ für alle } i \in I \quad \Longrightarrow \quad \bigcup_{i \in I} X_i \in \mathcal{O}_M.$$

(O_{ii}) Sind X_i, $i \in I$ (Indexmenge endlich) endlich viele offene Teilmengen, so ist auch ihr Durchschnitt offen:

$$X_i \in \mathcal{O}_M \text{ für alle } i \in I \quad \Longrightarrow \quad \bigcap_{i \in I} X_i \in \mathcal{O}_M, \qquad \text{falls } I \text{ endlich.}$$

(O_{iii}) $\varnothing \in \mathcal{O}_M, \qquad M \in \mathcal{O}_M.$

15.4.4 Bemerkungen

(1) Genau genommen sind die beiden Eigenschaften (O_{iii}) in (O_i) bzw. (O_{ii}) enthalten, nämlich für den Fall leerer Indexmengen.

(2) Ein System von Teilmengen einer beliebigen Menge M (nicht notwendig metrischer Raum), dass die Eigenschaften (O_i) – (O_{iii}) erfüllt, heißt *Topologie* auf M.

15.4.5 Beweis Nur (ii) ist nicht ganz trivial. Es sei also $x \in \bigcap_{i \in I} X_i$ mit I endlich. Dann ist $x \in X_i$ für jedes $i \in I$. Da die X_i offen sind, gibt es jeweils ein ε_i, so dass $U_{\varepsilon_i}(x) \subseteq X_i$. Setzen wir jetzt

$$\varepsilon := \min\{\varepsilon_i | i \in I\} > 0,$$

so gilt $U_\varepsilon(x) \subseteq U_{\varepsilon_i}(x) \subseteq X_i$ für alle $i \in I$, also auch $U_\varepsilon(x) \subseteq \bigcap_{i \in I} X_i$.

15.4.6 Der Kern–Operator Wir definieren einen „Mengenoperator", der Teilmengen eines metrischen Raumes andere Teilmengen zuordnet.

15.4.7 Definition: Es sei (M, d) ein metrischer Raum. Wir definieren den *Kern–Operator*

$$\_^\circ \begin{cases} \mathcal{P}(M) & \to & \mathcal{P}(M) \\ X & \mapsto & X^\circ = \{x \in X \,\big|\, \text{Es existiert } \varepsilon > 0, \text{ so dass } U_\varepsilon(x) \subseteq X\} \end{cases}$$

15.4.8 Beispiele

(1) Es ist $\varnothing^\circ = \varnothing, \quad M^\circ = M$.

(2) Für Intervalle gilt

$$\begin{aligned} [a,b]^\circ &= \,]a,b]^\circ = [a,b[^\circ = \,]a,b[^\circ = \,]a,b[. \\ ]-\infty,b]^\circ &= \,]-\infty,b[^\circ = \,]-\infty,b[\\ [a,\infty[^\circ &= \,]a,\infty[^\circ = \,]a,\infty[. \end{aligned}$$

(3) Für die Teilmengen $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ gilt

$$\mathbb{Q}^\circ = \varnothing, \qquad (\mathbb{R} \setminus \mathbb{Q})^\circ = \varnothing.$$

15.4.9 Satz: Eigenschaften des Kern–Operators

Es seien X, Y beliebige Teilmengen von M. Dann gilt

(K_i) $\quad X \subseteq Y \quad \Longrightarrow \quad X^\circ \subseteq Y^\circ$ (Isotonie)

(K_{ii}) $\quad X^\circ \subseteq X$

(K_{iii}) $\quad (X^\circ)^\circ = X^\circ$ (Idempotenz)

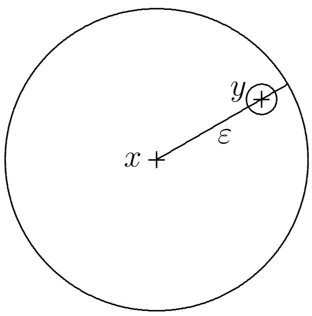
(K_{iv}) $\quad (X \cap Y)^\circ = X^\circ \cap Y^\circ$.

15.4.10 Beweis Wir zeigen (K_{iii}), die anderen Eigenschaften sind vergleichsweise einfach zu beweisen. Dabei ist die eine Richtung $(X^\circ)^\circ \subseteq X^\circ$ ein Spezialfall von (K_{ii}).

Es sei also $x \in X^\circ$. Gemäß Definition des Kern–Operators gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq X$.

Dieses $U_\varepsilon(x)$ ist aber ganz in X° enthalten: Zu jedem $y \in U_\varepsilon(x)$ gibt es nämlich ein $\varepsilon' := \frac{\varepsilon - d(x,y)}{2} > 0$, so dass $U_{\varepsilon'}(y) \subseteq U_\varepsilon(x) \subseteq X$. Das aber bedeutet $y \in X^\circ$.

Also gibt es zu jedem $x \in X^\circ$ eine Umgebung $U_\varepsilon(x) \subseteq X^\circ$. Das aber bedeutet $x \in (X^\circ)^\circ$.



15.4.11 Definition: Das System der abgeschlossenen Mengen

(1) Eine Teilmenge X eines metrischen Raumes (M, d) heißt *abgeschlossen (in M)*, wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ der Schluss

$$\left.\begin{array}{l} \forall_n \; x_n \in X \\ \lim_{n\to\infty} x_n = x \end{array}\right\} \quad \Longrightarrow \quad x \in X$$

gilt.

(2) Das System aller abgeschlossenen Mengen $X \subseteq M$ bezeichnen wir mit

$$\mathcal{A}_M \;:=\; \{X \in \mathcal{P}(M) \mid X \text{ abgeschlossen }\}.$$

15.4.12 Beispiele

(1) Die leere Menge und die gesamte Menge M sind immer abgeschlossen.

(2) Die folgenden Typen von Intervallen in $\mathbb{R}$ sind abgeschlossen

$$[a, b], \qquad]-\infty, b], \qquad [a, +\infty[.$$

(3) Die Teilmengen von $\mathbb{R}$

$$[a, b[, \qquad]a, b], \qquad \mathbb{Q}, \qquad \mathbb{R} \setminus \mathbb{Q},$$

sind nicht abgeschlossen.

(4) Abgeschlossene Quader im $\mathbb{R}^d$, das sind kartesische Produkte

$$[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d] \;\subseteq\; \mathbb{R}^d$$

aus abgeschlossenen Intervallen von $\mathbb{R}$, sind abgeschlossen im $\mathbb{R}^d$.

(5) Affine Unterräume im $\mathbb{R}^d$. Beispielsweise ist für fest gegebene Vektoren $v_1, \ldots, v_k \in \mathbb{R}^d$ und $b \in \mathbb{R}^d$ die Teilmenge

$$U \;=\; \left\{ x \in \mathbb{R}^d \;\middle|\; x \;=\; b + \alpha_1 v_1 + \ldots + \alpha_k v_k \quad \text{für} \quad \alpha_1, \ldots, \alpha_k \in \mathbb{R} \right\}$$

abgeschlossen.

(6) Ist eine Menge M mit der diskreten Metrik versehen, so sind alle Teilmengen abgeschlossen.

(7) Der Satz 12.2.1 sagt aus, dass die Menge $\mathcal{BC}(M, \mathbb{K})$ der beschränkten stetigen Funktionen $M \to \mathbb{K}$ abgeschlossen in der Menge $\mathcal{B}(M, \mathbb{K})$ der beschränkten Funktionen ist.

(8) Abgeschlossene Kugeln $B_r(a)$ sind immer abgeschlossen. Um das zu beweisen, nehmen wir an, dass der Grenzwert x einer in $B_r(a)$ liegenden Folge (x_n) nicht in $B_r(a)$ enthalten ist. Das bedeutet $r_1 \;:=\; d(x, a) > r$. Definiere $\varepsilon \;:=\; \frac{r_1 - r}{2} > 0$. In der ε–Umgebung $U_\varepsilon(x)$ muss ein x_n der Folge liegen: $d(x, x_n) < \varepsilon$. Es folgt

$$d(x_n, a) \overset{\text{DEU}}{\geq} \underbrace{d(x, a)}_{=r_1} - \underbrace{d(x, x_n)}_{<\varepsilon} > r_1 - \varepsilon = r_1 - \frac{r_1 - r}{2} = \frac{r_1 + r}{2} > r.$$

Das ist ein Widerspruch zu $x_n \in B_r(a)$.

15.4.13 Satz: Eigenschaften von $\mathcal{A}_M$

(A_i) Sind X_i, $i \in I$ (Indexmenge beliebig) beliebig viele abgeschlossene Teilmengen, so ist auch der Durchschnitt abgeschlossen:

$$X_i \in \mathcal{A}_M \text{ für alle } i \in I \quad \Longrightarrow \quad \bigcap_{i\in I} X_i \in \mathcal{A}_M.$$

(A_{ii}) Sind X_i, $i \in I$ (Indexmenge endlich) endlich viele abgeschlossene Teilmengen, so ist auch ihre Vereinigung abgeschlossen:

$$X_i \in \mathcal{A}_M \text{ für alle } i \in I \quad \Longrightarrow \quad \bigcup_{i\in I} X_i \in \mathcal{A}_M, \qquad \text{falls } I \text{ endlich.}$$

(A_{iii}) $\varnothing \in \mathcal{A}_M, \qquad M \in \mathcal{A}_M.$

15.4.14 Bemerkungen

(1) Genau genommen sind die beiden Eigenschaften (A_{iii}) in (A_{ii}) bzw. (A_i) enthalten, nämlich für den Fall leerer Indexmengen.

(2) Ein System von Teilmengen einer beliebigen Menge M, dass die Eigenschaften (i) – (iii) erfüllt, heißt *duale Topologie* auf M.

15.4.15 Beweis Nur (A_{ii}) ist nicht ganz trivial. Es sei also $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\bigcup_{i\in I} X_i$, I endlich, mit $x = \lim_{n\to\infty} x_n$. Da die Indexmenge I endlich ist, existiert mindestens ein $i \in I$ mit der Eigenschaft, dass X_i eine unendliche Teilfolge (x_{n_k}) von (x_n) enthält. Diese Teilfolge konvergiert ebenfalls gegen x, wegen der Abgeschlossenheit von X_i ist $x \in X_i$, also auch $x \in \bigcup_{i\in I} X_i$.

Wir definieren noch einen „Mengenoperator", der Teilmengen eines metrischen Raumes andere Teilmengen zuordnet.

15.4.16 Definition: Der Abschluss–Operator

(1) Es sei (M,d) ein metrischer Raum. Wir definieren den *Abschluss–Operator* = *topologischen Hüll–Operator*

$$\overline{} \left\{ \begin{array}{rcl} \mathcal{P}(M) & \to & \mathcal{P}(M) \\ X & \mapsto & \overline{X} = \{x \in M \,\Big|\, \text{Es existiert Folge } (x_n) \subseteq X \text{ mit } \lim\limits_{n\to\infty} x_n = x\} \end{array} \right.$$

(2) Ein Punkt $x \in M$ heißt *Berührpunkt* von X (innerhalb von M), wenn es eine Folge $(x_n) \subseteq X$ gibt, so dass $\lim_{n\to\infty} x_n = x$.

Also ist $\overline{X}$ die Menge der Berührpunkte von X.

(3) Wenn für zwei Teilmengen $X \subseteq Y \subseteq M$ gilt, dass $Y \subseteq \overline{X}$, so sagt man, X liege *dicht* in Y.

15.4.17 Beispiele

(1) Es ist $\overline{\varnothing} = \varnothing$, $\overline{M} = M$.

(2) Für Intervalle in $\mathbb{R}$ gilt

$$\begin{array}{rcccccccc} \overline{[a,b]} & = & \overline{]a,b]} & = & \overline{[a,b[} & = & \overline{]a,b[} & = & [a,b]. \\ \overline{]-\infty,b]} & = & \overline{]-\infty,b[} & = &]-\infty,b] \\ \overline{[a,\infty[} & = & \overline{]a,\infty[} & = & [a,\infty[. \end{array}$$

(3) Für die Teilmengen $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ von $\mathbb{R}$ gilt

$$\overline{\mathbb{Q}} = \mathbb{R}, \qquad \overline{\mathbb{R} \setminus \mathbb{Q}} = \mathbb{R}.$$

Die Begründung erfolgt einfach mit dem Satz 4.5.7 über die Dichtheit.

(4) Ist $(x_n) \subseteq M$ eine konvergente Folge mit Grenzwert x, so ist

$$\overline{\{x_n | n \in \mathbb{N}\}} = \{x_n | n \in \mathbb{N}\} \cup \{x\}.$$

(5) Ist $z \in \mathbb{R}^d$, so gilt $\overline{\mathbb{R}^d \setminus \{z\}} = \mathbb{R}^d$.

(6) Der Beweis von Satz 10.3.2 zeigt, dass die Teilmenge von $\mathcal{B}([c,d],\mathbb{R})$ der Treppenfunktionen dicht in der Teilmenge der Regelfunktionen liegt.

15.4.18 Satz: Eigenschaften des Abschluss–Operators

Es seien X, Y beliebige Teilmengen von M. Dann gilt

(H_i) $X \subseteq Y \implies \overline{X} \subseteq \overline{Y}$ (Isotonie)

(H_{ii}) $X \subseteq \overline{X}$

(H_{iii}) $\overline{\overline{X}} = \overline{X}$ (Idempotenz)

(H_{iv}) $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$.

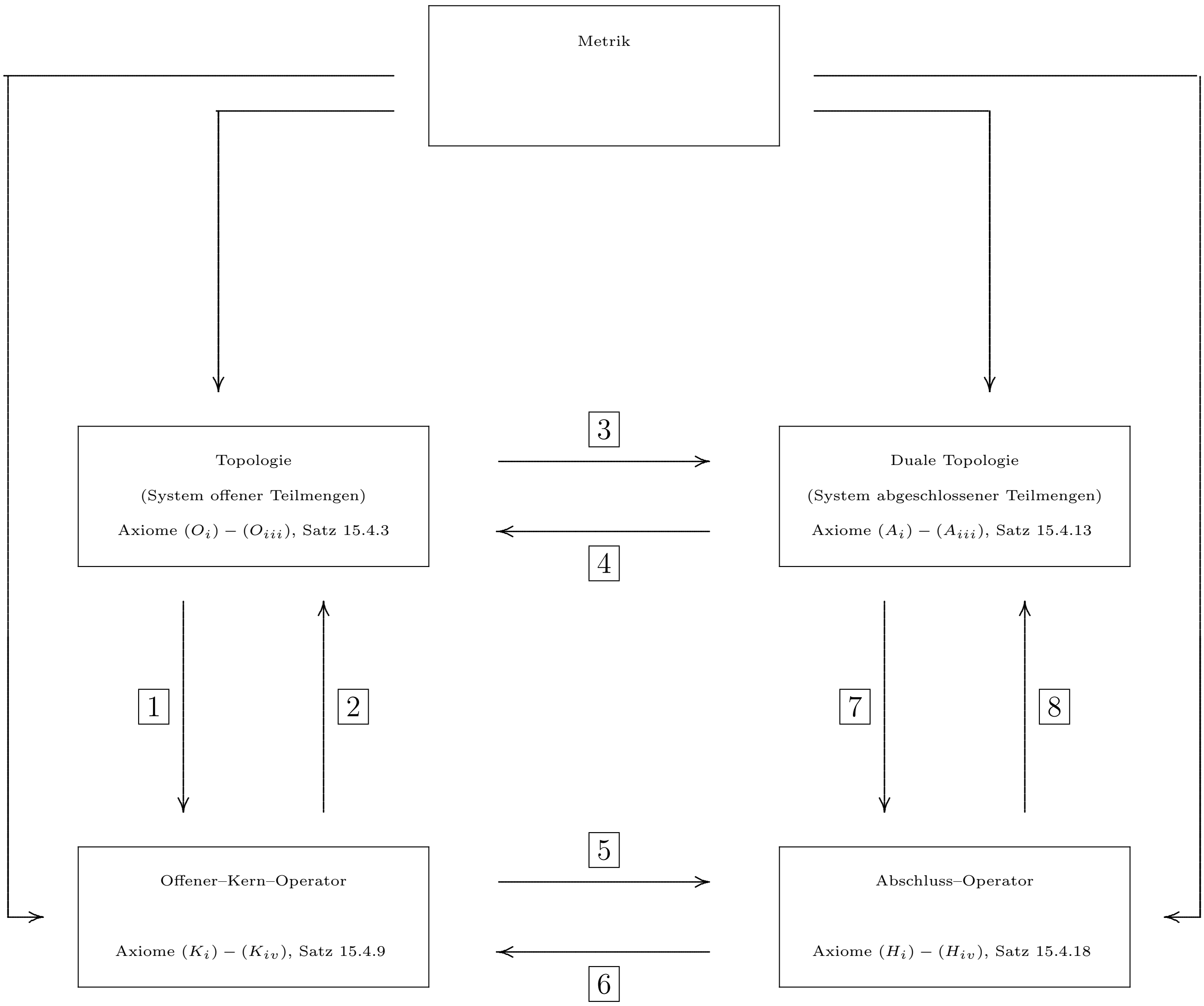
15.4.19 Beweis Wir zeigen nur die Idempotenz: Die eine Richtung $\overline{\overline{X}} \supseteq \overline{X}$ folgt aus (H_{ii}). Ist nun $x \in \overline{\overline{X}}$, so existiert definitionsgemäß eine Folge (x_n) in $\overline{X}$ mit $\lim_{n\to\infty} x_n = x$. Aufgrund der Definition des Abschluss–Operators gibt es zu jedem $n \in \mathbb{N}$ ein $y_n \in X$, so dass $d(y_n, x_n) < \frac{1}{n}$. Daraus folgt aber wegen

$$d(y_n, x) \le d(y_n, x_n) + d(x_n, x) < \frac{1}{n} + \underbrace{d(x_n, x)}_{\to 0},$$

dass die Folge (y_n) gegen x konvergiert. Also ist $x \in \overline{X}$.

15.4.20 Topologische Räume

Es sei M eine Menge. Wir betrachten dazu das folgende Diagramm



Die Pfeile sind als ,,induziert" bzw. ,,definiert" zu lesen. Beispielsweise bedeutet der Pfeil zwischen Metrik und Topologie, dass

- eine Metrik d auf einer Menge M
- eine Topologie $\mathcal{O}_M$ auf dieser Menge

definiert (induziert).

Die vielen anderen Pfeile bedeuten dann im einzelnen

$\boxed{1}$ Ist auf der Menge M ein System offener Teilmengen (Topologie) gegeben, so wird durch

$$X^\circ := \bigcup \Big\{ O \subseteq X \,\Big|\, O \text{ offen } \Big\}$$

ein Offener–Kern–Operator definiert: Der offene Kern von X entsteht dadurch, dass man alle in X enthaltenen offenen Mengen vereinigt.

