Mitschrift zur Analysis II Vorlesung von Prof. Dr. Wittbold im SS 08

Thomas El Khatib

2. August 2008

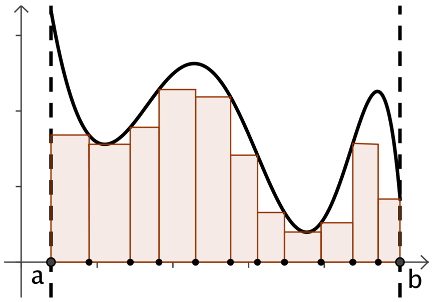
Inhaltsverzeichnis

5 Integration

5.1 Das bestimmte Riemann-Integral

5.1.1 Motivation

Wie berechnet man die Fläche, die von der x-Achse, den Graphen $x = a$, $x = b$ und dem Graphen einer Funktion $f : [a, b] \to \mathbb{R}$ begrenzt wird?



Eine naheliegende Idee ist, die Fläche durch einfacher zu berechnende Flächen, zum Beispiel Rechtecke, anzunähern. Die Genauigkeit der Näherung steigt anschaulich, je feiner man den betrachteten Bereich der x-Achse unterteilt.

5.1.2 Zerlegung, Unter- und Obersummen

Definition 5.1.1. Sei $I = [a, b]$, $a, b \in \mathbb{R}$, $f : I \to \mathbb{R}$ beschränkt.

a) $Z = (x_0, x_1, \ldots, x_n)$ mit $a = x_0 < x_1 < \ldots < x_n = b$, $n \in \mathbb{N}$, heißt **Zerlegung von** I in Teilintervalle $I_k := [x_{k-1}, x_k]$ für $k \in \{1, \ldots, n\}$.

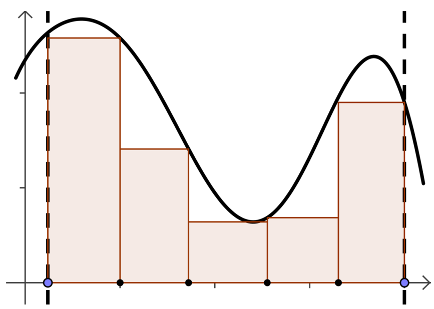
   $|Z| := \max\{|I_k| \mid k \in \{1, \ldots, n\}\}$, wobei $|I_k| := x_k - x_{k-1}$ heißt **Feinheit der Zerlegung** Z.

b) Seien

$$m_k := \inf f(I_k) = \inf_{x \in I_k} f(x) = \inf\{f(x) \mid x \in I_k\},$$
$$M_k := \sup f(I_k) = \sup_{x \in I_k} f(x) = \sup\{f(x) \mid x \in I_k\}$$

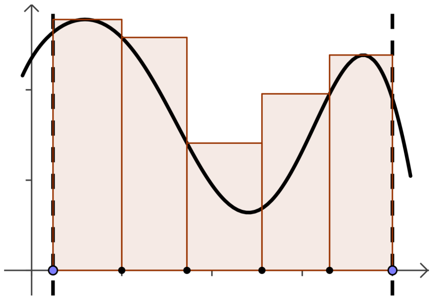
   für $k \in \{1, \ldots, n\}$, so definiert man die Unter- bzw. Obersummen von f bezüglich der Zerlegung Z:

   – **Untersumme**: $s(Z, f) := \sum_{k=1}^{n} m_k \cdot |I_k|$



und

- **Obersumme**: $S(Z,f) := \sum\limits_{k=1}^{n} M_k \cdot |I_k|$.



c) Seien Z, Z_1, Z_2 Zerlegungen von I.

   i) Z_1 heißt **Verfeinerung** von Z, wenn Z_1 alle Zwischenpunkte der Zerlegung Z enthält.

   ii) Die Zerlegung Z, die genau die Zwischenpunkte von Z_1 und Z_2 enthält, heißt **Überlagerung** von Z_1 und Z_2, geschrieben: $Z = Z_1 + Z_2$.

5.1.3 Überlagerungslemma

Lemma 5.1.2. Sei $f : [a,b] \to \mathbb{R}$ beschränkt, d.h. es gilt $|f(x)| \leq K$ für alle $x \in [a,b]$. Sei Z_1 eine Zerlegung von I mit p im Inneren von I liegenden Punkten ($p \in \mathbb{N}$). Dann gilt für jede Zerlegung Z von I:

$$\begin{aligned} s(Z,f) \leq s(Z+Z_1,f) \leq s(Z,f) + 2pK\,|Z| \\ S(Z,f) \geq S(Z+Z_1,f) \geq S(Z,f) - 2pK\,|Z| \end{aligned}$$

Insbesondere gilt: Bei Verfeinerungen nehmen Untersummen zu und Obersummen ab.