$\boxed{2}$ Ist auf der Menge M ein Offener–Kern–Operator gegeben, so wird durch

$$X \text{ offen} \quad :\Longleftrightarrow \quad X = X^\circ$$

ein System offener Teilmengen (Topologie) definiert: Die offenen Mengen sind genau die, die mit ihrem offenen Kern übereinstimmen.

$\boxed{3}$ Ist auf der Menge M ein System offener Teilmengen (Topologie) gegeben, so wird durch

$$X \text{ abgeschlossen} \quad :\Longleftrightarrow \quad M \setminus X \text{ offen}$$

ein System abgeschlossener Mengen (Duale Topologie) definiert: Die abgeschlossenen Mengen sind genau die Komplemente der offenen Mengen.

$\boxed{4}$ Ist auf der Menge M ein System abgeschlossener Teilmengen (Duale Topologie) gegeben, so wird durch

$$X \text{ offen} \quad :\Longleftrightarrow \quad M \setminus X \text{ abgeschlossen}$$

ein System offener Mengen (Topologie) definiert: Die offenen Mengen sind per definitionem genau die Komplemente der abgeschlossenen Mengen.

$\boxed{5}$ Ist auf der Menge M ein Offener–Kern–Operator gegeben, so wird durch

$$\overline{X} := M \setminus (M \setminus X)^\circ$$

ein Abschluss–Operator definiert: Man nimmt das Komplement des offenen Kern des Komplements.

$\boxed{6}$ Ist auf der Menge M ein Abschluss–Operator gegeben, so wird durch

$$X^\circ := M \setminus (\overline{M \setminus X})$$

ein Offener–Kern–Operator definiert: Man nimmt das Komplement des Abschlusses des Komplements.

$\boxed{7}$ Ist auf der Menge M ein System abgeschlossener Teilmengen (Duale Topologie) gegeben, so wird durch

$$\overline{X} \ := \bigcap \Big\{ A \supseteq X \ \Big| \ A \text{ abgeschlossen } \Big\}$$

ein Abschluss–Operator definiert: Der Abschluss von X entsteht dadurch, dass man alle Mengen schneidet, in denen X enthaltenen ist.

$\boxed{8}$ Ist auf der Menge M ein Abschluss–Operator gegeben, so wird durch

$$X \text{ abgeschlossen} \quad :\Longleftrightarrow \quad X = \overline{X}$$

ein System abgeschlossener Teilmengen (Duale Topologie) definiert: Die abgeschlossenen Mengen sind per definitionem genau die, die mit ihrem abgeschlossenen Kern übereinstimmen.

15.5 Stetigkeit

Wir betrachten zunächst stetige Funktionen zwischen zwei beliebigen metrischen Räumen.

15.5.1 Satz: Offene Mengen und Stetigkeit Es seien (M_1, d_1) und (M_2, d_2) zwei metrische Räume, $f : M_1 \to M_1$ eine Abbildung.

(i) Für $a \in M_1$ sind die folgenden Aussagen äquivalent:

- (A) f ist in a stetig.
- (B) Zu jeder Umgebung U_2 von $f(a)$ in M_2 gibt es eine Umgebung U_1 von a in M_1, so dass $f(U_1) \subseteq U_2$.
- (C) Zu jeder offenen Umgebung U_2 von $f(a)$ in M_2 gibt es eine offene Umgebung U_1 von a in M_1, so dass $f(U_1) \subseteq U_2$.
- (D) Zu jeder abgeschlossenen Umgebung U_2 von $f(a)$ in M_2 gibt es eine abgeschlossene Umgebung U_1 von a in M_1, so dass $f(U_1) \subseteq U_2$.

(ii) Die folgenden Aussagen sind äquivalent:

- (A) f ist stetig (auf M_1).
- (B) Das Urbild $f^{-1}(Y)$ jeder offenen Teilmenge $Y \subseteq M_2$ ist offen in M_1.
- (C) Das Urbild $f^{-1}(Y)$ jeder abgeschlossenen Teilmenge $Y \subseteq M_2$ ist abgeschlossen in M_1.

15.5.2 Bemerkung Bemerkenswert an diesen die Stetigkeit charakterisierenden Aussagen ist, dass sie nicht Bezug auf die Metrik nehmen, sondern nur die topologischen Begriffe der Offenheit bzw. Abgeschlossenheit enthalten. Würde man den Satz zur Definition von Stetigkeit heranziehen, so wäre diese in den Kontext der topologischen Räume verallgemeinert.

15.5.3 Beweis Zu (i): Die verschiedenen Versionen können leicht auf die ε–δ–Definition der Stetigkeit zurückgeführt werden.

Zu (ii): (A) $\Rightarrow$ (C): Es sei $Y \subseteq M_2$ abgeschlossen und $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge in $f^{-1}(Y)$ mit Grenzwert a. Dann ist auch die durch $b_n := f(a_n)$ definierte Folge $(b_n)_{n\in\mathbb{N}} \subseteq Y$ konvergent. Ihr Grenzwert b ist wegen der Abgeschlossenehit von Y in Y enthalten. Es gilt insgesamt:

$$f(a) = f(\lim_{n\to\infty} a_n) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} b_n = b \in Y,$$

also $a \in f^{-1}(Y)$. Also ist $f^{-1}(Y)$ abgeschlossen.

(C) $\Rightarrow$ (B) beweisen wir durch die folgende Implikationskette:

$$\begin{aligned} & Y \text{ offen in } M_2 \\ \Longrightarrow\ & M_2 \setminus Y \text{ abgeschlossen} \\ \Longrightarrow\ & f^{-1}(M_2 \setminus Y) \text{ abgeschlossen} \end{aligned}$$

$\Longrightarrow f^{-1}(M_2) \setminus f^{-1}(Y)$ abgeschlossen

$\Longrightarrow M_1 \setminus f^{-1}(Y)$ abgeschlossen

$\Longrightarrow f^{-1}(Y)$ offen in M_1.

Dabei haben wir für die dritte Schlussfolgerung benutzt, dass das Urbild einer Differenz zweier Mengen A und B gleich der Mengendifferenz der Urbilder ist:

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

Das ist eine Eigenschaft beliebiger Funktionen, die wir zu Beginn der Vorlesung AYS1 beweisen hätten sollen (können). Das ist eine Übung für Sie.

(B) $\Rightarrow$ (A): Wir betrachten ein beliebiges $a \in X$, es sei $\varepsilon > 0$ vorgegeben. Das Urbild der offenen Menge $U_\varepsilon(f(a))$ ist nach Voraussetzung (B) offen, es existiert also ein $\delta > 0$, so dass die offene δ–Kugel $U_\delta(a)$ um a in $f^{-1}(U_\varepsilon(f(a)))$ enthalten ist. Daraus folgt aber:

$$\begin{array}{rcl} \|x-a\| < \delta & \Longrightarrow & x \in U_\delta(a) \subseteq f^{-1}(U_\varepsilon(f(a))) \quad \Longrightarrow \quad f(x) \in U_\varepsilon(f(a)) \\ & \Longrightarrow & \|f(x) - f(a)\| < \varepsilon. \end{array}$$

Also ist f stetig in a. Da a beliebig war, ist f überhaupt stetig.

Wir wechseln jetzt zu der spezielleren Situation, dass die beteiligten metrischen Räume Teilmengen von euklidischen Räumen sind.

15.5.4 Definition: Es sei $X \subseteq \mathbb{R}^d$ und $f : X \to \mathbb{R}^w$ eine Funktion, $a \in X$.

(1) Für jedes $j \in \{1, \ldots, w\}$ können wir die j–te *Koordinatenfunktion*

$$f_j : \left\{ \begin{array}{rcl} X & \to & \mathbb{R} \\ x & \mapsto & f(x)_j \end{array} \right.$$

bilden.

(2) Für jedes $k \in \{1, \ldots, d\}$ heißt die Funktion

$$f(a_1, \ldots, a_{k-1}, \underset{\substack{\uparrow \\ k}}{\cdot}, a_{k+1}, \ldots, a_d) : \left\{ \begin{array}{rcl} X_k & \to & \mathbb{R}^w \\ x & \mapsto & f(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_d) \end{array} \right.$$

die *partielle Funktion für die k–te Stelle.* Dabei enthält X_k genau diejenigen Zahlen $x \in \mathbb{R}$, für die $(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_d)$ in X enthalten ist.

15.5.5 Satz: Stetigkeit bei partiellen Funktionen Es sei $a \in X \subseteq \mathbb{R}^d$.

(i) Eine Funktion $f : X \to \mathbb{R}^w$ ist genau dann stetig in a, wenn alle Koordinatenfunktionen $(j = 1, \ldots, w)$ stetig in $a \in X$ sind.

(ii) Aus der Stetigkeit aller partiellen Funktionen in a_k $(k = 1, \ldots, d)$ kann nicht auf die Stetigkeit der Funktionen f in a geschlossen werden.

15.5.6 Beweis

(i) Dies liegt daran, dass Folgen im $\mathbb{R}^w$ genau dann konvergieren, wenn alle Koordinatenfolgen konvergieren.

(ii) Für die Funktion

$$f : \left\{ \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x_1, x_2) & \mapsto & \left\{ \begin{array}{ll} 1, & \text{falls } x_1 \cdot x_2 \neq 0, \\ 0, & \text{sonst} \end{array} \right. \end{array} \right.$$

sind die partiellen Funktionen $f(\cdot\,, 0)$ und $f(0, \cdot\,)$ konstant Null, also stetig. f ist aber nicht stetig in $(0, 0)$, da in jeder Umgebung von $(0, 0)$ der Funktionswert 1 angenommen wird.

Ein noch interessanteres Gegenbeispiel ist folgendes: Betrachte (Zeichnung) die Funktion

$$f : \left\{ \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x_1, x_2) & \mapsto & \left\{ \begin{array}{ll} 1, & \text{falls } 0 < |x_2| < x_1^2, \\ 0, & \text{sonst.} \end{array} \right. \end{array} \right.$$

Jede Folge, die ganz auf einer Geraden durch den Ursprung gegen den Ursprung konvergiert, ist schließlich eine Nullfolge. Trotzdem ist f nicht stetig in $(0, 0)$.

15.6 Kompakte Mengen

15.6.1 Definition: Häufungspunkt von Folgen Es sei (a_n) ein Folge in einem metrischen Raum M. $a \in M$ heißt *Häufungspunkt* der Folge, wenn in jeder Umgebung von a **unendlich viele** Folgenglieder liegen.
Beachte den kleinen, aber feinen Unterschied zum Grenzwertbegriffs: $a \in M$ heißt *Grenzwert* der Folge, wenn in jeder Umgebung von a **fast alle** Folgenglieder liegen.

15.6.2 Beispiele (1) Die durch $a_n = (-1)^n$ definierte Folge besitzt die beiden Häufungspunkte -1 und $+1$.

(2) Die durch $a_n = (-1)^n + \frac{1}{n}$ definierte Folge besitzt ebenfalls die beiden Häufungspunkte -1 und $+1$.

(3) Die durch $a_n = n$ definierte Folge besitzt keinen Häufungspunkt.

(4) Gibt es Folgen mit unendlich vielen Häufungspunkten?

(5) Die Folge, definiert durch

$$a_n := \begin{cases} 1, & \text{falls } n \text{ gerade,} \\ n, & \text{falls } n \text{ ungerade,} \end{cases}$$

besitzt 1 als Häufungspunkt.

15.6.3 Satz Ist a Grenzwert einer Folge, so ist a der einzige Häufungspunkt. Das Beispiel (5) zeigt, dass die Umkehrung falsch ist.

15.6.4 Beweis

Dass a Häufungspunkt ist, folgt sofort aus der Definition. Würden ein weiterer Häufungspunkt $\widetilde{a}$ existieren, so würde die Umgebung

$$U_\varepsilon(\widetilde{a}) \qquad \text{mit } \varepsilon := \frac{d(a, \widetilde{a})}{2}$$

unendlich viele Folgenglieder enthalten. Diese sind dann aber nicht in der Umgebung $U_\varepsilon(a)$ enthalten, was ein WIDERSPRUCH dazu ist, dass a Grenzwert der Folge ist.

15.6.5 Satz: Es sei (a_n) eine Folge in M und $a \in M$. Die beiden folgenden Aussagen sind äquivalent:

(A) a ist Häufungspunkt.

(B) Es existiert eine konvergente Teilfolge mit Grenzwert a.

15.6.6 Beweis (A) $\Rightarrow$ (B): Es sei a Häufungspunkt. Für jedes $k \in \mathbb{N}$ gibt es in der Umgebung $U_{\frac{1}{k}}(a)$ ein Folgenglied a_n, dessen Index n wir als n_k auswählen. Die so definierte Teilfolge

$$\begin{array}{ccccc} \mathbb{N} & \to & \mathbb{N} & \to & M \\ k & \mapsto & n_k & \mapsto & a_{n_k} \end{array}$$

konvergiert offensichtlich gegen a.

(B) $\Rightarrow$ (A): Es sei $(a_{n_k})_{k\in\mathbb{N}}$ die Teilfolge und U eine Umgebung von a. Es existiert ein $K \in \mathbb{N}$, so dass für alle $k \geq K$ die Teilfolgenglieder a_{n_k} in U liegen. Das sind aber unendlich viele Folgenglieder der ursprünglichen Folge.

15.6.7 Definitionen: Kompakte Mengen

Es sei M ein metrischer (oder allgemeiner: ein topologischer) Raum. X sei eine Teilmenge von M.

1. Es sei I eine beliebige Menge. Sie heißt in diesem Zusammenhang *Indexmenge*. Für jedes $i \in I$ sei eine Teilmenge U_i gegeben. Wir betrachten das ,,Mengensystem"

   $$\mathcal{U} = \left\{ U_i \;\middle|\; i \in I \right\}.$$

   Das Mengensystem heißt *endlich* bzw. *abzählbar*, wenn das für die Indexmenge zutrifft.

2. Ein Mengensystem $\mathcal{U}$ heißt eine *offene Überdeckung für* X, wenn

   - jedes $U_i \in \mathcal{U}$ offen ist und
   - $X \subseteq \bigcup_{i\in I} U_i$.

3. Es seien $\mathcal{U}$ und $\mathcal{V}$ zwei offene Überdeckungen für X. Dann heißt $\mathcal{V}$ *offene Teilüberdeckung* von $\mathcal{U}$ für X, wenn

   $$\mathcal{V} \subseteq \mathcal{U},$$

   das heißt jede offene Menge des Mengensystems $\mathcal{V}$ gehört auch zum Mengensystem $\mathcal{U}$. Alternativ kann man sagen, dass es eine Teilmenge I' der Indexmenge I gibt, so dass

   $$\mathcal{V} = \left\{ U_i \;\middle|\; i \in I' \right\}.$$

4. Eine Teilmenge X eines metrischen Raumes M heißt *kompakt*, wenn es

   - zu jeder offenen Überdeckung $\mathcal{U}$ für X
   - eine **endliche** Teilüberdeckung $\mathcal{U}'$ gibt.

   Die in dieser Definition enthaltene Aussage wird als *Heine–Borel–Eigenschaft* bezeichnet.

5. Ist M als Teilmenge von sich selbst kompakt, so heißt M auch *kompakter Raum*.

15.6.8 Bemerkung

Diese Definitionen muten dem Anfänger als unzugänglich bis überzogen–gekünstelt an. Tatsächlich spielt die Kompaktheit (von Teilmengen) eines metrischen Raumes bei vielen Konstruktionen (und darauf beruhenden) Sätzen der Analysis eine eminent wichtige Rolle.

Der Übergang zu einer **endlichen** Teilüberdeckung ermöglicht dabei die Durchführung von **endlichen** Prozessen wie die Bildung von

- **endlichen** Summen von **endlich** vielen Summanden,
- **endlichen** Produkten von **endlich** vielen Faktoren,
- **endlichen** Suprema bzw. Infima bei **endlich** vielen Zahlen,
- Vereinigungen bzw. Schnitten von **endlich** vielen Teilmengen,

und so weiter ...

15.6.9 Beispiele

1. Eine endliche Teilmenge X eines metrischen Raumes M ist immer kompakt. Ist nämlich $\mathcal{U}$ eine offene Überdeckung für X, so wähle man aus $\mathcal{U}$ für jedes $x \in X$ einen Index $i(x)$, so dass $x \in U_{i(x)}$ aus. Dann ist

$$\mathcal{U}' = \{U_{i(x)} | x \in X\}$$

eine endliche Teilüberdeckung von $\mathcal{U}$ für X.

2. Es sei (x_n) eine konvergente Folge in einem metrischen Raum M mit $x = \lim_{n\to\infty} x_n$. Dann ist die Menge

$$X = \{x_n | n \in \mathbb{N}\} \cup \{x\}$$

kompakt. Ist nämlich $\mathcal{U}$ eine offene Überdeckung für X, so gibt es ein $V \in \mathcal{U}$ mit $x \in V$. Aufgrund der Definition der Konvergenz gibt es ein $N \in \mathbb{N}$, so dass $x_n \in V$ für alle $n \geq N$. Für jedes $n < N$ nehme man einen Index $i(n) \in I$ mit $x_n \in U_{i(n)}$. Das Mengensystem

$$\mathcal{U}' = \Big\{ U_{i(n)} \;\Big|\; n \in \mathbb{N}, n < N \Big\} \cup \Big\{ V \Big\}$$

ist dann eine endliche Teilüberdeckung.

3. $\mathbb{N}$ ist als Teilmenge des metrischen Raumes $\mathbb{R}$ nicht kompakt. Es ist nämlich

$$\mathcal{U} \;=\; \Big\{ \;]n - \tfrac{1}{2}, n + \tfrac{1}{2}[\;\Big|\; n \in \mathbb{N} \Big\}$$

eine offene Überdeckung für $\mathbb{N}$, zu der keine endliche Teilüberdeckung existiert.

4. Das offene Intervall $]0, 1[$ ist nicht kompakt. Es ist nämlich

$$\mathcal{U} = \Big\{ \;]\tfrac{1}{n}, 1[\;\Big|\; n \in \mathbb{N} \Big\}$$

eine offene Überdeckung, zu der es keine endliche Teilüberdeckung gibt.

5. Die Menge $[0, 1] \cap \mathbb{Q}$ der rationalen Zahlen im Einheitsintervall ist nicht kompakt. Es ist nämlich

$$\mathcal{U} = \Big\{ \;] - 1, \tfrac{\sqrt{2}}{2} - \tfrac{1}{n}[\;\Big|\; n \in \mathbb{N} \Big\} \cup \Big\{ \;]\tfrac{\sqrt{2}}{2} + \tfrac{1}{n}, 2[\;\Big|\; n \in \mathbb{N} \Big\}$$

eine offene Überdeckung für diese Menge, zu der es keine endliche Teilüberdeckung gibt.

6. Früher hatten wir bereits Intervalle der Form $[c, d]$ von $\mathbb{R}$ als kompakt bezeichnet. Wir werden gleich sehen, dass dies — im Sinne von Heine–Borel — zutrifft.

15.6.10 Hauptsatz über kompakte Mengen

Für eine Teilmenge X von $\mathbb{R}^d$ sind die folgenden Aussagen äquivalent:

(A) X ist kompakt (im Sinne der obigen Heine–Borel–Definition).

(B) Jede Folge in X besitzt einen Häufungspunkt in X.

(C) X ist beschränkt und abgeschlossen.

15.6.11 Bemerkungen

1. Ersetzt man in diesem Satz $\mathbb{R}^d$ durch einen beliebigen metrischen Raum M, so sind immer noch die Implikationen $(A) \Leftrightarrow (B) \Rightarrow (C)$ gültig. Wir werden diesen allgemeineren Kontext auch im Beweis durch die Notation zum Ausdruck bringen.

2. Für ein abgeschlossenes Intervall von $\mathbb{R}$ kennen wir die Implikation (C) $\Rightarrow$ (B) bereits aus der Analysis I. Das ist der Satz von Bolzano–Weierstraß.

3. Häufig wird „Kompaktheit" für Teilmengen des $\mathbb{R}^d$ durch die Eigenschaft (C) definiert. Die Implikation $(C) \Rightarrow (A)$ erscheint dann als Satz von Heine–Borel.

15.6.12 Beweis

Zu $(A) \Rightarrow (B)$

Es sei (a_n) eine Folge in X. Wir suchen einen Häufungspunkt.

(1) Wir definieren eine wachsende Folge offener Mengen

$$Y_1 \subseteq Y_2 \subseteq Y_3 \subseteq \ldots\ldots$$

durch

$$Y_n \; := \; M \setminus \overline{\{a_n, a_{n+1}, a_{n+2}, \ldots\ldots\}}$$

und zeigen:

(2) $\{Y_n | n \in \mathbb{N}\}$ ist keine Überdeckung von M.

Wäre diese Menge nämlich doch eine, so gäbe es aufgrund der Voraussetzung (A) eine endliche offene Teilüberdeckung $\{Y_{n_1}, Y_{n_2}, \ldots, Y_{n_\ell}\}$, also

$$M \subseteq Y_{n_1} \cup Y_{n_2} \cup Y_{n_\ell} = Y_{\min\{n_1,\ldots,n_\ell\}} \subsetneqq M \qquad (\text{Widerspruch}).$$

(3) Es gibt also (mindestens) einen Punkt

$$a \;\in\; M \setminus \left(\bigcup_{n\in\mathbb{N}} Y_n \right) = M \setminus \left(\bigcup_{n\in\mathbb{N}} (M \setminus \overline{\{a_n, a_{n+1}, a_{n+2}, \ldots\ldots\}}\,) \right) =$$

$$M \setminus \left(M \setminus \bigcap_{n\in\mathbb{N}} \overline{\{a_n, a_{n+1}, a_{n+2}, \ldots\ldots\}} \right) = \bigcap_{n\in\mathbb{N}} \overline{\{a_n, a_{n+1}, a_{n+2}, \ldots\ldots\}}\,.$$

(4) Das aber bedeutet, dass a Häufungspunkt von (a_n) ist: Für jedes $\varepsilon > 0$ und jedes $n \in \mathbb{N}$ ist nämlich

$$a \in \overline{\{a_n, a_{n+1}, a_{n+2}, \ldots\ldots\}},$$

deshalb existiert ein $\ell \geq n$, so dass

$$d(a_\ell, a) \leq \varepsilon.$$

(5) Um zu zeigen, dass $a \in X$ ist, nehmen wir das Gegenteil an: $a \in M \setminus X$.

Zu jedem $x \in X$ gibt es

(a) eine offene Umgebung U_x von x und

(b) eine offene Umgebung V_x von a, so dass

(c) $U_x \cap V_x = \varnothing$.

(6) Das Mengensystem

$$\mathcal{U} = \left\{ U_x \;\middle|\; x \in X \right\}$$

ist eine offene Überdeckung für X. Da X kompakt ist, gibt es eine endliche Teilmenge $\{x_1, \ldots, x_n\} \subseteq X$, so dass

$$\mathcal{U}' = \left\{ U_{x_1}, \ldots, U_{x_n} \right\}$$

eine endliche offene Teilüberdeckung für X ist.

(7) Die Menge

$$V = V_{x_1} \cap \ldots \cap V_{x_n}$$

ist eine offene Umgebung von a in $M \setminus X$, was wir noch genauer überlegen:

- $a \in V$, da $a \in V_{x_j}$ für alle $j = 1, \ldots, n$.
- V ist offen, da V ein endlicher Schnitt der offenen Mengen V_{x_j}, $j = 1, \ldots, n$, ist.
- $V \subseteq M \setminus X$.

Zur Begründung des letzten Punktes: Wäre dies nicht der Fall, so gäbe es ein

$$\begin{aligned}
& z \in V \cap X \\
\Longrightarrow \quad & z \in V_{x_j} \cap X \text{ für alle } j = 1, \ldots, n \\
\Longrightarrow \quad & z \notin U_{x_j} \cap X \text{ für alle } j = 1, \ldots, n \\
\Longrightarrow \quad & z \notin \Big(\bigcup_{j=1}^{n} U_{x_j}\Big) \cap X = X \qquad (\textsc{Widerspruch}).
\end{aligned}$$

(8) Da also eine offene Umgebung von a in $M \setminus X$ existiert, kann a nicht Häufungspunkt der Folge $(a_n) \subseteq X$ sein. Wir müssen die in (5) gemachte Annahme verwerfen.

Zu (B) $\Rightarrow$ (C):

Wenn X unbeschränkt wäre, so gäbe es zu jedem $n \in \mathbb{N}$ ein $a_n \in X$ mit $\|a_n - a_1\| \geq n$. Die Folge (a_n) besitzt dann aber keine konvergente Teilfolge, also auch keinen Häufungspunkt.

Wenn (a_n) eine Folge in X ist, die in M konvergiert, so ist der Grenzwert ein (sogar der einzige) Häufungspunkt, also gemäß (B) in X enthalten. Das aber bedeutet, dass X abgeschlossen ist.

Zu $(C) \Rightarrow (A)$:

(1) Als Vorbereitung zeigen wir:

> $\circledast_1$ Es sei $X \subseteq Y \subseteq M$, wobei M ein metrischer Raum, Y eine kompakte Teilmenge von M ist. Dann gilt: X abgeschlossen $\Rightarrow$ X kompakt.

Es sei also X eine abgeschlossene Teilmenge von Y und $\mathcal{U}$ eine offene Überdeckung für X. Wir nehmen zu $\mathcal{U}$ eine einzige offene Menge dazu

$$\mathcal{V} = \mathcal{U} \cup \{M \setminus X\}$$

und erhalten so eine offene Überdeckung von Y. Da Y kompakt ist, besitzt diese Überdeckung eine endliche Teilüberdeckung $\mathcal{V}'$. Dann ist aber

$$\mathcal{U}' = \mathcal{U} \cap \mathcal{V}'$$

eine endliche offene Teilüberdeckung für X.

(2)

> $\circledast_2$ (Schachtelungsprinzip) Es sei M ein vollständig metrischer Raum. Ist (A_n) eine absteigende Folge von nichtleeren abgeschlossenen Teilmengen
>
> $$A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots \qquad \text{mit} \qquad \lim_{n\to\infty} \operatorname{diam} A_n = 0,$$
>
> so gibt es einen Punkt $a \in M$ mit
>
> $$\bigcap_{n\in\mathbb{N}} A_n = \{a\}.$$

Enthielte die Menge in der letzten Zeile zwei Punkte, so könnte der Durchmesser nicht beliebig klein werden.

Es sei (a_n) eine Folge mit $a_n \in A_n$. Zu $\varepsilon > 0$ gibt es $N \in \mathbb{N}$, so dass $\operatorname{diam} A_N \leq \varepsilon$ und deshalb

$$d(a_n, a_m) \leq \operatorname{diam} A_N = \varepsilon \qquad \text{für alle } n, m \geq N.$$

Damit ist (a_n) eine Cauchy–Folge, sie ist also konvergent mit Grenzwert a. Für ein festes $\ell \in \mathbb{N}$ ist $(a_n)_{n\geq\ell}$ eine Folge in A_ℓ mit $\lim_{n\to\infty} a_n = a$. Da A_ℓ abgeschlossen ist, gilt $a \in A_\ell$.

Wir beweisen jetzt:

$\circledast_3$ Der abgeschlossene Würfel

$$Q = [-a, +a]^d \subseteq \mathbb{R}^d$$

ist kompakt.

Es sei $\mathcal{U}$ eine offene Überdeckung dieses Würfels.
Wir nehmen an, dass es keine endliche Teilüberdeckung gibt.

(1) Wir definieren rekursiv eine Folge von Würfeln Q_m, so dass folgende Eigenschaften erfüllt sind:

(a) $Q_{m+1} \subseteq Q_m$

(b) $\operatorname{diam} Q_{m+1} = \frac{1}{2} \cdot \operatorname{diam} Q_m$

(c) Q_{m+1} besitzt keine endliche Teilüberdeckung $\mathcal{U}' \subseteq \mathcal{U}$.

(2) Zunächst setzen wir $Q_0 = Q$.

(3) Es sei jetzt der Quader Q_m schon gefunden. Wir zerlegen ihn in 2^d abgeschlossene Teilquader dadurch, dass wir ihn bzgl. jeder der d Dimensionen halbieren. Als Q_{m+1} suchen wir einen dieser Teilquader heraus, der keine endliche Teilüberdeckung aus $\mathcal{U}$ hat. Es muss ja mindestens einer diese Eigenschaft haben, sonst hätte bereits der Vorgängerquader Q_m eine endliche Teilüberdeckung. Es ist klar, dass der neue Quader Q_{m+1} die drei obigen Eigenschaften (a) – (c) besitzt.

(4) Auf diese Quaderfolge wenden wir das Schachtelungsprinzip $\circledast_2$ an. Es gibt also einen Punkt q mit

$$\bigcap_{m \in \mathbb{N}} Q_m = \{q\}.$$

(5) Zu q gibt es ein $U \in \mathcal{U}$, so dass $q \in U$.

(6) Zu U gibt es ein $\varepsilon > 0$, so dass $U_\varepsilon(q) \subseteq U$.

(7) Zu $\varepsilon > 0$ gibt es ein $m \in \mathbb{N}$, so dass $Q_m \subseteq U_\varepsilon(q)$.

(8) Insgesamt gilt also

$$Q_m \subseteq U_\varepsilon(q) \subseteq U \in \mathcal{U},$$

was bedeutet, dass Q_m doch eine endliche Teilüberdeckung aus $\mathcal{U}$, nämlich die ein–elementige $\{U\}$ hat. Widerspruch.

Die beiden Aussagen $\circledast_1$ und $\circledast_3$ liefern jetzt die Implikation (C) $\Rightarrow$ (A).

15.7 Stetige Funktionen auf kompakten Mengen

15.7.1 Satz: Stetige Funktionen auf kompakten Mengen

Ist $K \subseteq \mathbb{R}^d$ eine kompakte Teilmenge und $f : K \to \mathbb{R}^\ell$ stetig, so ist auch die Bildmenge $f(K)$ kompakt.

15.7.2 Beweis Er erfolgt mittels der Heine–Borel–Eigenschaft. Es sei also

$$\mathcal{U} = \left\{ U_i \,\middle|\, i \in I \right\}$$

eine offene Überdeckung des Bildes $f(K)$. Dann ist aber

$$\mathcal{V} = \left\{ f^{-1}(U_i) \,\middle|\, i \in I \right\}$$

eine offene Überdeckung von K (vgl. Satz 15.5.1). Da K kompakt ist, gibt es zu $\mathcal{V}$ eine endliche Teilüberdeckung

$$\mathcal{V}' = \left\{ f^{-1}(U_i) \,\middle|\, i \in I' \right\}$$

mit $I' \subseteq I$. Das Bild dieser Teilüberdeckung unter f

$$\left\{ f(f^{-1}(U_i)) \,\middle|\, i \in I' \right\} = \left\{ U_i \,\middle|\, i \in I' \right\}$$

ist dann eine offene Teilüberdeckung von $f(K)$.

Zur Illustration sei noch ein zweiter Beweis auf der Grundlage der Bolzano–Weiserstrass–Eigenschaft vorgeführt:
Es sei $(b_n) \subseteq f(K)$ eine beliebige Folge im Bild von K. Zu jedem $n \in \mathbb{N}$ gibt es ein $a_n \in K$, so dass $f(a_n) = b_n$. Die Folge $n \mapsto a_n$ besitzt gemäß Satz 15.6.10 eine Teilfolge $k \mapsto a_{n_k}$, deren Grenzwert

$$a = \lim_{k\to\infty} a_{n_k} \in K$$

existiert. Da f stetig ist, existiert auch

$$b := \lim_{k\to\infty} b_{n_k} = \lim_{k\to\infty} f(a_{n_k}) = f(\lim_{k\to\infty} a_{n_k}) = f(a) \in f(K).$$

Also besitzt $(b_n)_{n\in\mathbb{N}}$ eine in $f(K)$ konvergente Teilfolge. Wieder mit Satz 15.6.10 folgt die Behauptung.

15.7.3 Korollar Eine stetige Funktion $f : K \to \mathbb{R}$ auf einer kompakten Menge $K \subseteq \mathbb{R}^d$ nimmt ihr Minimum und Maximum an, d.h. es gibt $a, b \in K$, so dass

$$f(a) = \min f(K) \quad \text{und} \quad f(b) = \max f(K).$$

15.7.4 Beweis Nach Satz 15.7.1 ist die Teilmenge $f(K) \subseteq \mathbb{R}$ kompakt. Da also die Menge $f(K)$ beschränkt und abgeschlossen ist, enthält sie ihr Minimum und Maximum.

15.8 Gleichmäßig stetige Funktionen nicht klausurrelevant

15.8.1 Definition: Gleichmäßig stetige Funktionen Es sei $X \subseteq \mathbb{R}^d$ und $f : X \to \mathbb{R}^w$ eine Funktion. f heißt *gleichmäßig stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass

$$\|f(x) - f(y)\| < \varepsilon, \qquad \text{falls } \|x - y\| < \delta.$$

Beachte, dass das δ **unabhängig** von einer festen Stelle $a \in [c, d]$ gewählt sein muss.

15.8.2 Satz:

(i) Ist $f : [c, d] \to \mathbb{R}$ stetig, so ist f sogar gleichmäßig stetig.

(ii) Allgemeiner: Ist X kompakt, so ist $f : X \to \mathbb{R}^\ell$ gleichmäßig stetig.

15.8.3 Beweis Wir beweisen nur die erste Aussage. Dafür sei ein $\varepsilon > 0$ vorgegeben. Wir betrachten die folgende Familie von Aussagen $(\mathcal{A}_z)_{z \in [c,d]}$:

Es existiert ein δ_z, so dass

$$\|f(x) - f(y)\| < \varepsilon, \qquad \text{falls } x, y \in [c, z] \text{ und } |x - y| < \delta_z.$$

Wenn $\mathcal{A}_d$ gezeigt ist, ist die Aussage (1) des Satzes bewiesen. Wir nehmen an, es gibt ein $z \in [c, d]$, so dass $\mathcal{A}_z$ nicht gilt, und setzen dann

$$a \; := \; \sup\{z | \mathcal{A}_z \text{ gilt nicht } \}.$$

Da f in a stetig ist, gibt es ein $\gamma > 0$, so dass

$$\|f(x) - f(a)\| < \frac{\varepsilon}{2}, \qquad \text{falls } x \in [a - \gamma, a + \gamma]. \qquad (*)$$

Wir setzen jetzt

$$\delta_{a+\gamma} \; := \; \min\{\delta_{a-\frac{\gamma}{2}}, \frac{\gamma}{2}\} > 0$$

und zeigen, dass $\mathcal{A}_{a+\gamma}$ gilt. Es seien also $x, y \in [c, a + \gamma]$ mit $|x - y| < \delta_{a+\gamma}$.

1. Fall: $x, y \in [c, a - \frac{\gamma}{2}]$. Da $\mathcal{A}_{a-\frac{\gamma}{2}}$ wahr und $\delta_{a+\gamma} < \delta_{a-\frac{\gamma}{2}}$ ist, gilt $\|f(x) - f(y)\| < \varepsilon$.

2. Fall: $x, y \in [a - \gamma, a + \gamma]$. Dann gilt aufgrund von $(*)$

$$\|f(x) - f(y)\| \le \|f(x) - f(a)\| + \|f(a) - f(y)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

3. Fall: $x \in [c, a - \gamma], y \in [a - \gamma, a + \gamma]$. Dann folgt aber mit $|x - y| < \delta_{a+\gamma} \le \frac{\gamma}{2}$, dass

$$y < x + \frac{\gamma}{2} \le a - \gamma + \frac{\gamma}{2} = a - \frac{\gamma}{2},$$

also $x, y \in [c, a - \frac{\gamma}{2}]$. Das ist aber wieder Fall 1.

16 Differenzierbarkeit von Funktionen mehrerer Variabler

Im folgenden sei X eine offene Teilmenge von $\mathbb{R}^d$, $a \in X$ und $f{:}X \to \mathbb{R}^w$ eine Funktion. Die Vektoren aus $\mathbb{R}^d$ (**D**efinitionsmenge) bzw. $\mathbb{R}^w$ (**W**ertemenge) werden generell als Spaltenvektoren aufgefasst

16.1 Partielle Differenzierbarkeit

16.1.1 Definition: Partielle Differenzierbarkeit

Es seien eine Koordinatennummer $j \in \{1,\ldots,w\}$ für die Wertemenge und eine Koordinatennummer $k \in \{1,\ldots,d\}$ für die Definitionsmenge fixiert.