Beweis. für Untersummen und $p = 1$, d.h. $Z_1 = (a,z,b)$, und $Z = (x_0,\dots,x_n)$.

1. Fall: $z \in \{x_0,\dots,x_n\}$ ist trivial.

2. Fall: $z \in \,]x_{k-1},x_k[$ für ein $k \in \{1,\dots,n\}$. Dann gilt

$$\begin{aligned} s(Z,f) &= \sum_{i=1}^{n} m_k \cdot (x_i - x_{i-1}) \\ &= \underbrace{\sum_{i=1, i\neq k}^{n} m_k \cdot (x_i - x_{i-1})}_{\Sigma} + \inf f([x_{k-1},x_k]) \cdot (x_k - x_{k-1}) \\ &= \Sigma + \inf f([x_{k-1},x_k]) \cdot (x_k - z + z - x_{k-1}) \\ &= \Sigma + \inf f([x_{k-1},x_k]) \cdot (x_k - z) + \inf f([x_{k-1},x_k]) \cdot (z - x_{k-1}) \\ &\leq \Sigma + \inf f([z,x_k]) \cdot (x_k - z) + \inf f([x_{k-1},z]) \cdot (z - x_{k-1}) \\ &= s(Z+Z_1,f) \end{aligned}$$

Weiterhin hat man

$$
\begin{aligned}
s(Z+Z_1,f) &\leq \sum_{i=1,i\neq k}^{n} m_k\cdot(x_i-x_{i-1}) + K\cdot(x_k-z) + K\cdot(z-x_{k-1})\\
&= \sum_{i=1,i\neq k}^{n} m_k\cdot(x_i-x_{i-1}) + K\cdot(x_k-x_{k-1})\\
&\leq \sum_{i=1,i\neq k}^{n} m_k\cdot(x_i-x_{i-1}) + K\cdot|Z|\\
&= \underbrace{\sum_{i=1}^{n} m_k\cdot(x_i-x_{i-1})}_{s(Z,f)} - \underbrace{\inf f([x_{k-1},x_k])}_{\geq -K}\cdot\underbrace{(x_k-x_{k-1})}_{\leq |Z|} + K\cdot|Z|\\
&\leq s(Z,f) + 2K\cdot|Z|
\end{aligned}
$$

Per Induktion folgt die Behauptung für beliebige $p \in \mathbb{N}$. □

Korollar 5.1.3. Jede Obersumme ist größer als jede Untersumme, d.h. es gilt $S(Z,f) \geq s(\tilde{Z},f)$ für alle Zerlegungen $Z, \tilde{Z}$.

Beweis. Seien $Z, \tilde{Z}$ Zerlegungen, dann folgt nach dem Lemma

$$s(\tilde{Z},f) \leq s(Z+\tilde{Z},f) \leq S(Z+\tilde{Z},f) \leq S(Z,f)$$

□

5.1.4 Unteres- und oberes Riemann-Integral

Definition 5.1.4. Sei $f:\ [a,b] \to \mathbb{R}$ beschränkt.

a)
$$J_*(f) := \sup_{Z \text{ Zerlegung von } [a,b]} s(Z,f)$$

heißt **unteres Riemann-Integral** von f über $[a,b]$. Entsprechend ist

$$J^*(f) := \inf_{Z \text{ Zerlegung von } [a,b]} S(Z,f)$$

das **obere Riemann-Integral** von f über $[a,b]$. Klar ist: $J_*(f) \leq J^*(f)$.

b) Wenn $J^*(f) = J_*(f)$, dann heißt f über $[a,b]$ **Riemann-integrierbar** und in diesem Fall definiert man:

$$\int_a^b f(x)\,\mathrm{d}x = J_*(f) = J^*(f)$$

c) $R([a,b]) := \{f:\ [a,b] \to \mathbb{R} \mid f \text{ R-integrierbar}\}$

5.1.5 Riemann-Integral als Grenzwert

Satz 5.1.5. Sei $f : [a, b] \to \mathbb{R}$ beschränkt.

Für jede Zerlegungsnullfolge, d.h. für jede Folge von Zerlegungen $(Z_n)_{n\in\mathbb{N}}$ mit $|Z_n| \overset{n\to\infty}{\longrightarrow} 0$, gilt:

$$\lim_{n\to\infty} s(Z_n, f) = J_*(f) \quad \text{und} \quad \lim_{n\to\infty} S(Z_n, f) = J^*(f)$$

Insbesondere gilt für jede Funktion $f \in R([a, b])$

$$\lim_{n\to\infty} s(Z_n, f) = \lim_{n\to\infty} S(Z_n, f) = \int\limits_a^b f(x)\,\mathrm{d}x$$

Beweis. für Untersummen:

Sei $(Z_n)_{n\in\mathbb{N}}$ eine beliebige Zerlegungsnullfolge; sei weiter Z eine beliebige Zerlegung von $[a, b]$ mit p inneren Zwischenpunkten. Nach Lemma 5.1.2 gilt

$$s(Z_n, f) \le s(Z + Z_n, f) \le s(Z_n, f) + 2pK\,|Z_n| \quad \forall\, n \in \mathbb{N} \quad (*)$$

wobei $K \ge 0$, sodass $|f(x)| \le K \quad \forall\, x \in [a, b]$.

Angenommen, $s(Z_n, f) \overset{n\to\infty}{\longrightarrow} \alpha \in \mathbb{R}$, dann folgt aus $(*)$, dass auch $s(Z + Z_n, f) \overset{n\to\infty}{\longrightarrow} \alpha$. Es ist aber

$$s(Z, f) \le s(Z + Z_n, f) \le J_*(f) \quad \forall\, n \in \mathbb{N}$$

Somit folgt

$$s(Z, f) \le \alpha \le J_*(f)$$

für jede Zerlegung Z von $[a, b]$. Folglich ist auch

$$\sup_{Z \text{ Zerlegung}} s(Z, f) \le \alpha \le J_*(f)$$

also $\alpha = J_*(f)$.

Der allgemeiner Fall folgt nun mit Hilfe des aus Analysis I bekannten Teilfolgenprinzips: jede Teilfolge von $(s(Z_n, f))_{n\in\mathbb{N}}$ besitzt eine weitere Teilfolge $(s(Z_{n_k}, f))_{k\in\mathbb{N}}$, die konvergiert (da $(s(Z_n, f))_{n\in\mathbb{N}}$ beschränkt ist). Nach dem 1. Schritt ist aber

$$\lim_{n\to\infty} s(Z_{n_k}, f) = J_*(f)$$

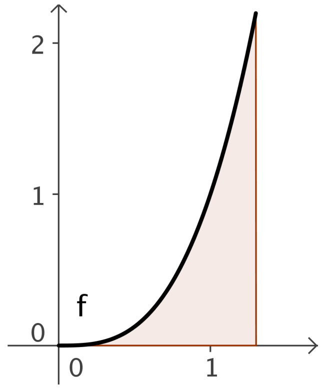
Somit gilt

$$\lim_{n\to\infty} s(Z_n, f) = J_*(f)$$

□

5.1.6 Beispiel 1

$$\begin{aligned} f : [0, b] &\to \mathbb{R} \\ x &\mapsto x^3 \end{aligned}$$



Sei $(Z_n)_{n\in\mathbb{N}}$ äquidistante die Zerlegungsfolge

$$x_0 = 0,\ x_1 = \frac{b}{n}, \dots, \frac{kb}{n}, \dots, x_n = b$$

Dann ist

$$\begin{aligned}
s(Z_n, f) &= \sum_{k=1}^{n} \inf f([x_{k-1}, x_k]) \cdot \frac{b}{n} \\
&= \sum_{k=1}^{n} \left(\frac{k-1}{n} b\right)^3 \cdot \frac{b}{n} \\
&= \frac{b^4}{n^4} \sum_{k=1}^{n} (k-1)^3 \\
&= \frac{b^4}{n^4} \sum_{k=0}^{n-1} k^3 \\
&= \frac{b^4}{n^4} \left(\frac{n(n-1)}{2}\right)^2 \xrightarrow{n\to\infty} \frac{b^4}{4}
\end{aligned}$$

Also ist $J_*(f) = \frac{b^4}{4}$. Analog gilt

$$\begin{aligned}
S(Z_n, f) &= \sum_{k=1}^{n} \sup f([x_{k-1}, x_k]) \cdot \frac{b}{n} \\
&= \sum_{k=1}^{n} \left(\frac{k}{n} b\right)^3 \cdot \frac{b}{n} \\
&= \frac{b^4}{n^4} \sum_{k=1}^{n} k^3 \\
&= \frac{b^4}{n^4} \left(\frac{(n+1)n}{2}\right)^2 \xrightarrow{n\to\infty} \frac{b^4}{4}
\end{aligned}$$

und damit $J^*(f) = \frac{b^4}{4}$. f ist also R-integrierbar mit

$$\int_0^b x^3 \,\mathrm{d}x = \frac{b^4}{4}$$

5.1.7 Beispiel 2

Dirichlet-Funktion:

$$\begin{aligned} D:\ [a,b] &\to \mathbb{R} \\ x &\mapsto \begin{cases} 0, & x \in \mathbb{R}\setminus\mathbb{Q} \\ 1, & x \in \mathbb{Q} \end{cases} \end{aligned}$$

Dann gilt für jede beliebige Zerlegung Z von $[a,b]$

$$s(Z,D) = \sum_{k=1}^{n} \underbrace{\inf D([x_{k-1},x_k])}_{=0}(x_{k-1}-x_k) = 0$$

$$S(Z,D) = \sum_{k=1}^{n} \underbrace{\sup D([x_{k-1},x_k])}_{=1}(x_{k-1}-x_k) = b-a$$

Also gilt $J_*(D) = 0$ und $J^*(D) = b-a$, d.h. $D \notin R([a,b])$.