Die j–te Koordinatenfunktion $f_j : X \to \mathbb{R}$ heißt in a *partiell differenzierbar* nach der k–ten Variablen x_k, wenn die partielle Funktion

$$f_j(a_1,\ldots,a_{k-1},\underset{\underset{k}{\uparrow}}{\cdot},a_{k+1},\ldots,a_d) : \begin{cases} X_k & \to & \mathbb{R}^w \\ x & \mapsto & f_j(a_1,\ldots,a_{k-1},x,a_{k+1},\ldots,a_d) \end{cases}$$

an der Stelle a_k (wie gewöhnlich) differenzierbar ist. Die Ableitung heißt *partielle Ableitung der j–ten Koordinatenfunktion nach der k–ten Variablen* und wird notiert als:

$$\frac{\partial f_j}{\partial x_k}(a) \quad := \quad \partial_k f_j(a) \;:=\; \frac{d}{dx} f_j(a_1,\ldots,a_{k-1},x,a_{k+1},\ldots,a_d)\Big|_{x=a_k}.$$

Die Funktion f heißt *partiell differenzierbar auf* X, wenn die obige Definition für alle Stellen in X und für alle Variablen $j \in \{1,\ldots,w\}$ und $k \in \{1,\ldots,d\}$ zutrifft. In diesem Fall bilden die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k} : X \to \mathbb{R}$ einen Satz von $w \cdot d$ Funktionen auf X.

16.1.2 Beispiel Wir betrachten die Funktion

$$\begin{cases} \mathbb{R}^+ \times \mathbb{R} & \to & \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \mapsto & \begin{pmatrix} x_1^2\,x_2 + \sin(x_1)\cdot e^{x_2} \\ x_1^{x_2} \end{pmatrix} \end{cases}$$

Unter Beachtung von $x_1^{x_2} = e^{x_2 \cdot \ln x_1}$ kann man berechnen:

$$\frac{\partial f_1}{\partial x_1}(x_1,x_2) \;=\; 2x_1\,x_2 + \cos(x_1)\cdot e^{x_2} \qquad \frac{\partial f_1}{\partial x_2}(x_1,x_2) \;=\; x_1^2 + \sin(x_1)\cdot e^{x_2}$$

$$\frac{\partial f_2}{\partial x_1}(x_1,x_2) \;=\; x_2 \cdot x_1^{x_2-1} \qquad \frac{\partial f_2}{\partial x_2}(x_1,x_2) \;=\; \ln(x_1)\cdot x_1^{x_2}$$

16.2 Totale Differenzierbarkeit

16.2.1 Definition: Totale Differenzierbarkeit an einer Stelle

f heißt *(total) differenzierbar an der Stelle* a mit der *Jacobi–Matrix* (oder *Funktional–Matrix*)

$$J_a = J = \begin{pmatrix} J_{11} & \cdots & J_{1d} \\ \vdots & & \vdots \\ J_{w\,1} & \cdots & J_{w\,d} \end{pmatrix} \in \mathbb{R}^{w \times d}$$

als *Ableitung*, wenn eine der folgenden äquivalenten Aussagen zutrifft:

(A) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für $x \in X$ mit $\|x - a\| \leq \delta$ gilt:

$$\|f(x) - f(a) - J \cdot (x - a)\| \leq \varepsilon \cdot \|x - a\|.$$

(B) Es ist $\displaystyle \lim_{x \overset{\bullet}{\to} a} \frac{\|f(x) - f(a) - J(x-a)\|}{\|x - a\|} = 0.$

(C) Es ist $\displaystyle \lim_{h \overset{\bullet}{\to} 0} \frac{\|f(a+h) - f(a) - Jh\|}{\|h\|} = 0.$

16.2.2 Satz Es existiert höchstens eine Matrix J_a mit dieser Eigenschaft.

16.2.3 Beweis Angenommen, es gibt zwei solcher Matrizen J und $\widetilde{J}$. Fixiere einen Vektor $H \in \mathbb{R}^d \setminus \{0\}$. Dann gilt

$$\begin{aligned} \frac{\|(J - \widetilde{J})H\|}{\|H\|} &= \lim_{t \overset{\bullet}{\to} 0} \frac{\|(J - \widetilde{J})tH\|}{\|tH\|} \\ &= \lim_{t \overset{\bullet}{\to} 0} \frac{\|[f(a+h) - f(a) - JtH] - [f(a+h) - f(a) - \widetilde{J}tH]\|}{\|tH\|} \\ &\leq \lim_{t \overset{\bullet}{\to} 0} \frac{\|f(a+h) - f(a) - JtH\| + \|f(a+h) - f(a) - \widetilde{J}tH\|}{\|tH\|} \quad = \quad 0, \end{aligned}$$

woraus folgt, dass $(J - \widetilde{J})H = 0$. Da H beliebig war, muss $J - \widetilde{J}$ gleich der Nullmatrix sein.

16.2.4 Definition: Totale Differenzierbarkeit auf der Definitionsmenge

Wenn f an allen Stellen von X differenzierbar ist, dann heißt f *differenzierbar* schlechthin. In diesem Fall ist die Ableitung als Funktion

$$f' = Df : \begin{cases} X & \to & \mathbb{R}^{w \times d} \\ a & \mapsto & J_a \end{cases}$$

definiert.

16.2.5 Beispiele

(1) Ist $d = w = 1$, so zeigt ein Vergleich der definierenden Ungleichung mit der in Satz/Definition 9.1.1, dass dann der neue Differenzierbarkeitsbegriff mit dem alten übereinstimmt. Die ,,alte" Ableitung $f'(a)$ ist als 1×1–Matrix ein Spezialfall der ,,neuen" Ableitung.

(2) Ist $F : \mathbb{R}^d \to \mathbb{R}^w$ eine lineare Abbildung, die dann als Multiplikation mit einer Matrix A aufgefasst werden kann, so ist

$$\|F(x) - F(a) - A \cdot (x-a)\| = 0 \qquad \text{für alle } x, a \in \mathbb{R}^d.$$

Damit ist die Differenzierbarkeit an allen Stellen $a \in \mathbb{R}^d$ gegeben. Die Ableitungsfunktion ist konstant: $F'(a) = DF(a) = A$ für alle $a \in \mathbb{R}^d$.

Weitere Beispiele folgen nach den nächsten Sätzen.

Der folgende Satz gestattet es, die (totale) Differenzierbarkeit ohne Rückgriff auf Matrixrechnung zu fassen:

16.2.6 Satz:

(i) Die beiden folgenden Aussagen sind äquivalent:

(A) Die Funktion $f : X \to \mathbb{R}^w$ ist differenzierbar in $a \in X$ mit Ableitung $J \in \mathbb{R}^{w \times d}$.

(B) Jede der w Koordinatenfunktionen $f_j : X \to \mathbb{R}$ ist differenzierbar in a mit Ableitung

$$\begin{pmatrix} J_{j1} & \cdots & J_{jd} \end{pmatrix} \qquad (j\text{–te Zeile von } J)$$

(ii) Die beiden folgenden Aussagen sind äquivalent:

(A) Die Funktion $f : X \to \mathbb{R}$ (also $w = 1$) ist differenzierbar in $a \in X$ mit Ableitung

$$J = \begin{pmatrix} J_1 & \cdots & J_d \end{pmatrix} \in \mathbb{R}^{1 \times d} \qquad \text{(Zeilenvektor).}$$

(B) Es gibt d Zahlen $J_1, \ldots, J_d$, so dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|f(x) - f(a) - \underbrace{[J_1(x_1 - a_1) + \ldots + J_d(x_d - a_d)]}| \;\le\; \varepsilon \cdot \|x - a\|,$$
$$\text{falls } x \in X, \|x - a\| \;\le\; \delta.$$

16.2.7 Beweis Die Aussage (i) kann man darauf zurückführen, dass Vektoren im $\mathbb{R}^w$ genau dann „klein" bzgl. der Norm gemacht werden können, wenn die Koordinaten „klein" bzgl. des Betrags gemacht werden können.

Die Aussage (ii) ist einfach eine Umschreibung der Definition der Ableitung, das Matrixprodukt $J \cdot (x - a)$ ist als reelle Zahl (oberhalb der geschweiften Klammer) ausgeschrieben.

16.2.8 Bemerkungen Der Ausdruck oberhalb der geschweiften Klammer in Aussage (B) von (ii) kann auch als Skalarprodukt aufgefasst werden:

$$\Big[J_1(x_1 - a_1) + \ldots + J_d(x_d - a_d)\Big] = \langle \begin{pmatrix} J_1 \\ \vdots \\ J_d \end{pmatrix}, \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_d - a_d \end{pmatrix} \rangle.$$

Wenn man herausstellen will, dass das Skalarprodukt involviert ist, benutzt man die Schreibweise $\operatorname{grad} f$ oder ∇f anstelle von Df für die als Spaltenvektor aufzufassende Ableitung. Die Unterscheidung von Matrixmultiplikation und Skalarprodukt scheint überflüssig, sie gewinnt aber an Relevanz, wenn man andere Skalarprodukte als das Standard–Skalarprodukt betrachtet oder Variablentransformationen in der Definitionsmenge $X \subseteq \mathbb{R}^d$ durchführt.

Der folgende Satz stellt eine Verbindung zwischen partieller und totaler Differenzierbarkeit sowie Stetigkeit her.

16.2.9 Satz: Verschiedene Differenzierbarkeitsbegriffe

Betrachte die folgenden Aussagen:

(A) Alle Koordinatenfunktionen f_j ($j = 1, \ldots, w$) sind auf einer offenen Umgebung $U_\delta(a)$ von a (nach allen Variablen) partiell differenzierbar und alle partiellen Ableitungen

$$\frac{\partial f_j}{\partial x_k}, \quad j = 1, \ldots, w, \quad k = 1, \ldots, d$$

sind auf dieser Umgebung stetig.

(B) f ist in a total differenzierbar.

(C) Alle Koordinatenfunktionen f_j ($j = 1, \ldots, w$) sind in a (nach allen Variablen) partiell differenzierbar.

(D) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\|f(x) - f(a)\| \leq \big(\underbrace{\|Df(a)\| + \varepsilon}_{=:L}\big) \cdot \|x - a\| \qquad \text{für alle } x \in U_\delta(a),$$

d.h. f ist in a Lipschitz–stetig.

(E) f ist stetig in a.

Dann gelten die Implikationen:

$$\begin{array}{ccccc} (A) & \Longrightarrow & (B) & \Longrightarrow & (C) \\ & & \Downarrow & & \\ & & (D) & \Longrightarrow & (E) \end{array}$$

Ist (B) erfüllt, so gilt

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_d}(a) \\ \vdots & & \vdots \\ \frac{\partial f_w}{\partial x_1}(a) & \cdots & \frac{\partial f_w}{\partial x_d}(a) \end{pmatrix}.$$

16.2.10 Beweis $(A) \Rightarrow (B)$ Das wesentliche an diesem Beweis tritt schon beim Fall $d = 2$ und $w = 1$ auf. Es sei ein $\varepsilon > 0$ vorgegeben. Wähle ein $\delta > 0$, so dass die beiden folgenden Abschätzungen erfüllt sind:

$$\left|\frac{\partial f}{\partial x_1}(x) - \frac{\partial f}{\partial x_1}(a)\right| < \frac{\varepsilon}{2}, \qquad \text{falls } x \in U_\delta(a),$$

$$\left|f(a_1, x_2) - f(a_1, a_2) - \frac{\partial f}{\partial x_2}(a_1, a_2) \cdot (x_2 - a_2)\right| \leq \frac{\varepsilon}{2} \cdot |x_2 - a_2|, \qquad \text{falls } |x_2 - a_2| \leq \delta.$$

Es sei jetzt $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in U_\delta(a)$ beliebig. Wir wenden den Mittelwertsatz der Differentialrechnung auf die partielle Funktion $f(\cdot\,, x_2)$ an: Es gibt ein ξ zwischen x_1 und a_1, so dass

$$f(x_1, x_2) - f(a_1, x_2) = \frac{\partial f}{\partial x_1}(\xi, x_2) \cdot (x_1 - a_1).$$

Wegen $\begin{pmatrix} \xi \\ x_2 \end{pmatrix} \in U_\delta(a)$ gilt dann weiter

$$\begin{aligned}
&\Big| f(x) - f(a) - [\frac{\partial f}{\partial x_1}(a) \cdot (x_1 - a_1) + \frac{\partial f}{\partial x_2}(a) \cdot (x_2 - a_2)] \Big| \leq \\
&\Big| \overbrace{f(x_1, x_2) - f(a_1, x_2)}^{\frac{\partial f}{\partial x_1}(\xi, x_2)\cdot(x_1 - a_1)} - \frac{\partial f}{\partial x_1}(a_1, a_2) \cdot (x_1 - a_1) \Big| + \\
&\Big| f(a_1, x_2) - f(a_1, a_2) - \frac{\partial f}{\partial x_2}(a_1, a_2) \cdot (x_2 - a_2) \Big| \leq \\
&\frac{\varepsilon}{2} \cdot |x_1 - a_1| + \frac{\varepsilon}{2} \cdot |x_2 - a_2| \leq \frac{\varepsilon}{2} \cdot \|x - a\| + \frac{\varepsilon}{2} \cdot \|x - a\| = \varepsilon \cdot \|x - a\|.
\end{aligned}$$

$(B) \Rightarrow (C)$ Wir zeigen die Aussage nur für $d = 2$ und stellvertretend für $j = k = 1$. Dazu sei $\varepsilon > 0$ vorgegeben. Es gibt nach Voraussetzung eine $w \times 2$–Matrix $J = Df(a)$ und $\delta > 0$, so dass für x mit $\|x - a\| \leq \delta$

$$\left\| f(x_1, x_2) - f(a_1, a_2) - J \cdot \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \end{pmatrix} \right\| \leq \varepsilon \cdot \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\|$$

Da der Betrag einer Koordinate eines Vektors immer kleiner oder gleich der Norm ist, gilt weiter

$$\begin{aligned}
&\Big| f_1(x_1, x_2) - f_1(a_1, a_2) - \Big[J_{11}(x_1 - a_1) + J_{12}(x_2 - a_2) \Big] \Big| \leq \\
&\qquad\qquad \varepsilon \cdot \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\|
\end{aligned}$$

Dann gilt aber speziell für alle $z \in \mathbb{R}$ mit $\begin{pmatrix} z \\ a_2 \end{pmatrix} \in U_\delta(a)$

$$|f_1(z, a_2) - f_1(a_1, a_2) - J_{11}(z - a_1)| \leq \varepsilon \cdot \left\| \begin{pmatrix} z \\ a_2 \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\| = \varepsilon |z - a_1|.$$

Das aber bedeutet gerade, dass die partielle Funktion $f_1(\cdot\,, a_2)$ in a_1 differenzierbar ist mit Ableitung J_{11}. Wegen der Eindeutigkeit der Ableitung gilt:

$$\frac{\partial f_1}{\partial x_1}(a_1, a_2) \quad = \quad J_{11} \quad = \quad (Df(a))_{11}.$$

Der Beweis von $(B) \Rightarrow (D) \Rightarrow (E)$ ist fast gleich wie der des entsprechenden Satzes 9.3.1 aus der 1–dimensionalen Differentialrechnung. Wir führen ihn wegen der geänderten Schreibweisen noch einmal durch.

$(B) \Rightarrow (D)$ Zu $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\|f(x) - f(a) - Df(a) \cdot (x - a)\| \leq \varepsilon \cdot \|x - a\|, \qquad \text{falls } \|x - a\| \leq \delta.$$

Für diese x gilt dann weiter

$$\begin{aligned} \|f(x) - f(a)\| &\leq \|f(x) - f(a) - Df(a) \cdot (x - a)\| + \|Df(a) \cdot (x - a)\| \\ &\leq \varepsilon \cdot \|x - a\| + \|Df(a)\| \cdot \|x - a\| = \underbrace{(\varepsilon + \|Df(a)\|)}_{=:L} \cdot \|x - a\|. \end{aligned}$$

$(D) \Rightarrow (E)$ Ist $\varepsilon > 0$ vorgegeben, so wählt man $\delta_1 = \min\{\delta, \frac{\varepsilon}{L}\}$ mit δ und L gemäß (D). Dann gilt für alle x mit $\|x - a\| < \delta_1$

$$\|f(x) - f(a)\| \leq L \cdot \|x - a\| \leq \frac{L \cdot \varepsilon}{L} = \varepsilon.$$

16.2.11 Beispiele

(1) Die Implikation $(A) \Rightarrow (B)$ zeigt, dass die Funktion f aus dem Beispiel in Abschnitt 16.1 auf X total differenzierbar ist mit Ableitung

$$\begin{aligned} Df(x_1, x_2) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_1, x_2) & \frac{\partial f_1}{\partial x_2}(x_1, x_2) \\ \frac{\partial f_2}{\partial x_1}(x_1, x_2) & \frac{\partial f_2}{\partial x_2}(x_1, x_2) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 \, x_2 + \cos(x_1) \cdot e^{x_2} & x_1^2 + \sin(x_1) \cdot e^{x_2} \\ x_2 \cdot x_1^{x_2 - 1} & \ln(x_1) \cdot x_1^{x_2} \end{pmatrix}. \end{aligned}$$

(2) Bereits aus der 1–dimensionalen Differentialrechnung wissen wir, dass die Implikation $(A) \Rightarrow (B)$ nicht umgekehrt werden kann.

(3) Das Beispiel aus dem Beweis von Satz 15.5.5(ii)

$$f : \begin{cases} \mathbb{R}^2 & \to \mathbb{R} \\ (x_1, x_2) & \mapsto \begin{cases} 1, & \text{falls } x_1 \cdot x_2 \neq 0, \\ 0, & \text{sonst} \end{cases} \end{cases}$$

zeigt, dass die Implikation $(C) \Rightarrow (B)$ im allgemeinen falsch ist.

16.2.12 Satz: Totale Differenzierbarkeit bzgl. ,,abgetrennter" Variabler

Es seien $X_1 \subseteq \mathbb{R}^{d_1}$, $X_2 \subseteq \mathbb{R}^{d_2}$ jeweils offen. Dann ist auch $X_1 \times X_2 \subseteq \mathbb{R}^{d_1+d_2}$ offen. Die Abbildung $f : X_1 \times X_2 \to \mathbb{R}^{w_1} \times \mathbb{R}^{w_2}$ sei an der Stelle $a = (a_1, a_2) \in X_1 \times X_2$ differenzierbar.

Dann sind die Funktionen

$$f_1(\cdot\,, a_2) : \left\{ \begin{array}{rcl} X_1 & \to & \mathbb{R}^{w_1} \\ x & \mapsto & f_1(x, a_2) \end{array} \right. \qquad f_1(a_1, \cdot\,) : \left\{ \begin{array}{rcl} X_2 & \to & \mathbb{R}^{w_1} \\ x & \mapsto & f_1(a_1, x) \end{array} \right.$$

$$f_2(\cdot\,, a_2) : \left\{ \begin{array}{rcl} X_1 & \to & \mathbb{R}^{w_2} \\ x & \mapsto & f_2(x, a_2) \end{array} \right. \qquad f_2(a_1, \cdot\,) : \left\{ \begin{array}{rcl} X_2 & \to & \mathbb{R}^{w_2} \\ x & \mapsto & f_2(a_1, x) \end{array} \right.$$

in a_1 bzw. a_2 differenzierbar. Die Ableitungen hängen über die Gleichung

$$Df(a) = \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) \end{pmatrix}.$$

zusammen. $D_1 f_j$ und $D_2 f_j$ sind die ,,Teil–Ableitungen" der Funktion f_j nach dem X_1–Argument bzw. dem X_2–Argument.