5.1.8 Riemann'sches Integrabilitätskriterium

Satz 5.1.6. Sei $f : [a,b] \to \mathbb{R}$ beschränkt. Dann gilt: f ist genau dann Riemann-integrierbar, wenn

$$\forall \epsilon > 0 \quad \exists Z \text{ Zerlegung von } [a,b] \ : \ S(Z,f) - s(Z,f) < \epsilon$$

Beweis.

$\Rightarrow$: Ist $f \in R([a,b])$, so gilt

$$\sup_{Z \text{ Zerlegung}} s(Z,f) = \inf_{Z \text{ Zerlegung}} S(Z,f)$$

Sei nun $\epsilon > 0$. Dann existiert eine Zerlegung Z_1 von $[a,b]$ mit

$$s(Z_1,f) \geq J_*(f) - \frac{\epsilon}{2}$$

Weiter existiert eine Zerlegung Z_2 von $[a,b]$, sodass

$$S(Z_2,f) \leq J^*(f) + \frac{\epsilon}{2}$$

Es ist aber

$$s(Z_1+Z_2,f) \geq s(Z_1,f) \geq J_*(f) - \frac{\epsilon}{2}$$

$$S(Z_1+Z_2,f) \leq S(Z_2,f) \leq J^*(f) + \frac{\epsilon}{2}$$

Addiert man die Ungleichungen folgt mit $J_*(f) = J^*(f)$

$$S(Z_1+Z_2,f) + J_*(f) - \frac{\epsilon}{2} \leq J^*(f) + \frac{\epsilon}{2} + s(Z_1+Z_2,f) \qquad \Leftrightarrow$$

$$S(Z_1+Z_2,f) - s(Z_1+Z_2,f) \leq \epsilon$$

Mit $Z = Z_1 + Z_2$ folgt also die Behauptung.

$\Leftarrow$: Sei $\epsilon > 0$. Dann existiert eine Zerlegung $\tilde{Z}$ mit

$$S(\tilde{Z}, f) \leq s(\tilde{Z}, f) + \epsilon$$

Somit folgt

$$\inf_{Z \text{ Zerlegung}} S(Z, f) \leq S(\tilde{Z}, f) \leq s(\tilde{Z}, f) + \epsilon \leq \sup_{Z \text{ Zerlegung}} s(Z, f) + \epsilon$$

d.h. es gilt $J^*(f) \leq J_*(f) + \epsilon$ für alle $\epsilon > 0$. Wegen $J^*(f) \geq J_*(f)$ folgt $J^*(f) = J_*(f)$.

$\square$

5.1.9 Klassen R-integrierbarer Funktionen

Satz 5.1.7. Seien $a, b \in \mathbb{R}$, $a < b$.

i) Jede auf $[a, b]$ stetige Funktion ist R-integrierbar.

ii) Jede auf $[a, b]$ beschränkte und bis auf endlich viele Stellen stetige Funktion ist R-integrierbar.

iii) Jede auf $[a, b]$ monotone Funktion ist R-integrierbar.

Beweis.

i) Sei f eine stetige Funktion auf $[a, b]$. Dann ist f nach Kenntnissen aus Analysis I gleichmäßig stetig auf $[a, b]$, d.h.

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x, y \in [a, b] \ : \ |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$$

Sei $\epsilon > 0$, dann wähle $\delta > 0$ dementsprechend. Sei nun $Z = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$ mit $|Z| < \delta$. Dann ist

$$\begin{aligned} S(Z, f) - s(Z, f) &= \sum_{k=1}^{n} \underbrace{(\sup f([x_{k-1}, x_k]) - \inf f([x_{k-1}, x_k]))}_{<\epsilon} \cdot (x_k - x_{k-1}) \\ &< \epsilon \cdot \sum_{k=1}^{n} (x_k - x_{k-1}) \\ &< \epsilon \cdot (b - a) \end{aligned}$$

d.h. das Riemann'sche Integrabilitätskriterium ist erfüllt.

ii) Sei f wie in ii) gegeben, seien $y_1, \dots, y_p$ die Unstetigkeitsstellen von f und $K \geq 0$, sodass $|f(x)| < K \quad \forall x \in [a, b]$.