Der Beweis besteht letztlich in einer Umschreibung des Beweises von $(B) \Rightarrow (C)$ des vorhergehenden Satzes 16.2.9.

16.2.13 Satz: Kettenregel

Es seien $X \subseteq \mathbb{R}^d, Y \subseteq \mathbb{R}^n$ offen und $f : X \to Y$, $g : Y \to \mathbb{R}^w$ Funktionen,

$$X \xrightarrow{f} Y \xrightarrow{g} \mathbb{R}^w.$$

Ist f in $a \in X$ differenzierbar und g in $b = f(a)$ differenzierbar, so ist auch $g \circ f : X \to \mathbb{R}^w$ in a differenzierbar mit Ableitung

$$\underbrace{(g \circ f)'(a)}_{w \times d} = \underbrace{g'(b)}_{w \times n} \cdot \underbrace{f'(a)}_{n \times d}.$$

16.2.14 Beweis Der Beweis besteht in einer geeigneten Modifizierung des Beweises 9.5.4 von Satz 9.5.1 über die 1–dimensionale Kettenregel: Im wesentlichen müssen Gleichheiten der Form

$$|a \cdot b| = |a| \cdot |b|$$

für reelle Zahlen durch Abschätzungen der Form

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|$$

(vgl. Submultiplikativität der Schur–Norm, Sätzchen 15.1.7) ersetzt werden.
Zunächst zerlegen wir den Ausdruck in der Definition der Differenzierbarkeit geeignet, setzen dabei zur Abkürzung $y = f(x)$

$$\begin{aligned}
&\Big\| g(f(x)) - g(f(a)) - g'(b)f'(a)(x-a) \Big\| = \\
&\Big\| g(y) - g(b) - g'(b)(y-b) + g'(b)(f(x) - f(a)) - g'(b)f'(a)(x-a) \Big\| \leq \\
&\Big\| g(y) - g(b) - g'(b)(y-b) \Big\| + \|g'(b)\| \cdot \Big\| f(x) - f(a) - f'(a)(x-a) \Big\| \overset{\text{Ziel}}{\leq} \varepsilon \|x-a\|.
\end{aligned}$$

Unser Ziel ist es jetzt, zu einem vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ zu finden, so dass die letzte Ungleichung für $\|x - a\| \leq \delta$ erfüllt ist. Dafür müssen die Teile des Ausdrucks vor dem Ungleichheitszeichen geeignet abgeschätzt werden. Dazu gehen wir der Reihe nach so vor:

1. Es existiert (vgl. Satz 16.2.9 (D)) ein $\delta_1 > 0$ und ein $L > 0$, so dass

   $$\|y - b\| = \|f(x) - f(a)\| \leq L \cdot \|x - a\|, \qquad \text{falls } \|x - a\| \leq \delta_1.$$

2. Es existiert gemäß Definition der Differenzierbarkeit ein $\delta_2 > 0$, so dass

   $$\Big\| f(x) - f(a) - f'(a)(x-a) \Big\| \leq \tfrac{\varepsilon}{2 \cdot (\|g'(b)\| + 1)} \cdot \|x - a\|, \quad \text{falls} \quad \|x - a\| \leq \delta_2.$$

3. Es existiert ein $\delta_3 > 0$, so dass

   $$\Big\| g(y) - g(b) - g'(b)(y-b) \Big\| \leq \tfrac{\varepsilon}{2L} \|y - b\|, \qquad \text{falls } \|y - b\| \leq \delta_3.$$

Setzen wir jetzt $\delta := \min\{\delta_1, \delta_2, \frac{\delta_3}{L}\}$, so gilt für $x \in D$ mit $\|x-a\| \le \delta$ zunächst

$$\|y-b\| = \|f(x)-f(a)\| \le L \cdot \|x-a\| \le L \cdot \tfrac{\delta_3}{L} \le \delta_3$$

und dann

$$\underbrace{\underbrace{\Big\|g(y)-g(b)-g'(b)(y-b)\Big\|}_{\le \frac{\varepsilon}{2L}\|y-b\|}}_{\le \frac{\varepsilon}{2}\|x-a\|} + \underbrace{\|g'(b)\| \underbrace{\Big\|f(x)-f(a)-f'(a)(x-a)\Big\|}_{\le \frac{\varepsilon}{2\cdot(\|g'(b)\|+1)}\|x-a\|}}_{\le \frac{\varepsilon}{2}\|x-a\|} \le \varepsilon\|x-a\|.$$

16.2.15 Korollar: Spezialfall der Kettenregel Es seien $X \subseteq \mathbb{R}, Y \subseteq \mathbb{R}^n$ offen und $f: X \to Y$, $g: Y \to \mathbb{R}$ Funktionen,

$$X \xrightarrow{f} Y \xrightarrow{g} \mathbb{R}.$$

Ist f in $a \in X$ differenzierbar und g in $b = f(a)$ differenzierbar, so ist auch die reelle Funktion $g \circ f : X \to \mathbb{R}$ differenzierbar mit Ableitung

$$(g \circ f)'(a) = \tfrac{\partial g}{\partial y_1}(b) \cdot \tfrac{\partial f_1}{\partial x}(a) + \ldots + \tfrac{\partial g}{\partial y_n}(b) \cdot \tfrac{\partial f_n}{\partial x}(a).$$

16.2.16 Bemerkungen

(1) Bezeichnet man die Funktion f mit y, da sie ja „Werte y" in dem Raum Y annimmt, so lautet die obige Formel

$$\frac{d}{dx}\, g(y_1(x), y_2(x), y \ldots, y_n(x)) = \tfrac{\partial g}{\partial y_1}(b) \cdot \tfrac{\partial y_1}{\partial x}(a) + \ldots + \tfrac{\partial g}{\partial y_n}(b) \cdot \tfrac{\partial y_n}{\partial x}(a).$$

Das ist mathematisch etwas ungenau, aber rechentechnisch eingängig. Physiker und Ingenieure rechnen so.

(2) Wählt man im obigen Korollar die Funktion g als

$$g : \left\{ \begin{array}{rcl} Y & \to & \mathbb{R} \\ \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} & \mapsto & y_1 \cdot \ldots \cdot y_n, \end{array} \right.$$

so ist gemäß obiger Formel

$$\begin{aligned} (g \circ f)(x) &= f_1(x) \cdot f_2(x) \cdot \ldots \cdot f_n(x) \\ (g \circ f)'(a) &= \tfrac{\partial g}{\partial y_1}(b) \cdot \tfrac{\partial y_1}{\partial x}(a) + \ldots + \tfrac{\partial g}{\partial y_n}(b) \cdot \tfrac{\partial y_n}{\partial x}(a) \\ &= f_2(a) \cdots f_n(a) f_1'(a) + \cdots + f_1(a) \cdots f_{n-1}(a) f_n'(a), \end{aligned}$$

wir haben die Regel für Produkte erneut bewiesen.

16.3 Der Abschätzungssatz

16.3.1 Satz: Abschätzungssatz (1) Es sei $J \subseteq \mathbb{R}$ ein offenes Intervall. Die beiden Abbildungen $f : J \to \mathbb{R}^w$ und $g : J \to \mathbb{R}$ seien differenzierbar.

(i) Gilt $\quad \|f'(x)\| \leq g'(x) \quad$ für alle $x \in J$,

so folgt $\quad \|f(y) - f(x)\| \leq g(y) - g(x) \quad$ für alle $x, y \in J$ mit $y \geq x \qquad (*)$

(ii) Gilt $\quad \|f'(x)\| \leq M \quad$ für alle $x \in J$,

so folgt $\quad \|f(y) - f(x)\| \leq M(y - x) \quad$ für alle $x, y \in J$ mit $y \geq x$.

16.3.2 Beweis (ii) ist eine einfache Folgerung aus (i) für $g(t) = M \cdot t$.

(1) Es seien $x, y \in J$ mit $y > x$ fixiert. Wir beweisen die Ungleichung (*) mit einem ganz speziellen Trick. Statt ihrer beweisen wir für jedes $\varepsilon > 0$ eine Ungleichung mit einem zusätzlichen Spielraum

$$\|f(y) - f(x)\| \leq g(y) - g(x) + \varepsilon(y - x). \qquad (**)$$

Ist nämlich diese Spielraum–Ungleichung $(**)$ für jedes $\varepsilon > 0$ erfüllt, so kann man auf beiden Seiten zum Grenzwert für $\varepsilon \searrow 0$ übergehen und erhält $(*)$.

(2) Es sei also $\varepsilon > 0$ fest. Wir betrachten die Teilmenge

$$\mathcal{S} := \Big\{ u \in [x, y] \;\Big|\; \|f(u) - f(x)\| \leq g(u) - g(x) + \varepsilon(u - x) \Big\}$$

der Stellen u, für die die obige Spielraum–Ungleichung $(**)$ erfüllt ist.

(3) Wir nehmen nun an, dass die Behauptung des Satzes nicht stimmt. Für das Supremum s von $\mathcal{S}$ gilt dann

$$s := \sup \mathcal{S} < y.$$

(4) Wir zeigen, dass s selbst zu $\mathcal{S}$ gehört.
Ist nämlich (s_n) eine Folge in $\mathcal{S}$ mit $\lim_{n\to\infty} s_n = s$, so gilt für alle $n \in \mathbb{N}$

$$\|f(s_n) - f(x)\| \leq g(s_n) - g(x) + \varepsilon(s_n - x)$$

Der Übergang $n \to \infty$ liefert dann

$$\|f(s) - f(x)\| \leq g(s) - g(x) + \varepsilon(s - x).$$

(5) Da die Funktionen f und g in s differenzierbar sind, gibt es zu $\frac{\varepsilon}{2}$ ein $\delta > 0$ mit $s + \delta < y$, so dass für alle $u \in [s, s + \delta[$ die Ungleichungen

$$\begin{aligned} \|f(u) - f(s) - f'(s)(u - s)\| &\leq \tfrac{\varepsilon}{2} \cdot \|u - s\| \\ |g(u) - g(s) - g'(s)(u - s)| &\leq \tfrac{\varepsilon}{2} \cdot (u - s) \end{aligned}$$

gelten, woraus die Abschätzungen:

$$\begin{aligned} \|f(u) - f(s)\| &\leq (\|f'(s)\| + \tfrac{\varepsilon}{2}) \cdot \|u - s\| \leq (\|f'(s)\| + \tfrac{\varepsilon}{2}) \cdot (u - s) \\ g'(s) \cdot (u - s) &\leq g(u) - g(s) + \tfrac{\varepsilon}{2}(u - s) \end{aligned}$$

folgen.

(6) Für alle $u \in [s, s+\delta[$ folgt weiter:

$$\begin{aligned}
\|f(u) - f(x)\| &\leq \|f(u) - f(s)\| + \|f(s) - f(x)\| \\
&\leq (\|f'(s)\| + \tfrac{\varepsilon}{2}) \cdot (u-s) + g(s) - g(x) + \varepsilon(s-x) \\
&\leq (\ \ g'(s) \ \ + \tfrac{\varepsilon}{2}) \cdot (u-s) + g(s) - g(x) + \varepsilon(s-x) \\
&\leq g(u) - g(s) + \varepsilon \cdot (u-s) + g(s) - g(x) + \varepsilon(s-x) \\
&= g(u) - g(x) + \varepsilon \cdot (u-x).
\end{aligned}$$

Dies zeigt aber, dass das Intervall $[s, s+\delta[$ zu $\mathcal{S}$ gehört im WIDERSPRUCH dazu, dass s das Supremum von $\mathcal{S}$ sein soll.

16.3.3 Definition: konvexe Teilmengen Eine Teilmenge X eines Vektorraums heißt *konvex*, wenn zu je zwei Punkten $x, y \in X$ auch immer die gesamte Strecke zwischen diesen beiden Punkten in X enthalten ist.

$$x, y \in X \quad \Longrightarrow \quad (1-\lambda)x + \lambda y \in X \qquad \text{für alle } \lambda \in [0,1].$$

16.3.4 Satz: Abschätzungssatz (2)

Es sei X eine offene konvexe Teilmenge von $\mathbb{R}^d$ und $f : X \to \mathbb{R}^w$ eine differenzierbare Abbildung. Die Ableitungsfunktion $f' : X \to \mathbb{R}^{w \times d}$ sei in X durch die Konstante M beschränkt:

$$\|f'(x)\| \leq M \ \text{ für alle } x \in X.$$

Dann gilt für alle $x, y \in X$:

$$\|f(y) - f(x)\| \leq M \cdot \|y - x\|.$$

16.3.5 Beweis Es seien $x, y \in X$ festgehalten.

(1) Wir definieren eine Funktion

$$g : \begin{cases} J & \to & X \\ \lambda & \mapsto & g(\lambda) \ := \ (1-\lambda)x + \lambda y, \end{cases}$$

wobei J ein offenes reelles Intervall J mit $[0,1] \subseteq J$ und $g(J) \subseteq X$ ist. Die Existenz von J und g ist dadurch gesichert, dass X offen und konvex ist.

(2) Wir wenden Satz 16.3.4 auf die zusammengesetzte Abbildung $f \circ g$ in

$$J \xrightarrow{g} X \xrightarrow{f} \mathbb{R}^w$$

an. Es gilt dann

$$\begin{aligned}
\|f(y) - f(x)\| &= \|(f \circ g)(1) - (f \circ g)(0)\| \\
&\overset{\text{S 16.3.4}}{\leq} \sup_{\lambda \in J} \{\|(f \circ g)'(\lambda)\|\} \cdot (1-0) \\
&\overset{\text{S 16.2.13}}{=} \sup_{\lambda \in J} \{\| \underbrace{(f'(g(\lambda))}_{w \times d} \cdot \underbrace{g'(\lambda)}_{d \times 1} \|\}
\end{aligned}$$

$$
\begin{aligned}
&= \sup_{\lambda\in J}\{\| \underbrace{(f'(g(\lambda))}_{w\times d} \cdot \underbrace{(y-x)}_{d\times 1} \|\} \\
&\overset{\text{S 15.1.7}}{\leq} \sup_{\lambda\in J}\{\| \underbrace{(f'(g(\lambda))}_{w\times d} \| \cdot \| \underbrace{y-x}_{d\times 1} \|\} \\
&\leq \sup_{z\in X}\{\| \underbrace{(f'(z)}_{w\times d} \|\} \cdot \| \underbrace{y-x}_{d\times 1} \| \\
&\leq M \cdot \|y-x\|.
\end{aligned}
$$

16.4 Mehrfache Differenzierbarkeit

Es sei wieder $X \subseteq \mathbb{R}^d$ offen und $f : X \to \mathbb{R}$ eine reellwertige Funktion, $a \in X$.

16.4.1 Definitionen

(1) Es sei $(k_1, k_2, \ldots, k_{r-1}, k_r)$ eine endliche Folge von Zahlen aus $\{1, \ldots, d\}$. Bei entsprechenden Differenzierbarkeitsvoraussetzungen nennt man die Funktion

$$\frac{\partial^r f}{\partial x_{k_1} \cdots \partial x_{k_r}} = \partial_{k_r} \partial_{k_{r-1}} \ldots \partial_{k_2} \partial_{k_1} f \; : X \to \mathbb{R} \qquad (*)$$

eine *höhere partielle Ableitung*.

(2) Die Funktion $f : X \to \mathbb{R}$ heißt *r–mal differenzierbar*, wenn für beliebige Zahlenfolgen $(k_1, k_2, \ldots, k_r)$ und alle $x \in X$ die zugehörigen partiellen Ableitungen in $(*)$ existieren.

(3) Die Funktion $f : X \to \mathbb{R}$ heißt *r–mal stetig differenzierbar*, wenn die Funktionen in $(*)$ existieren und stetig sind.

Die Menge der r–mal stetig differenzierbaren Funktionen wird mit $\mathcal{C}^r(X, \mathbb{R})$ oder $\mathcal{C}^r(X)$ bezeichnet.

(4) Die Funktion $f : X \to \mathbb{R}$ heißt ∞–oft (stetig) differenzierbar, wenn sie für alle $r \in \mathbb{N}$ r–mal (stetig) differenzierbar ist. Symbolik: $\mathcal{C}^\infty(X, \mathbb{R})$ oder $\mathcal{C}^\infty(X)$.

Für die Theorie sind vor allem die Definitionen (3) und (4) von Bedeutung.

16.4.2 Beispiel Wir betrachten die Funktion

$$f : \begin{cases} \mathbb{R}^2 & \to \quad \mathbb{R} \\ (x_1, x_2) & \mapsto \quad \begin{cases} \frac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2}, & \text{falls } (x_1, x_2) \neq (0,0), \\ 0, & \text{falls } (x_1, x_2) = (0,0) \end{cases} \end{cases}$$

und berechnen zunächst die partiellen Ableitungen nach der ersten Variablen.

Für $x_2 \neq 0$ können wir einfache Differentiationsregeln anwenden:

$$\begin{aligned} \partial_1 f(x_1, x_2) &= \frac{(3x_1^2 x_2 - x_2^3)(x_1^2 + x_2^2) - (x_1^3 x_2 - x_1 x_2^3) 2x_1}{(x_1^2 + x_2^2)^2} \\ &= \frac{x_1^4 x_2 + 4x_1^2 x_2^3 - x_2^5}{(x_1^2 + x_2^2)^2} = \frac{x_2(x_1^4 + 4x_1^2 x_2^2 - x_2^4)}{(x_1^2 + x_2^2)^2} \end{aligned}$$

Für $x_2 = 0$ ist die partielle Funktion $f(\cdot\,, 0) \equiv 0$, daher für $x_1 \in \mathbb{R}$

$$\partial_1 f(x_1, 0) = 0.$$

Wir berechnen dann weiter

$$\partial_2 \partial_1 f((0,0) = \lim_{h \overset{\bullet}{\to} 0} \frac{\partial_1 f(0,h) - \partial_1 f(0,0)}{h} = \lim_{h \overset{\bullet}{\to} 0} \frac{-\frac{h^5}{h^4} - 0}{h} = -1.$$

Genau dual dazu wird berechnet oder mit Hilfe der Beziehung

$$f(u,v) - f(v,u), \qquad \text{für } (u,v) \in \mathbb{R}^2$$

durchdacht, dass

$$\partial_1 \partial_2 f((0,0) = +1.$$

Das heißt also, dass bei Vertauschung der Reihenfolge der partiellen Ableitungen andere Ergebnisse zu erwarten sind. Zur Beruhigung dient der folgende Satz.

16.4.3 Satz: Vertauschung der partiellen Ableitungen

Ist die Funktion $f : X \to \mathbb{R}$ in $\mathcal{C}^r(X)$ enthalten, so gilt für eine beliebige Permutation σ (bijektive Abbildung $\{1, \ldots, r\} \to \{1, \ldots, r\}$) und alle $x \in X$:

$$\partial_{k_r} \partial_{k_{r-1}} \ldots \partial_{k_2} \partial_{k_1} f(x) = \partial_{k_{\sigma(r)}} \partial_{k_{\sigma(r-1)}} \ldots \partial_{k_{\sigma(2)}} \partial_{k_{\sigma(1)}} f(x)$$

Auf die Reihenfolge der partiellen Ableitungen kommt es also nicht an.

16.4.4 Beweis Man kann den Beweis dieses Satzes in mehrerer Hinsicht auf die folgende spezielle Situation reduzieren:

- Man betrachtet die Stelle $a = 0$.
- Es ist $r = 2$: Es wird nur zweimal differenziert.
- Es ist $d = 2$: Die Funktion hängt nur von 2 Variablen ab.

Da man die Voraussetzung des Satzes auch inhaltlich noch abschwächen kann, formulieren wir dies in einem weiteren Satz.

16.4.5 Satz: Partielle Ableitung Es sei $X \subseteq \mathbb{R}^2$ eine offene Umgebung von $0 \in \mathbb{R}^2$ und $f : X \to \mathbb{R}$. Betrachte die folgenden Aussagen:

(A_1) Die partielle Ableitung $\partial_1 f : X \to \mathbb{R}$ existiert.

(A_2) Die partielle Ableitung $\partial_2 f : X \to \mathbb{R}$ existiert.

(B_1) Die partielle Ableitung $\partial_2 \partial_1 f : X \to \mathbb{R}$ existiert und ist stetig.

(B_2) Die partielle Ableitung $\partial_1 \partial_2 f : X \to \mathbb{R}$ existiert und ist stetig.

Wenn (A_1) und (A_2) erfüllt sind, dann sind (B_1) und (B_2) äquivalent und es gilt:

$$\partial_2 \partial_1 f(x) = \partial_1 \partial_2 f(x) \quad \text{für alle } x \in X.$$

16.4.6 Beweis Es seien die Aussagen (A_1), (A_2) und (B_2) erfüllt.

(1) Es sei $a = (a_1, a_2) \in X$ fest gewählt. Eine Transformation

$$\widetilde{f}(x,y) = f(x + a_1, y + a_2) - \partial_2\partial_1 f(a_1, a_2)\, x\, y$$

zeigt, dass man o.B.d.A. $a = 0$ und dann $\partial_2\partial_1 f(0,0) = 0$ annehmen kann.

(2) Wir definieren die Hilfsfunktionen auf X

$$\begin{aligned} g(x,y) &:= f(x,y) - f(x,0) \\ h(x,y) &:= f(x,y) - f(0,y), \end{aligned}$$

die dann die Identität

$$h(x,y) - h(x,0) = g(x,y) - g(0,y)$$

erfüllen.

(3) Es sei $\varepsilon > 0$ vorgegeben. Da $\partial_2\partial_1 f$ stetig ist, gibt es $\delta > 0$ so, dass

$$|\partial_2\partial_1 f(x,y)| \leq \varepsilon \text{ für alle } (x,y) \in U_{2\delta}(0).$$

(4) Es sei jetzt $x \in [-\delta, +\delta]$ fest und $(y_n)_{n\in\mathbb{N}}$ eine Folge mit $0 < |y_n| \leq \delta$ und $\lim_{n\to\infty} y_n = 0$.

(5) Für alle $n \in \mathbb{N}$ gilt dann $(x, y_n) \in U_{2\delta}(0)$ und deshalb

$$\left|\frac{h(x,y_n) - h(x,0)}{y_n}\right| = \left|\frac{g(x,y_n) - g(0,y_n)}{y_n}\right| = \left|\frac{\partial_1 g(\xi, y_n)\, x}{y_n}\right| =$$

$$\left|\frac{\partial_1 f(\xi, y_n) - \partial_1 f(\xi, 0)}{y_n}\, x\right| = |\partial_2\partial_1 f(\xi, \eta_n)\, x| \leq \varepsilon |x|,$$

wobei ξ zwischen 0 und x und η_n zwischen 0 und y_n geeignete Zwischenstellen (gemäß Mittelwertsatz der 1–dimensionalen Differentialrechnung) sind.

(6) Es gilt dann weiter für $x \in [-\delta, +\delta]$

$$|\partial_2 f(x,0) - \partial_2 f(0,0) - 0 \cdot x| = |\partial_2 h(x,0)| = \lim_{n\to\infty} \left|\frac{h(x,y_n) - h(x,0)}{y_n}\right| \leq \varepsilon |x|.$$

Das bedeutet, dass die Funktion $\partial_2 f(\cdot\,, 0)$ in $0 \in \mathbb{R}$ differenzierbar ist mit Ableitung

$$\partial_1\partial_2 f(0,0) \;=\; 0 \;=\; \partial_2\partial_1 f(0,0).$$

(7) Die Rücktransformation bzgl. (1) zeigt dann, dass die zweite partielle Ableitung $\partial_1\partial_2 f(a_1, a_2)$ für alle $(a_1, a_2) \in X$ existiert und mit $\partial_2\partial_1 f(a_1, a_2)$ übereinstimmt.

Dann folgt aber aus der Stetigkeit der Funktion $\partial_2\partial_1 f$ auch die von $\partial_1\partial_2 f$.
Das war zu beweisen.

17 Existenzsätze

17.1 Der Banach'sche Fixpunktsatz

17.1.1 Definition: Kontrahierende Abbildung Es sei $T : M \to M$ eine Abbildung innerhalb eines metrischen Raumes M. Diese Abbildung heißt *kontrahierend*, wenn es eine Zahl $\theta \in [0, 1[$ gibt, so dass

$$d(T(y), T(\widetilde{y})) \leq \theta \cdot d(y, \widetilde{y}) \qquad \text{für alle } y, \widetilde{y} \in M.$$

Es ist klar, dass eine kontrahierende Abbildung stetig ist, man muss für vorgegebenes $\varepsilon > 0$ nur $\delta := \frac{\varepsilon}{\theta}$ wählen. Überlege den Fall $\theta = 0$.

17.1.2 Satz: Banach'scher Fixpunktsatz

Es sei (M, d) ein vollständig metrischer Raum und $T : M \to M$ eine kontrahierende Abbildung.

(i) T besitzt genau einen Fixpunkt $y^* \in M$, d.h. $T(y^*) = y^*$.

(ii) Jede durch Iteration von T mit beliebigem Startpunkt y_0 definierte Folge

$$y_0 \in M \text{ (beliebig)}, \qquad y_{n+1} := T(y_n), \quad n \in \mathbb{N}_0,$$

konvergiert gegen diesen Fixpunkt.

17.1.3 Beweis
(1) **Eindeutigkeit**: Hätte T zwei Fixpunkte y^*, y^{**}, so würde dies bedeuten:

$$d(y^*, y^{**}) = d(T(y^*), T(y^{**})) \leq \theta \cdot d(y^*, y^{**}).$$

Wäre $d(y^*, y^{**}) > 0$, so folgte $1 \leq \theta$ im Widerspruch zur Kontraktionseigenschaft. Also muß $y^* = y^{**}$ sein.

(2) Es sei nun (y_n) eine beliebige wie in (ii) beschriebene Folge.

(2a) Wir zeigen durch Induktion: Der Abstand zweier Nachbarglieder der Folge geht geometrisch gegen Null:

$$d(y_{n+1}, y_n) \leq \theta^n \cdot d(y_1, y_0).$$

Der Induktionsanfang ($n = 0$) ist sonnenklar. Der Induktionsschluss $n \mapsto n + 1$ ist auch nicht schwer:

$$\begin{aligned} d(y_{n+2}, y_{n+1}) &= d(T(y_{n+1}), T(y_n)) \leq \theta \cdot d(y_{n+1}, y_n) \\ &\overset{\text{IndV}}{\leq} \theta \cdot \theta^n \cdot d(y_1, y_0) = \theta^{n+1} \cdot d(y_1, y_0). \end{aligned}$$

(2b) Wir führen eine Abschätzung des Abstands zweier beliebiger Folgenglieder durch. Für $m, n \in \mathbb{N}_0$ (O.B.d.A $m > n$) gilt:

$$\begin{aligned} d(y_m, y_n) &\overset{\text{DEU}}{\leq} d(y_m, y_{m-1}) + d(y_{m-1}, y_{m-2}) + \ldots + d(y_{n+1}, y_n) \\ &\leq (\theta^{m-1} + \theta^{m-2} + \ldots + \theta^n) \cdot d(y_1, y_0) \\ &= \theta^n \cdot (\theta^{m-n-1} + \theta^{m-n-2} + \ldots + 1) \cdot d(y_1, y_0) \\ &\overset{\text{geoRei}}{\leq} \frac{\theta^n}{1-\theta} \cdot d(y_1, y_0) \end{aligned}$$

(2c) Die Folge ist eine Cauchy–Folge. Ist nämlich $\varepsilon > 0$ gegeben, so wähle $N \in \mathbb{N}$, so groß, dass

$$\frac{\theta^N}{1-\theta} \cdot d(y_1, y_0) < \varepsilon.$$

Es folgt dann mit (2b), dass für alle $m > n \geq N$:

$$d(y_m, y_n) \leq \frac{\theta^n}{1-\theta} \cdot d(y_1, y_0) \leq \frac{\theta^N}{1-\theta} \cdot d(y_1, y_0) < \varepsilon.$$

(4) Da M vollständig ist, besitzt die Folge (y_n) einen Grenzwert y'.

(5) Wir zeigen, dass dieser Grenzwert ein Fixpunkt ist. Aufgrund der Stetigkeit von T gilt nämlich:

$$T(y') = T(\lim_{n\to\infty} y_n) = \lim_{n\to\infty} T(y_n) = \lim_{n\to\infty} y_{n+1} = y'.$$

Damit ist die **Existenz** eines Fixpunktes gezeigt.

(7) Wir wissen also jetzt, dass es genau einen Fixpunkt gibt. Die Schritte (2) – (5) zeigen, dass **jede beliebige Iterationsfolge** gegen einen Fixpunkt konvergiert, sie **konvergiert** also **gegen den Fixpunkt**.

17.1.4 Beispiel Es sei

$$T : \begin{cases} \mathbb{R}^d & \to & \mathbb{R}^d \\ y & \mapsto & Ay + b \end{cases}$$

eine affine Abbildung mit der Matrix $A \in \mathbb{R}^{d \times d}$ und dem Vektor $b \in \mathbb{R}^d$. Gilt $\|A\| = \theta < 1$, so folgt mit der Submultiplikativität (Satz 15.1.7) für alle $y, \widetilde{y} \in \mathbb{R}^d$

$$\|T(y) - T(\widetilde{y})\| = \|(Ay + b)) - (A\widetilde{y} + b)\| \leq \|A\| \|y - \widetilde{y}\| = \theta \|y - \widetilde{y}\|.$$

Die Abbildung T hat damit genau einen Fixpunkt, die Gleichung

$$Ay + b = y$$

ist also eindeutig lösbar. Es muss also die Inverse der Matrix $A - I$ existieren, die Lösung ist gegeben durch

$$y = (A - I)^{-1} b.$$

17.1.5 Satz: Anwendung: Newton–Verfahren Es sei $I \subseteq \mathbb{R}$ ein (echtes) offenes Intervall und $f : I \to \mathbb{R}$ zweimal stetig differenzierbar. Es sei $a \in I$ eine Stelle mit

$$f(a) \quad = \quad 0 \quad \text{und} \quad f'(a) \neq 0.$$

Dann gibt es ein $\delta > 0$ und eine wohldefinierte Abbildung

$$T_f : \begin{cases} B_\delta(a) & \to & B_\delta(a) \\ x & \mapsto & x - \frac{f(x)}{f'(x)}, \end{cases}$$

so dass eine durch T_f rekursiv definierte Folge $(x_n)_{n\in\mathbb{N}}$ mit beliebiger Startstelle $x_0 \in B_\delta(a)$ gegen die Nullstelle a konvergiert.

17.1.6 Beweis

(0) Wir betrachten zunächst die Abbildung T_f als mit einer anderen Wertemenge versehen:

$$T_f : \begin{cases} B_\delta(a) & \to & \mathbb{R} \\ x & \mapsto & x - \frac{f(x)}{f'(x)} \end{cases}$$

(1) Fixiere ein $\theta \in \,]0, 1[$ und wähle dann $\delta > 0$, so dass $B_\delta(a) \subseteq I$ und für $x \in B_\delta(a)$

$$f'(x) \neq 0, \qquad \left| \frac{f(x) \cdot f''(x)}{(f'(x))^2} \right| \leq \theta.$$

(2) Dann ist für $x \in B_\delta(a)$

$$|T'_f(x)| = \left| 1 - \frac{f'(x)f'(x) - f(x) \cdot f''(x)}{(f'(x))^2} \right| = \left| \frac{f(x) \cdot f''(x)}{(f'(x))^2} \right| \leq \theta.$$

(3) Mit dem Abschätzungssatz 16.3.4 gilt dann für $x, \widetilde{x} \in B_\delta(a)$

$$|T_f(x) - T_f(\widetilde{x})| \leq \theta \cdot |x - \widetilde{x}|.$$

(4) Es ist weiter

$$|T_f(x) - a| = |T_f(x) - T_f(a)| \leq \theta \cdot |x - a|,$$

woraus folgt, dass $T_f(B_\delta(a)) \subseteq B_\delta(a)$, also die Wertemenge von T_f wie im Satz angegeben ist.

(5) Der Banach'sche Fixpunktsatz liefert dann die Behauptung.

17.2 Parameter–Abhängigkeit des Fixpunkts

17.2.1 Definition: Gleichmäßige Kontraktion, Fixpunktabbildung

Es sei (M, d) ein vollständiger metrischer Raum, X ein weiterer metrischer Raum.

Wir betrachten die Abbildung

$$T : \begin{cases} X \times M & \to & M \\ (x, y) & \mapsto & T(x, y) = T_x(y). \end{cases}$$

(1) T heißt eine *gleichmäßige Kontraktion*, wenn es ein $\theta < 1$ gibt, so dass für alle $x \in X$, $y, \widetilde{y} \in M$

$$d(T(x, y), T(x, \widetilde{y})) \leq \theta \cdot d(y, \widetilde{y}).$$

(2) Ist T eine gleichmäßige Kontraktion, so gibt es gemäß Banach'schem Fixpunktsatz 17.1.2 zu jedem $x \in X$ genau einen Fixpunkt $y^* \in M$ der partiellen Abbildung $T(x, \cdot\,)$:

$$T(x, y^*) = y^*.$$

Auf diese Weise ist also eine Abbildung

$$T^* : \begin{cases} X & \to & M \\ x & \mapsto & y^* \end{cases}$$

definiert, die wir die zu T gehörige *Fixpunktabbildung* nennen.

17.2.2 Satz: Stetigkeit der Fixpunktabbildung

Es sei T wie oben eine gleichmäßige Kontraktion.

(i) Ist die partielle Abbildung $T(\cdot\,, y) : X \to M$ für alle $y \in M$ stetig, so ist auch die Fixpunktabbildung $T^* : X \to M$ stetig.

(ii) Ist die partielle Abbildung $T(\cdot\,, y) : X \to M$ für alle $y \in M$ Lipschitz–stetig, so ist auch die Fixpunktabbildung $T^* : X \to M$ Lipschitz–stetig.

17.2.3 Beweis Wir betrachten eine fest gewählte Stelle $a \in X$. Für beliebiges $x \in X$ gilt dann

$$\begin{aligned} & d(T^*(x), T^*(a)) \\ = \; & d(T(x, T^*(x)), T(a, T^*(a))) \\ \leq \; & d(T(x, T^*(x)), T(x, T^*(a))) + d(T(x, T^*(a)), T(a, T^*(a))) \\ \leq \; & \theta \cdot d(T^*(x), T^*(a)) + d(T(x, T^*(a)), T(a, T^*(a))). \end{aligned}$$

Daraus folgt aber

$$d(T^*(x), T^*(a)) \quad \leq \quad \frac{1}{1 - \theta} \cdot d(T(x, T^*(a)), T(a, T^*(a))).$$

Deshalb überträgt sich die Stetigkeit bzw. Lipschitz–Stetigkeit der Abbildung $T(\cdot\,, T^*(a)) : X \to M$ auf die Abbildung $T^* : X \to M$ wie im Satz angegeben.

17.2.4 Satz: Differenzierbarkeit der Fixpunktabbildung

Es seien offene Teilmengen $X \subseteq \mathbb{R}^d, Y \subseteq \mathbb{R}^w$ gegeben. Die Abbildung

$$T : X \times Y \to \mathbb{R}^w$$

sei stetig differenzierbar. Die Ableitung von T bzgl. der zweiten Variablen erfülle

$$\|D_2T(x,y)\| \leq \theta < 1 \qquad \text{für alle } (x,y) \in X \times Y.$$

(i) Ist Y konvex, so besitzt die partielle Abbildung $T(x,\cdot\,) : Y \to \mathbb{R}^w$ für jedes $x \in X$ höchstens einen Fixpunkt $y^* \in Y$.

(ii) Es sei zusätzlich $M \subseteq Y$ eine abgeschlossene Teilmenge mit $T(X \times M) \subseteq M$. Dann ist die Fixpunktabbildung

$$T^* : \begin{cases} X & \to & M \\ x & \mapsto & y^* \end{cases}$$

stetig differenzierbar, für die Ableitung $T^{*\prime}(x) \in \mathbb{R}^{w \times d}$ an der Stelle x gilt die Gleichung

$$\underbrace{T^{*\prime}(x)}_{w\times d} = \underbrace{D_2T(x,T^*(x))}_{w\times w} \cdot \underbrace{T^{*\prime}(x)}_{w\times d} + \underbrace{D_1T(x,T^*(x))}_{w\times d},$$

die sich durch Differentiation der Fixpunktgleichung

$$T^*(x) = T(x,T^*(x))$$

nach x ergibt.

17.2.5 Beweis

Zu (i): Wenn $T(x,\cdot\,) : Y \to \mathbb{R}^w$ zwei Fixpunkte y^* und y^{**} hätte, so würde gemäß Abschätzungssatz 16.3.4 gelten:

$$\begin{aligned} \|y^* - y^{**}\| &= \|T(x,y^*) - T(x,y^{**})\| \leq \sup_{y\in Y}\{\|D_2T(x,y)\|\} \, \|y^* - y^{**}\| \\ &\leq \theta \cdot \|y^* - y^{**}\|, \end{aligned}$$

was aufgrund von $\theta < 1$ unmöglich ist.