Seien $(I_k')_{k=1}^{p}$ Intervalle um die Unstetigkeitsstellen mit

$$\sum_{k=1}^{p} |I_k'| < \epsilon$$

Dann gilt

$$\begin{aligned}\sum_{k=1}^{p}\sup f(I_k')\cdot|I_k'|-\sum_{k=1}^{p}\inf f(I_k')\cdot|I_k'| &\leq \sum_{k=1}^{p}\left(|\sup f(I_k')|+|\inf f(I_k')|\right)\cdot|I_k'|\\ &\leq 2K\cdot\sum_{k=1}^{p}|I_k'|<2K\epsilon\end{aligned}$$

Auf den restlichen abgeschlossenen Teilintervallen $(I_k'')_k$ ist f gleichmäßig stetig. Nach i) existiert deshalb eine Zerlegung Z der Intervall $(I_k'')_k$, sodass

$$S(Z_{(I_k'')},f)-s(Z_{(I_k'')},f)<\epsilon$$

Somit ergibt sich eine Zerlegung $\tilde{Z}$ des gesamten Intervalls $[a,b]$, für die gilt

$$S(\tilde{Z},f)-s(\tilde{Z},f)<(2K+1)\epsilon$$

iii) Sei f monoton wachsend auf $[a,b]$. Dann ist

$$f(a)\leq f(x)\leq f(b)\quad\forall\, x\in[a,b]$$

Sei $Z=(x_0,\ldots,x_n)$ eine Zerlegung von $[a,b]$. Dann gilt

$$\begin{aligned}S(Z,f)-s(Z,f) &= \sum_{k=1}^{n}\left(\sup f([x_{k-1},x_k])-\inf f([x_{k-1},x_k])\right)\cdot(x_k-x_{k-1})\\ &= \sum_{k=1}^{n}\left(f(x_k)-f(x_{k-1})\right)\cdot(x_k-x_{k-1})\\ &\leq |Z|\sum_{k=1}^{n}\left(f(x_k)-f(x_{k-1})\right)\\ &= |Z|\left(f(b)-f(a)\right)\end{aligned}$$

Zu $\epsilon>0$ wähle nun eine Zerlegung Z mit $|Z|<\frac{\epsilon}{f(b)-f(a)}$, dann folgt nach Riemann'schen Integrabilitätskriterium die Behauptung.

□

Wie man im Beweis von ii) sieht, ist eine beschränkte Funktion sogar dann R-integrierbar, wenn ihre Unstetigkeitsstellen für jedes $\epsilon>0$ in höchstens abzählbar vielen Intervallen $(I_j)_{j\in J}$ (J also höchstens abzählbare Indexmenge) mit

$$\sum_{j\in J}|I_j|<\epsilon$$

enthalten sind (oder auch: „von diesen überdeckt werden können“). Das ist die Aussage des ***Lebesgue'schen Integrabilitätskriteriums****.*

5.1.10 Riemann'sche Zwischensumme

Definition 5.1.8. Sei $Z = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$, $\xi_k \in [x_{k-1}, x_k]$ für $k = 1, \dots, n$ und $f : [a, b] \to \mathbb{R}$ beschränkt. Dann heißt

$$\sigma(Z, \xi_k, f) := \sum_{k=1}^{n} f(\xi_k) \cdot (x_k - x_{k-1})$$

eine **Riemann'sche Zwischensumme** von f bezüglich der Zerlegung Z.

Klar ist:

$$s(Z, f) \leq \sigma(Z, \xi_k, f) \leq S(Z, f)$$

Lemma 5.1.9. Sei $f : [a, b] \to \mathbb{R}$ beschränkt. Zu jedem $\epsilon > 0$ und zu jeder Zerlegung Z von $[a, b]$ existieren Zwischenpunkte $(\xi_k)_{k\in\mathbb{N}}$ und $(\eta_k)_{k\in\mathbb{N}}$, sodass

$$\begin{aligned} s(Z, f) \leq \sigma(Z, \xi_k, f) \leq s(Z, f) + \epsilon \\ S(Z, f) \geq \sigma(Z, \eta_k, f) \geq S(Z, f) - \epsilon \end{aligned}$$

Beweis. für Untersummen: Sei $Z = (x_0, \dots, x_n)$ eine Zerlegung. Nach Definition des Infimums existiert zu jedem $\epsilon > 0$ ein $\xi_k \in [x_{k-1}, x_k]$ mit

$$\begin{aligned} &0 \leq f(\xi_k) - \inf f([x_{k-1}, x_k]) \leq \frac{\epsilon}{b-a} && \Rightarrow \\ &0 \leq (f(\xi_k) - \inf f([x_{k-1}, x_k]))\,(x_k - x_{k-1}) \leq \frac{\epsilon}{b-a}(x_k - x_{k-1}) \end{aligned}$$

Über die $k = 1, \dots, n$ summiert ergibt sich

$$\begin{aligned} &0 \leq \sum_{k=1}^{n} (f(\xi_k) - \inf f([x_{k-1}, x_k]))\,(x_k - x_{k-1}) \leq \sum_{k=1}^{n} \frac{\epsilon}{b-a}(x_k - x_{k-1}) && \Rightarrow \\ &0 \leq \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) - \sum_{k=1}^{n} \inf f([x_{k-1}, x_k])(x_k - x_{k-1}) \leq \epsilon && \Rightarrow \end{aligned}$$