Zu (ii): Wir betrachten eine feste Stelle $a \in X$ und setzen $b := T^*(a)$.

(1) Die Abbildung

$$\begin{cases} \mathbb{R}^{w\times d} & \to & \mathbb{R}^{w\times d} \\ J & \mapsto & \underbrace{D_2T(a,b)}_{w\times w} \cdot J + \underbrace{D_1T(a,b)}_{w\times d} \end{cases}$$

in dem vollständigen metrischen Raum $\mathbb{R}^{w\times d}$ der $w \times d$–Matrizen ist eine Kontraktion, da $\|D_2T(a,b)\| \leq \theta$ und deswegen für $J, \widetilde{J} \in \mathbb{R}^{w\times d}$

$$\begin{aligned} &\left\| \, [D_2T(a,b) \cdot J + D_1T(a,b)] - \left[D_2T(a,b) \cdot \widetilde{J} + D_1T(a,b)\right] \right\| = \\ &\left\| D_2T(a,b) \cdot (J - \widetilde{J}) \right\| \leq \left\| D_2T(a,b) \right\| \cdot \left\| J - \widetilde{J} \right\| \leq \theta \|J - \widetilde{J}\|. \end{aligned}$$

Sie hat also genau einen Fixpunkt in $\mathbb{R}^{w\times d}$, den wir J^* nennen.

(2) Es sei jetzt ein $\varepsilon > 0$ vorgegeben. Aufgrund von Satz 16.2.9(D) gibt es eine Umgebung U' von a in X und eine Konstante $L > 0$, so dass für $x \in U'$, $y := T^*(x)$

$$\|y - b\| = \|T^*(x) - T^*(a)\| \leq L \cdot \|x - a\|$$

und dann eine weitere Umgebung $U'' \subseteq U'$ von a, so dass für alle $x \in U''$, $y := T^*(x)$

$$\begin{aligned}
\|T(x,y) - T(a,y) - D_1T(a,y)(x-a)\| &\leq \frac{(1-\theta)\varepsilon}{3} \cdot \|x-a\| \\
\|T(a,y) - T(a,b) - D_2T(a,b) \cdot (y-b)\| &\leq \frac{(1-\theta)\varepsilon}{3L} \cdot \|y-b\| \\
&\leq \frac{(1-\theta)\varepsilon}{3} \cdot \|x-a\| \\
\|D_1T(a,y) - D_1T(a,b)\| &\leq \frac{(1-\theta)\varepsilon}{3}.
\end{aligned}$$

(3) Für beliebiges $x \in U''$, $y := T^*(x)$ gilt dann

$$\begin{aligned}
& \|y - b - J^* \cdot (x-a)\| \\
= \; & \|T(x,y) - T(a,b) - [D_2T(a,b) \cdot J^* + D_1T(a,b)](x-a)\| \\
\leq \; & \|T(x,y) - T(a,y) - D_1T(a,y)](x-a)\| + \\
& \|T(a,y) - T(a,b) - D_2T(a,b) \cdot (y-b)\| + \\
& \underbrace{\|D_2T(a,b)\|}_{\leq\theta} \cdot \|y - b - J^* \cdot (x-a)\| + \\
& \|D_1T(a,y) - D_1T(a,b)\| \cdot \|x-a\| \\
\leq \; & \theta \cdot \|y - b - J^* \cdot (x-a)\| + \\
& \frac{(1-\theta)\varepsilon}{3} \cdot \|x-a\| + \frac{(1-\theta)\varepsilon}{3} \cdot \|x-a\| + \frac{(1-\theta)\varepsilon}{3} \cdot \|x-a\| \\
= \; & \theta \cdot \|y - b - J^* \cdot (x-a)\| + (1-\theta)\varepsilon \cdot \|x-a\|.
\end{aligned}$$

Eine Umstellung der Gesamt–Ungleichung ergibt

$$\|y - b - J^* \cdot (x-a)\| \leq \varepsilon \cdot \|x-a\|$$

Rück–ersetzt man die Abkürzungen $y = T^*(x)$, $b = T^*(a)$, so zeigt dies, dass für alle $x \in U''$

$$\|T^*(x) - T^*(a) - J^* \cdot (x-a)\| = \|y - b - J^* \cdot (x-a)\| \leq \varepsilon \cdot \|x-a\|,$$

also ist die Abbildung T^* differenzierbar in a mit Ableitung $T^{*\prime}(a) = J^*$, die als Fixpunkt der Abbildung (1) die im Satz angegebene Gleichung erfüllt.

17.2.6 Satz: Störung einer invertierbarer Matrix

(i) Es sei $A \in \mathbb{R}^{\ell\times\ell}$ eine invertierbare Matrix. Ist $B \in \mathbb{R}^{\ell\times\ell}$ eine Matrix mit

$$\|B - A\| < \frac{1}{\|A^{-1}\|},$$

so ist auch B invertierbar.

(ii) Die Teilmenge $\mathbf{GL}(\ell,\mathbb{R}) \subseteq \mathbb{R}^{\ell\times\ell}$ der invertierbaren Matrizen ist offen in der Menge aller Matrizen.

(iii) Die Abbildung

$$\begin{cases} \mathbf{GL}(\ell,\mathbb{R}) & \to & \mathbf{GL}(\ell,\mathbb{R}) \\ B & \mapsto & B^{-1} \end{cases}$$

ist stetig differenzierbar.

17.2.7 Beweis Wähle ein $\theta < 1$ und definiere dann X als die offene Kugel

$$X := \left\{ x \in \mathbb{R}^{\ell\times\ell} \,\middle|\, \|x - A\| < \frac{\theta}{\|A^{-1}\|} \right\}.$$

Wir definieren im vollständig metrischen Raum $\mathbb{R}^{\ell\times\ell}$ die Abbildung

$$T : \begin{cases} X \times \mathbb{R}^{\ell\times\ell} & \to & \mathbb{R}^{\ell\times\ell} \\ (x,y) & \mapsto & y - A^{-1}(xy - I) \end{cases}$$

Für festes x hat sie genau dann y als Fixpunkt, wenn y invers zu x ist.

Es bleibt also nur zu zeigen, dass T eine gleichmäßige Kontraktion ist. Tatsächlich ist für $x \in X$

$$\begin{aligned} & \|T(x,y) - T(x,\widetilde{y})\| \\ = \;& \|[y - A^{-1}(xy - I)] - [\widetilde{y} - A^{-1}(x\widetilde{y} - I)]\| \\ = \;& \|(I - A^{-1}x)(y - \widetilde{y})\| \\ = \;& \|A^{-1}(A - x)(y - \widetilde{y})\| \\ \leq \;& \|A^{-1}\| \cdot \|A - x\| \cdot \|y - \widetilde{y}\| \\ \leq \;& \theta \cdot \|y - \widetilde{y}\|. \end{aligned}$$

Die Aussagen des Satzes folgen nun mit Satz 17.2.4, man setzt dort

$$M = Y = \mathbb{R}^{\ell\times\ell} \cong \mathbb{R}^d = \mathbb{R}^w.$$

17.2.8 Bemerkungen Die beiden Aussagen (ii) und (iii) können auch mit Hilfe der Stetigkeit der Determinantenfunktion $\det : \mathbb{R}^{\ell\times\ell} \to \mathbb{R}$ und der Cramer'schen Formel für die Inverse einer Matrix bewiesen werden. Bemerkenswert an dem obigen Beweis ist, dass er ganz ohne Rückgriff auf eine Basis von $\mathbb{R}^{\ell\times\ell}$, d.h. auf die Einträge der Matrizen, auskommt.

Der obige Beweis kann auch ohne Rückgriff auf den Banach'schen Fixpunktsatz, beispielsweise mit Hilfe der sogenannten *von–Neumann–Reihe* geführt werden.

17.3 Der Satz über implizite Funktionen

Einstieg: Wir betrachten ein Höhenprofil (beliebige Funktion)

$$h : \begin{cases} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & h(x,y). \end{cases}$$

Der Graph dieser Funktion bildet eine ,,2–dimensionale Mannigfaltigkeit"

$$\{(x,y,z)|(x,y) \in \mathbb{R}^2, z = h(x,y)\}$$

im dreidimensionalen Raum. Es sei $C \in \mathbb{R}$ eine Konstante. Wir stellen uns die Frage, ob wir die ,,Höhe–gleich–C–Linie" als Graph $G_f = \{(x, f(x))|x \in \mathbb{R}\}$ einer Funktion $f : \mathbb{R} \to \mathbb{R}$ darstellen können.

Anders ausgedrückt: Kann man die Gleichung $h(x,y) = C$ nach y auflösen?

17.3.1 Beispiele

1) Es seien α, β Konstanten und

$$h(x,y) = \alpha x + \beta y.$$

Dieses Höhenprofil ist eine Ebene durch den Ursprung $(0,0)$. Oberhalb der x–Achse (d.h. $y = 0$) bildet diese Ebene die Gerade $z = \alpha x$ aus, oberhalb der y–Achse ($x = 0$) ist sie durch Gerade $z = \beta y$ gegeben.

Für $\beta \neq 0$ ist die Linie $h(x,y) = C$ gegeben durch

$$y = \frac{C - \alpha x}{\beta},$$

es handelt sich — erwartungsgemäß — um eine Gerade.

Im Fall $\beta = 0, \alpha \neq 0$ ist die Linie $h(x,y) = C$ eine Gerade parallel zur y–Achse. Sie ist nicht als Graph einer Funktion mit der ersten Koordinate x als abhängiger Variabler darstellbar.

Für $\alpha = \beta = 0$ ist die ,,Höhenlinie keine Linie" mehr, die obige Frage kann so gar nicht gestellt werden.

(2) Als Höhenprofil sei nun durch

$$h(x,y) = x^2 + y^2$$

ein Paraboloid gegeben, es ist vorstellbar als Spur einer um die z–Achse rotierenden Normalparabel. Es ist klar, dass für $C > 0$ die durch $x^2 + y^2 = C$ definierte Höhenlinie eine Kreislinie darstellt. Bei der Ermittlung der Funktion $y = f(x)$ treten unschöne Begleiterscheinungen auf. Die ,,Auflösung"

$$f : \begin{cases} [-\sqrt{C}, +\sqrt{C}] & \to & \mathbb{R} \\ x & \mapsto & \pm\sqrt{C - x^2} \end{cases}$$

ist doppeldeutig, sie ist an den Rändern der Definitionsmenge nicht differenzierbar.

Es stellt sich die Frage, unter welchen Bedingungen an die Funktion h eine „schöne" implizite Funktion f definiert wird.

17.3.2 Satz über implizite Funktionen (ImFT)

Es seien $X \subseteq \mathbb{R}^d$ und $Y \subseteq \mathbb{R}^w$ offen, $h : X \times Y \to \mathbb{R}^w$ eine stetig differenzierbare Abbildung.

Wir betrachten eine feste Stelle $(a, b) \in X \times Y$ mit den folgenden Eigenschaften

- Es ist $h(a, b) = 0$ (Nullstelle).
- Es ist $D_2h(a, b) \in \mathbb{R}^{w \times w}$ invertierbar.

(i) Dann existieren offene Mengen $U \subseteq X, V \subseteq Y$ mit $(a, b) \in U \times V$ und eine stetig differenzierbare Abbildung

$$f : \begin{cases} U & \to & V \\ x & \mapsto & f(x), \end{cases}$$

so dass

$$h(x, y) = 0 \qquad \Longleftrightarrow \qquad y = f(x) \qquad \text{für alle } x \in U.$$

Man sagt dann auch:

- Durch die Gleichung $h(x, y) = 0$ ist (lokal: in einer Umgebung U von a) die Funktion $y = f(x)$ *implizit definiert.*
- Die Gleichung $h(x, y) = 0$ kann (lokal: in einer Umgebung U von a) durch die Funktion $f : U \to V$ nach y aufgelöst werden.

(ii) Für die Ableitung von f gilt:

$$\underbrace{f'(x)}_{w\times d} = -\underbrace{[D_2h(x, f(x))]^{-1}}_{w\times w} \cdot \underbrace{D_1h(x, f(x))}_{w\times d}, \qquad x \in U.$$

17.3.3 Beweis

Wir kürzen ab $J := D_2h(a, b)$.

(1) Da die Abbildung $D_2h(a, \cdot\,) : Y \to \mathbb{R}^{w\times w}$ stetig ist, gibt es eine offene Umgebung $\widetilde{U}$ von a in X und eine offene konvexe Umgebung V von b in Y, so dass

$$\|D_2h(x, y) - J\| = \|D_2h(x, y) - D_2h(a, b)\| \le \frac{1}{2\,\|J^{-1}\|}, \qquad \text{falls } (x, y) \in \widetilde{U} \times V.$$

(2) Der Beweis beruht auf einem Wechselspiel zwischen den beiden Funktionen

$$h|_{\widetilde{U}\times V} : \begin{cases} \widetilde{U} \times V & \to & \mathbb{R}^w \\ (x, y) & \mapsto & h(x, y), \end{cases} \qquad F : \begin{cases} \widetilde{U} \times V & \to & \mathbb{R}^w \\ (x, y) & \mapsto & y - J^{-1}h(x, y). \end{cases}$$

Für $(x,y) \in \widetilde{U} \times V$ besteht die Äquivalenz $h(x,y) = 0 \iff F(x,y) = y$, die wie folgt gelesen werden sollte:

$$(*) \quad \begin{cases} \text{Für festes } x \in \widetilde{U} \text{ ist } y \in V \text{ genau dann Lösung der Gleichung } h(x,y) = 0, \\ \text{wenn } y \text{ ein Fixpunkt der Abbildung } F(x,\cdot\,) : V \to \mathbb{R}^w \text{ ist.} \end{cases}$$

(3) Wir gelangen zu der Abschätzung für $(x,y) \in \widetilde{U} \times V$

$$\|D_2F(x,y)\| = \|I - J^{-1}D_2h(x,y)\| \leq \|J^{-1}\| \cdot \|J - D_2h(x,y)\| \stackrel{(1)}{\leq} \frac{1}{2} \qquad (**)$$

Für festes $x \in \widetilde{U}$ hat gemäß Satz 17.2.4(i) die partielle Funktion $F(x,\cdot\,)$ höchstens einen Fixpunkt in V, dieser ist aufgrund der Äquivalenz $(*)$ die eindeutige Lösung der Gleichung $h(x,y) = 0$.

(4) Wir definieren nun U als die Menge derjenigen $x \in \widetilde{U}$, für die die Gleichung $h(x,y) = 0$ tatsächlich eine Lösung hat. Damit ist die Funktion

$$f : \begin{cases} U & \to & V \\ x & \mapsto & y, \text{ die eindeutige Lösung in } V \text{ von } h(x,y) = 0 \end{cases}$$

wohldefiniert.

(5) Es bleibt noch zu zeigen, dass U offen und $f : U \to V$ stetig differenzierbar ist.

Dazu sei $(p,q) \in U \times V$ mit $h(p,q) = 0$ (äquivalent: $q = f(p)$) fixiert.

Da V offen ist, gibt es ein $\varepsilon > 0$ mit

$$q \in B_\varepsilon(q) \subseteq V.$$

Es gibt dann weiter ein $\delta > 0$, so dass $U_\delta(p) \subseteq U$ und für $x \in U_\delta(p)$

$$\|h(x,q)\| = \|h(x,q) - \underbrace{h(p,q)}_{=0}\| \leq \frac{\varepsilon}{2\,\|J^{-1}\|}.$$

(6) Wir überzeugen uns jetzt davon, dass die Abbildung

$$F : \begin{cases} U_\delta(p) \times V & \to & \mathbb{R}^w \\ (x,y) & \mapsto & y + J^{-1}h(x,y), \end{cases}$$

mit $M := B_\varepsilon(q) \subseteq V$ den Bedingungen von Satz 17.2.4(ii) genügt: Zunächst ist F stetig differenzierbar und es gilt die obige Bedingung $(**)$.

Für ein beliebiges Paar $(x,y) \in U_\delta(p) \times B_\varepsilon(q)$ gilt

$$\begin{aligned} \|F(x,y) - q\| &\leq \|F(x,y) - F(x,q)\| + \|F(x,q) - q\| \\ &\stackrel{\text{AbS 16.3.4,}(**)}{\leq} \frac{1}{2} \cdot \|y - q\| + \|J^{-1}h(x,q)\| \\ &\leq \frac{1}{2} \cdot \underbrace{\|y - q\|}_{\leq \varepsilon} + \|J^{-1}\| \cdot \underbrace{\|h(x,q)\|}_{\leq \frac{\varepsilon}{2\,\|J^{-1}\|}} \\ &\leq \varepsilon, \end{aligned}$$

also $F(x,y) \in B_\varepsilon(q)$. Insgesamt ist also $F(U_\delta(p) \times B_\varepsilon(q)) \subseteq B_\varepsilon(q)$.

(7) Gemäß Satz 17.2.4(ii) existiert eine stetig differenzierbare Fixpunktabbildung

$$F^* : \begin{cases} U_\delta(p) & \to & B_\varepsilon(q) \\ x & \mapsto & F^*(x), \end{cases}$$

die aufgrund der Äquivalenz $(*)$ gerade mit der auf $U_\delta(p)$ eingeschränkten Abbildung f übereinstimmt. Damit wissen wir, dass f in p differenzierbar ist. Außerdem sehen wir, dass $U_\delta(p) \subseteq U$, also U offen, ist.

(8) Die Aussage (ii) ergibt sich durch Differentiation der Gleichung $h(x, f(x)) = 0$, $x \in U$:

$$D_1 h(x, f(x)) + D_2 h(x, f(x)) \cdot f'(x) = 0$$

und anschließende Auflösung nah $f'(x)$. Beachte dabei, dass die Invertierbarkeit von $D_2 h(x, f(x))$ durch Satz 17.2.6(i) und die Überlegung in Schritt (1) gesichert ist.

17.4 Der Satz über inverse Funktionen

17.4.1 Satz über inverse Funktionen (InFT)

Es sei $X \subseteq \mathbb{R}^d$ offen und $f : X \to \mathbb{R}^d$ eine stetig differenzierbare Abbildung.

An der Stelle $a \in X$ sei die Ableitung $f'(a) \in \mathbb{R}^{d\times d}$ invertierbar.

(i) Dann existieren offene Mengen U, V mit

$$a \in U \subseteq X, \qquad b := f(a) \in V \subseteq \mathbb{R}^d$$

und eine stetig differenzierbare Abbildung $g : V \to U$, die invers zu der auf U eingeschränkten Abbildung $f|_U$ ist.