$s(Z, f) \leq \sigma(Z, \xi_k, f) \leq s(Z, f) + \epsilon$

□

5.1.11 Riemann'sche Summenfolge

Definition 5.1.10. Sei $(Z_n)_{n\in\mathbb{N}}$ eine Zerlegungsnullfolge auf dem Intervall $[a, b]$, $(\xi_k^n)_{n,k}$ Zwischenpunkte und $f : [a, b] \to \mathbb{R}$ beschränkt. Dann heißt $(\sigma(Z_n, \xi_k^n, f))_n$ eine **Riemann'sche Summenfolge**.

Satz 5.1.11. Sei $f : [a, b] \to \mathbb{R}$ beschränkt. Dann gilt:

$\boxed{f \in R([a, b]) \iff \text{Jede Riemann'sche Summenfolge ist konvergent.}}$

In diesem Fall konvergieren alle Riemann'schen Summenfolgen gegen ein und denselben Grenzwert $\int_a^b f(x)\,\mathrm{d}x$.

Beweis.

$\Rightarrow$: Klar, da für jede Zerlegung Z, für beliebige Zwischenpunkte $(\xi_k)_{k\in\mathbb{N}}$ gilt:

$$s(Z, f) \leq \sigma(Z, \xi_k, f) \leq S(Z, f)$$

$\Leftarrow$: Sei $(Z_n)_{n\in\mathbb{N}}$ eine Zerlegungsnullfolge. Nach Lemma 5.1.9 angewandt mit $Z := Z_n$, $\epsilon = \frac{1}{n}$ existieren zu jedem $n \in \mathbb{N}$ Zwischenpunkte (ξ_k^n) und (η_k^n), sodass

$$s(Z_n, f) \leq \sigma(Z_n, \xi_k^n, f) \leq s(Z_n, f) + \frac{1}{n}$$
$$S(Z_n, f) \geq \sigma(Z_n, \eta_k^n, f) \geq S(Z_n, f) - \frac{1}{n}$$

Für $n \to \infty$ folgt nach Sandwich-Lemma

$$\lim_{n\to\infty} \sigma(Z_n, \xi_k^n, f) = J_*(f)$$
$$\lim_{n\to\infty} \sigma(Z_n, \eta_k^n, f) = J^*(f)$$

Betrachte die „Mischfolge“

$$\sigma(Z_n, \xi_k^1, f),\ \sigma(Z_n, \eta_k^1, f),\ \sigma(Z_n, \xi_k^2, f),\ \ldots$$

Dies ist wieder eine Riemann'sche Summenfolge. Nach Voraussetzung ist diese konvergent. Damit konvergiert auch jede ihrer Teilfolgen gegen den gleichen Grenzwert. Insbesondere konvergieren damit die beiden Teilfolgen $\sigma(Z_n, \xi_k^n, f)$ und $\sigma(Z_n, \xi_k^n, f)$ gegen denselben Grenzwert. Folglich ist $J_*(f) = J^*(f)$, d.h. aber $f \in R([a, b])$.

□

5.1.12 Beispiel

Sei $0 < a < b$. Betrachte $f : [a, b] \to \mathbb{R}$, $x \mapsto x^\alpha$ mit $\alpha \neq -1$. Berechne $\int\limits_a^b f(x)\,\mathrm{d}x$ (da f stetig ist, gilt $f \in R([a, b])$).

Betrachte dazu $q_n := \sqrt[n]{\frac{b}{a}}$. Zerlege das Intervall mittels geometrischer Progression:

$$a = x_0,\ x_1 = q_n,\ x_2 = aq_n^2,\ \ldots,\ x_n = aq_n^n = b$$

Dann gilt für jedes k

$$x_{k+1} - x_k = aq_n^k(q_n - 1) \overset{n\to\infty}{\longrightarrow} 0$$

Setze $\xi_k^n := aq_n^{k-1}$, dann ist

$$\begin{aligned}
\sigma(Z_n, \xi_k^n, f) &= \sum_{k=0}^{n-1} f(\xi_k^n) \cdot aq_n^k(q_n - 1) \\
&= a(q_n - 1) \sum_{k=0}^{n-1} (aq_n^k)^\alpha \cdot q_n^k \\
&= a^{\alpha+1}(q_n - 1) \sum_{k=0}^{n-1} (q_n^{\alpha+1})^k \\
&= a^{\alpha+1}(q_n - 1) \frac{1 - (q_n^{\alpha+1})^n}{1 - q_n^{\alpha+1}} \\
&= a^{\alpha+1} \cdot \left(1 - \left(\frac{b}{a}\right)^{\alpha+1}\right) \frac{q_n - 1}{1 - q_n^{\alpha+1}}
\end{aligned}$$