(ii) Für $(x, y) \in U \times V$, $y = f(x)$ gilt:

$$f'(x) \cdot g'(y) = I.$$

17.4.2 Beweis Um die Situation dieses Satzes InFT besser in die des Satzes ImFT über implizite Funktionen einpassen zu können, ,,spiegeln" wir die Notation und formulieren erneut:

Es sei $Y \subseteq \mathbb{R}^d$ offen und $g : Y \to \mathbb{R}^d$ eine stetig differenzierbare Abbildung

$$\mathbb{R}^d \xleftarrow{g} Y$$

An der Stelle $b \in Y$ sei die Ableitung $g'(b) \in \mathbb{R}^{d\times d}$ invertierbar.

Dann existieren offene Mengen U, V mit

$$g(b) \in U \subseteq \mathbb{R}^d, \qquad b \in V \subseteq Y$$

und eine stetig differenzierbare Abbildung $f : U \to V$, die invers zu der auf V eingeschränkten Abbildung $g|_V$ ist.

Wenn wir jetzt

$$X = \mathbb{R}^d = \mathbb{R}^w$$

$$a := g(b)$$

$$h : \begin{cases} \mathbb{R}^d \times Y & \to & \mathbb{R}^d \\ (x, y) & \mapsto & x - g(y) \end{cases}$$

setzen, dann sind die Voraussetzungen des ImFT erfüllt:

$$\begin{aligned} h(a, b) &= a - g(b) = 0 \\ D_2 h(a, b) &= -g'(b) \quad \text{invertierbar.} \end{aligned}$$

Es existieren also U, V offen und eine stetig differenzierbare Funktion $f : U \to V$ mit

$$0 = h(x, y) = x - g(y) \quad \Longleftrightarrow \quad y = f(x), \qquad x \in U,$$

also ist f invers zu $g|_V$. Die Behauptung über die Ableitungen ist eine einfache Konsequenz der Kettenregel.

17.4.3 Satz Sei $\theta \in [0,1[$ und $Y \subseteq \mathbb{R}^d$ offen und konvex.

Weiter sei $f : Y \to \mathbb{R}^d$ eine differenzierbare Funktion mit

$$|f'(y)| \quad \leq \quad \theta \quad \text{für} \quad y \in Y.$$

Weiter sei eine der Aussagen

(A) $f(Y) \subseteq Y$

(B) Es existiert (mindestens) ein Fixpunkt $y^{**} \in Y$.

Dann konvergiert jede durch f iterierte Folge gegen den (dann eindeutigen) Fixpunkt y^*.

17.4.4 Beweis Aufgrund des Abschätzungssatzes 16.3.4 gilt für alle $y, \widetilde{y} \in Y$:

$$\|f(y) - f(\widetilde{y})\| = \sup_{z \in Y}\{\|f'(z)\|\} \cdot \|y - \widetilde{y}\| \leq \theta \cdot \|y - \widetilde{y}\|.$$

und unter der Voraussetzung (A) folgt das Ergebnis direkt mit dem Banach'schen Fixpunktsatz.

Die zweite Voraussetzung (B) impliziert (A). Ist nämlich $y \in Y$, so gilt

$$|f(y) - y^{**}| = |f(y) - f(y^{**})| \leq \theta \cdot |y - y^{**}|,$$

es folgt $f(y) \in Y$.

17.4.5 Korollar Die Funktion $f : I \to \mathbb{R}$ habe einen Fixpunkt $a \in I$ und sei in a stetig differenzierbar mit $|f'(a)| < 1$. Dann gibt es eine Umgebung U von a, so dass bei einem Start in U die durch f iterierte Folge gegen a konvergiert.

17.4.6 Beweis Wähle ein $\theta \in \mathbb{R}$ mit $|f'(a)| < \theta < 1$ und dann eine abgeschlossene Umgebung Y von a, so dass

$$|f'(x)| \leq \theta \qquad \text{für alle} \quad x \in Y.$$

Es sind dann die Voraussetzungen von Satz 17.4.3 erfüllt und damit folgt die Behauptung.

17.4.7 Hinweis Graphische Umsetzung dieser Idee. Bezug zum Cobweb-Theorem

17.5 Extrema unter Nebenbedingungen

17.5.1 Definition: Extrema Es sei X Teilmenge von $\mathbb{R}^d$ (oder eines allgemeineren metrischen Raumes) und $f : X \to \mathbb{R}$ eine reellwertige Funktion, $a \in X$.

(1) Der Punkt $(a, f(a))$ des Graphen heißt wie in der ersten Spalte angegeben, wenn es ein $\varepsilon > 0$ gibt, so dass die Bedingung der zweiten Spalte gilt:

lokales Maximum	$f(x) \leq f(a)$ für alle $x \in D$ mit $\lvert x-a\rvert < \varepsilon$
lokales Minimum	$f(x) \geq f(a)$ für alle $x \in D$ mit $\lvert x-a\rvert < \varepsilon$
strenges lokales Maximum	$f(x) < f(a)$ für alle $x \in D$ mit $0 < \lvert x-a\rvert < \varepsilon$
strenges lokales Minimum	$f(x) > f(a)$ für alle $x \in D$ mit $0 < \lvert x-a\rvert < \varepsilon$

(2) Der Punkt $(a, f(a))$ des Graphen heißt wie in der ersten Spalte angegeben, wenn die Bedingung der zweiten Spalte erfüllt ist:

globales Maximum	$f(x) \leq f(a)$ für alle $x \in D$
globales Minimum	$f(x) \geq f(a)$ für alle $x \in D$
strenges globales Maximum	$f(x) < f(a)$ für alle $x \in D \setminus \{a\}$
strenges globales Minimum	$f(x) > f(a)$ für alle $x \in D \setminus \{a\}$

(3) Man spricht jeweils von einem *Extremum*, wenn es sich um ein Maximum oder Minimum handelt.

(4) In diesem Zusammenhang heißt a auch *Stelle* des Maximums, Minimums oder Extremums und $f(a)$ der *Wert* des Maximums, Minimums oder Extremums.

17.5.2 Satz: Notwendige Bedingung für Extremum Es sei $a \in X$ ein innerer Punkt. Ist die Funktion $f : X \to \mathbb{R}$ in a differenzierbar und hat sie dort ein Extremum, so gilt

$$f'(a) = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}.$$

17.5.3 Beweis Es sei $j \in \{1, \ldots, d\}$ fest gewählt. Da a ein innerer Punkt ist, gibt es ein $\delta > 0$, so dass die partielle Funktion

$$f(a_1, \ldots, a_{j-1}, \underset{\underset{j}{\uparrow}}{\cdot}, a_{j+1}, \ldots, a_d) : \begin{cases}]a_j - \delta, a_j + \delta[, & \to \mathbb{R} \\ x & \mapsto f(a_1, \ldots, a_{j-1}, x, a_{j+1}, \ldots, a_d) \end{cases}$$

wohl–definiert ist. Sie hat ein Extremum bei a_j, deshalb ist aufgrund von Satz 9.7.2(i) $\partial_j f(a) = 0$. Da dies für alle $j \in \{1, \ldots, d\}$ gilt, ist aufgrund von Satz **??**

$$f'(a) = \begin{pmatrix} \partial_1 f(a) & \partial_2 f(a) & \cdots & \partial_d f(a) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}.$$

17.5.4 Satz: Extrema unter Nebenbedingungen / Spezialfall

Es sei $X \subseteq \mathbb{R}^d$ offen, die Funktion $f : X \to \mathbb{R}$ sei stetig differenzierbar.

Es sei $h : X \to \mathbb{R}$ eine weitere stetig differenzierbare Funktion, die über

$$X_N := \{x \in X | h(x) = 0\}$$

(N für Nullniveau oder Nebenbedingung) eine Teilmenge von X definiert. Auf X_N ist dann die Einschränkung von f

$$f_N : \begin{cases} X_N & \to & \mathbb{R} \\ x & \mapsto & f(x) \end{cases}$$

definiert.

Es seien für eine Stelle $a \in X_N$ die folgenden Aussagen gegeben:

- f_N habe (als Funktion auf X_N) in a ein lokales Extremum.
- Es ist $h'(a) \neq 0$.

Dann gibt es ein $\lambda \in \mathbb{R}$, so dass

$$f'(a) = \lambda\, h'(a),$$

d.h. die beiden Zeilenvektoren sind linear abhängig.

17.5.5 Bemerkungen

1. Dieser Satz kann auf eine höhere Zahl von Nebenbedingungen verallgemeinert werden.
2. Die Zahl λ heißt ein *Lagrange'scher Multiplikator*.
3. Der Sachverhalt, dass die Funktion $f_N : X_N \to \mathbb{R}$ in a ein Extremum aufweist, wird auch dadurch beschrieben, dass f in a ein *Extremum unter der Nebenbedingung* $h = 0$ hat.

17.5.6 Beweis (1) Wegen $h'(a) \neq 0$ ist eine der partiellen Ableitungen $\partial_k h(a) \neq 0$. Wir nehmen o.B.d.A. an, dass $k = d$.

(2) Wir wählen offene Mengen $\widetilde{X}$, $\widetilde{Y}$ mit

$$\widetilde{a} := (a_1, \ldots, a_{d-1}) \in \widetilde{X} \subseteq \mathbb{R}^{d-1}, \qquad \widetilde{b} := a_d \in \widetilde{Y} \subseteq \mathbb{R}, \qquad \widetilde{X} \times \widetilde{Y} \subseteq X$$

und wenden den Satz 17.3.2 über implizite Funktionen auf die eingeschränkte Funktion

$$\widetilde{h} : \begin{cases} \widetilde{X} \times \widetilde{Y} & \to & \mathbb{R} \\ (\widetilde{x}, \widetilde{y}) & \mapsto & h(\widetilde{x}, \widetilde{y}) \end{cases}$$

an der Stelle $a = (\widetilde{a}, \widetilde{b})$ an. Die beiden Voraussetzungen aus Satz 17.3.2 sind erfüllt, da

$$h(\widetilde{a}, \widetilde{b}) = h(a) = 0, \qquad D_2 h(\widetilde{a}, \widetilde{b}) = \partial_d h(a) \neq 0 \quad (\Rightarrow \ \text{invertierbar}).$$

Es existieren also offene Mengen U, V mit

$$(\widetilde{a}, \widetilde{b}) \in U \times V \subseteq \widetilde{X} \times \widetilde{Y} \subseteq X$$

und eine stetig differenzierbare Abbildung $g : U \to V$, so dass

$$h(\widetilde{x}, \widetilde{y}) = 0 \quad \iff \quad \widetilde{y} = g(\widetilde{x}) \quad \text{für alle } \widetilde{x} \in U.$$

Wir notieren aus Satz 17.3.2(ii) den Ausdruck für die Ableitung im Punkt $\widetilde{a} \in U$

$$g'(\widetilde{a}) = -[D_2 h(\widetilde{a}, g(\widetilde{a}))]^{-1} D_1 h(\widetilde{a}, g(\widetilde{a})).$$

Mit Hilfe dieser Funktion g können wir die eingeschränkte Definitionsmenge X_N von f_N umparametrisieren:

$$F : \begin{cases} U & \to \quad \mathbb{R} \\ \widetilde{x} & \mapsto \quad f_N(\widetilde{x}, g(\widetilde{x})) = f(\widetilde{x}, g(\widetilde{x})). \end{cases}$$

Da die Funktion f_N bei a ein Extremum besitzt, hat F im Punkt $\widetilde{a}$ eines. Die Ableitung von F in $\widetilde{a}$ muss Null sein:

$$\begin{aligned} 0 = \underbrace{F'(\widetilde{a})}_{1\times(d-1)} &= \underbrace{D_1 f(\widetilde{a}, g(\widetilde{a}))}_{1\times(d-1)} + \underbrace{D_2 f(\widetilde{a}, g(\widetilde{a}))}_{1\times 1} \underbrace{g'(\widetilde{a})}_{1\times(d-1)} \\ &= D_1 f(\widetilde{a}, g(\widetilde{a})) - \underbrace{D_2 f(\widetilde{a}, g(\widetilde{a})) \cdot [D_2 h(\widetilde{a}, g(\widetilde{a}))]^{-1}}_{=:\lambda} D_1 h(\widetilde{a}, g(\widetilde{a})). \end{aligned}$$

Geeignet zusammengefasst bedeutet dies aber:

$$\begin{aligned} D_1 f(\widetilde{a}, g(\widetilde{a})) &= \lambda \cdot D_1 h(\widetilde{a}, g(\widetilde{a})) \\ D_2 f(\widetilde{a}, g(\widetilde{a})) &= \lambda \cdot D_2 h(\widetilde{a}, g(\widetilde{a})) \end{aligned}$$

und deshalb — unter Beachtung von $a = (\widetilde{a}, g(\widetilde{a}))$ —

$$f'(a) = \begin{pmatrix} D_1 f(a) & D_2 f(a) \end{pmatrix} = \lambda \cdot \begin{pmatrix} D_1 h(a) & D_2 h(a) \end{pmatrix} = \lambda \cdot h'(a).$$

17.5.7 Satz: Extrema unter Nebenbedingungen / Verallgemeinerung

Es sei $Z \subseteq \mathbb{R}^d$ offen, die Funktion $f : Z \to \mathbb{R}$ sei stetig differenzierbar.

Es sei $h : Z \to \mathbb{R}^w$, $w \leq d$, eine weitere stetig differenzierbare Funktion, die über

$$Z_N := \{x \in Z | h(x) = 0\}$$

(N für Null– oder Nebenbedingung) eine Teilmenge von Z definiert. Die Funktion f kann auf Z_N eingeschränkt werden:

$$f_N : \begin{cases} Z_N & \to & \mathbb{R} \\ x & \mapsto & f(x). \end{cases}$$

Es seien für eine Stelle $c \in Z_N$ die folgenden Aussagen erfüllt:

- f_N habe (als Funktion auf Z_N) in c ein lokales Extremum.
- Die Matrix $h'(c) \in \mathbb{R}^{w \times d}$ hat Vollrang w, d.h. es gibt w linear unabhängige Spaltenvektoren.

Dann gibt es einen Zeilenvektor $\lambda = \begin{pmatrix} \lambda_1 & \cdots & \lambda_w \end{pmatrix} \in \mathbb{R}^w$, so dass

$$\underbrace{f'(c)}_{1 \times d} = \underbrace{\lambda}_{1 \times w} \cdot \underbrace{h'(c)}_{w \times d}.$$

Anders ausgedrückt: Die Ableitung von f in c ist eine Linearkombination der Ableitungen der Koordinatenfunktionen von h in c.

17.5.8 Bemerkungen

1. Man veranschauliche sich diesen Satz zunächst einmal für den Fall $w = 1$.
2. Die Zahlen $\lambda_1, \dots, \lambda_w$ heißen *Lagrange'sche Multiplikatoren.*
3. Der Sachverhalt, dass die Funktion $f_N : Z_N \to \mathbb{R}$ in c ein Extremum aufweist, wird auch dadurch beschrieben, dass f in c ein *Extremum unter der Nebenbedingung* $h = 0$ hat.
4. Anstelle von einer Nebenbedingung $h = 0$ spricht man oft von den w Nebenbedingungen:

$$\begin{aligned} h_1(x) &= 0 \\ &\vdots \\ h_w(x) &= 0. \end{aligned}$$

17.5.9 Beweis

(1) Wir können o.B.d.A. annehmen, dass $c \in X \times Y = Z$ mit offenen Mengen $X \subseteq \mathbb{R}^{d-w}$, $Y \subseteq \mathbb{R}^w$. Im folgenden beziehen sich die Ableitungssymbole D_1 und D_2 auf diese beiden Mengen.

(2) Zusätzlich können wir aufgrund der Vollrangbedingung annehmen, dass der Anteil $D_2h(c)$ in

$$\underbrace{Dh(c)}_{w\times d} = \begin{pmatrix} \underbrace{D_1h(c)}_{w\times(d-w)} & \underbrace{D_2h(c)}_{w\times w} \end{pmatrix}$$

invertierbar ist. Anderenfalls nummerieren wir die Variablen in $X \times Y$ um.

(3) Wir wenden den Satz 17.3.2 über implizite Funktionen auf die Funktion $h : X \times Y \to \mathbb{R}^w$ an der Stelle $c = (a, b)$ an.

Es existieren also offene Mengen U, V mit

$$(a, b) \in U \times V \subseteq X \times Y \subseteq Z$$

und eine stetig differenzierbare Abbildung $g : U \to V$, so dass

$$h(x, y) = 0 \quad \iff \quad y = g(x) \quad \text{für alle } x \in U.$$

Wir notieren aus Satz 17.3.2(ii) den Ausdruck für die Ableitung im Punkt $a \in U$

$$\underbrace{g'(a)}_{w\times(d-w)} = -\underbrace{[D_2h(a, g(a))]^{-1}}_{w\times w} \cdot \underbrace{D_1h(a, g(a))}_{w\times(d-w)}. \qquad (*)$$

Mit Hilfe dieser Funktion g können wir die Definitionsmenge Z_N der Funktion f_N in einer Umgebung von a ,,mit weniger Koordinaten aus U schreiben", es entsteht die Funktion

$$F : \begin{cases} U & \to & \mathbb{R} \\ x & \mapsto & f_N(x, g(x)) = f(x, g(x)). \end{cases}$$

Da die Funktion f_N bei c ein Extremum besitzt, hat F im Punkt a eines. Die Ableitung von F in a muss Null sein:

$$\begin{aligned} 0 = \underbrace{F'(a)}_{1\times(d-w)} &= \underbrace{D_1f(a, g(a))}_{1\times(d-w)} + \underbrace{D_2f(a, g(a))}_{1\times w}\ \underbrace{g'(a)}_{w\times(d-w)} \\ &\overset{(*)}{=} \underbrace{D_1f(a, g(a))}_{1\times(d-w)} - \overbrace{\underbrace{D_2f(a, g(a))}_{1\times w} \cdot \underbrace{[D_2h(a, g(a))]^{-1}}_{w\times w}}^{=:\lambda} \cdot \underbrace{D_1h(a, g(a))}_{w\times(d-w)}. \end{aligned}$$

Geeignet zusammengefasst bedeutet dies aber:

$$\begin{aligned} D_1f(a, g(a)) &= \lambda \cdot D_1h(a, g(a)) \\ D_2f(a, g(a)) &= \lambda \cdot D_2h(a, g(a)) \end{aligned}$$

und deshalb — unter Beachtung von $c = (a, g(a))$ —

$$f'(c) = \begin{pmatrix} D_1f(c) & D_2f(c) \end{pmatrix} = \lambda \cdot \begin{pmatrix} D_1h(c) & D_2h(c) \end{pmatrix} = \lambda \cdot h'(c).$$

Inhaltsverzeichnis