Nun ist

$$\lim_{x\to 1} \frac{x-1}{1-x^{\alpha+1}} \overset{H}{=} \lim_{x\to 1} \frac{1}{-(\alpha+1)x^{\alpha}} = \frac{-1}{\alpha}$$

und damit

$$\int_a^b x^{\alpha}\,\mathrm{d}x = \lim_{n\to\infty} \sigma(Z_n, \xi_k^n, f) = \left(a^{\alpha+1} - b^{\alpha+1}\right) \frac{-1}{\alpha} = \frac{1}{\alpha} \cdot \left(b^{\alpha-1} - a^{\alpha+1}\right)$$

5.1.13 Cauchy'sches Integrabilitätskriterium

Der folgende Satz wurde in der Vorlesung an dieser Stelle nicht behandelt, wird aber später in allgemeinerem Kontext wieder aufgegriffen.

Satz. Sei $f : [a,b] \to \mathbb{R}$ beschränkt. f ist Riemann-integrierbar genau dann, wenn

$$\forall\, \epsilon > 0 \quad \exists\, \delta > 0 : \left|\sigma(Z, \xi, f) - \sigma(\tilde{Z}, \tilde{\xi}, f)\right| < \epsilon \quad \forall\, Z, \tilde{Z} \text{ Zerlegungen, } |Z|, \left|\tilde{Z}\right| < \delta$$

Beweis.

$\Rightarrow$ Sei also $f \in R([a,b])$ und $\epsilon > 0$ gegeben. Angenommen es gäbe zu jedem $n \in \mathbb{N}$ Zerlegungen $Z_n, \tilde{Z}_n$ mit $|Z_n|, \left|\tilde{Z}_n\right| < \frac{1}{n}$, aber

$$\left|\sigma(Z_n, \xi^n, f) - \sigma(\tilde{Z}_n, \tilde{\xi}^n, f)\right| \geq \epsilon$$

für gewisse Zwischensummenfolgen $(\xi^n)_{n\in\mathbb{N}}, \left(\tilde{\xi}^n\right)_{n\in\mathbb{N}}$. Die Mischfolge

$$\left(\sigma(Z_1, \xi^1, f), \sigma(\tilde{Z}_1, \tilde{\xi}^1, f), \sigma(Z_2, \xi^2, f), \ldots\right)$$

ist eine Riemann'sche Summenfolge, die nach dem letzten Satz konvergiert. Da $\mathbb{R}$ vollständig ist, ist diese Folge eine Cauchy-Folge, d.h. insbesondere müsste

$$\left|\sigma(Z_n, \xi^n, f) - \sigma(\tilde{Z}_n, \tilde{\xi}^n, f)\right| < \epsilon$$

ab einem gewissen $n_0 \in \mathbb{N}$ gelten, was der Annahme widerspricht. Also muss das Cauchy-Integrabilitätskriterium erfüllt sein.

$\Leftarrow$ Sei $\epsilon > 0$ gegeben. Sei $\sigma(Z_n, \xi^n, f)$ eine Riemann'sche Summenfolge. Nach Voraussetzung gibt es ein $\delta > 0$, sodass

$$|\sigma(Z_n, \xi^n, f) - \sigma(Z_m, \xi^m, f)| < \epsilon$$

wenn nur $|Z_n|, |Z_m| < \delta$. Da $(Z_n)_{n\in\mathbb{N}}$ eine Zerlegungsnullfolge ist, muss dies ab einem $n_0 \in \mathbb{N}$ gelten. Damit ist gezeigt, dass $\sigma(Z_n, \xi^n, f)$ eine Cauchy-Folge ist, also konvergiert. Da die Folge beliebig war, folgt aus dem letzten Satz die Integrabilität von f.

$\square$

5.1.14 Allgemeinere Alternativdefinition des Riemann-Integrals

Bemerkung. Satz 5.1.11 zeigt eine Möglichkeit auf, das Riemann-Integral bzw. die Riemann-Integrierbarkeit für nicht notwendig reellwertige Funktionen zu definieren, da die zur Definition nicht notwendig die Ordnungsstruktur von $\mathbb{R}$ ausgenutzt werden muss.

5.1.15 Integrierbarkeit komplexwertiger Funktionen

Beispiel: Komplexwertige Funktionen: Sei $f : [a, b] \to \mathbb{C}$ beschränkt. Dann kann man f in Real- und Imaginärteil zerlegen:

$$\begin{aligned} \mathrm{Re}(f) : [a, b] &\to \mathbb{R} \\ x &\mapsto \mathrm{Re}(f(x)) \end{aligned} \qquad \begin{aligned} \mathrm{Im}(f) : [a, b] &\to \mathbb{R} \\ x &\mapsto \mathrm{Im}(f(x)) \end{aligned}$$

Es ist also: $f = \mathrm{Re}(f) + \mathrm{i} \cdot \mathrm{Im}(f)$.

Definition. $f : [a, b] \to \mathbb{C}$ beschränkt heißt Riemann-integrierbar, wenn $\mathrm{Re}(f)$ und $\mathrm{Im}(f)$ integrierbar sind.

Alternativ kann man die Riemann-Integrierbarkeit über Riemann'sche Summenfolgen definiert werden

Definition. $f : [a, b] \to \mathbb{C}$ beschränkt heißt Riemann-integrierbar, wenn jede Riemannsche Summenfolge $(\sigma(Z_n, \xi_k^n, f))_n$ konvergiert.

Bemerkung. Beide Definitionen sind äquivalent, weil

$$\sigma(Z_n, \xi_k^n, f) = \sigma(Z_n, \xi_k^n, \mathrm{Re}(f)) + \mathrm{i}\sigma(Z_n, \xi_k^n, \mathrm{Im}(f))$$

und eine Folge $(c_n)_{n\in\mathbb{N}} \subset \mathbb{C}$ konvergiert, genau dann wenn $(\mathrm{Re}(c_n))_{n\in\mathbb{N}}$ und $(\mathrm{Im}(c_n))_{n\in\mathbb{N}}$ konvergieren.

5.2 Eigenschaften des Riemann-Integrals

5.2.1 Linearität des Riemann-Integrals

Lemma 5.2.1. Seien $f, g \in R([a, b])$, $\alpha, \beta \in \mathbb{R}$, dann ist $\alpha \cdot f + \beta \cdot g \in R([a, b])$ und

$$\int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x + \beta \int_a^b g(x)\,\mathrm{d}x$$

D.h. $R([a, b])$ ist ein $\mathbb{R}$-Vektorraum.

Beweis. Betrachte die Riemann'sche Summenfolge

$$\begin{aligned} \sigma(Z_n, \xi_k^n, \alpha f + \beta g) &= \sum_{k=1}^{n} [(\alpha f(\xi_k^n) + \beta g(\xi_k^n)) \cdot (x_k - x_{k-1})] \\ &= \alpha \cdot \sum_{k=1}^{n} f(\xi_k^n) \cdot (x_k - x_{k-1}) + \beta \sum_{k=1}^{n} g(\xi_k^n) \cdot (x_k - x_{k-1}) \end{aligned}$$

Grenzübergang $n \to \infty$ ergibt die Behauptung. □

5.2.2 Erhaltung der Riemann-Integrierbarkeit

Satz 5.2.2. Seien $f, g \in R([a,b])$, $h: [a,b] \to \mathbb{R}$ beschränkt. Dann gilt

i) Änderungen des Wertes von h an endlichen vielen Stellen $x \in [a,b]$ ändert nichts an der Riemann-Integrierbarkeit der Funktion h; insbesondere ändern sich die Werte $J_*(h)$, $J^*(h)$ nicht.

ii) Wenn $f(x) = g(x)$ für alle x aus einer in $[a,b]$ dichten Teilmenge A, dann ist

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x$$

iii) Wenn $\varphi: f([a,b]) \to \mathbb{R}$ lipschitz-stetig ist, dann ist auch $\varphi \circ f \in R([a,b])$.

iv) $|f|$, $f^+ := \max(0,f)$, $f^- := \min(0,f)$ und f^2 sind Riemann-integrierbar. Falls $|f(x)| \geq \delta$ für alle $x \in [a,b]$ für ein $\delta > 0$, dann ist auch $\frac{1}{f} \in R([a,b])$.

v) $f \cdot g$, $\max(f,g)$, $\min(f,g) \in R([a,b])$.

Beweis.

i) Es genügt zu zeigen, dass eine Änderung von h an einem einzigen Punkt $x_0 \in [a,b]$ nicht den Wert von $J_*(h)$ verändert. Definiere also

$$\tilde{h} := \begin{cases} h(x), & x \in [a,b] \setminus \{x_0\} \\ h(x_0) + \gamma, & x = x_0 \end{cases}$$

Dann ändert sich der Wert einer beliebigen Untersumme $s(Z,h)$ von h maximal um den Wert $\pm\gamma\,|Z|$, da sich die Änderung nur in einem Teilintervall der Zerlegung auswirkt. D.h. es gilt

$$\left| s(Z,\tilde{h}) - s(Z,h) \right| \leq \gamma\,|Z|$$

Für eine Zerlegungsnullfolge $(Z_n)_{n\in\mathbb{N}}$ folgt

$$0 \leq \left| J_*(\tilde{h}) - J_*(h) \right| = \lim_{n\to\infty} \left| s(Z_n,\tilde{h}) - s(Z_n,h) \right| \leq 0$$

also $J_*(\tilde{h}) = J_*(h)$.

ii) A ist dichte Teilmenge von $[a,b]$, d.h. für alle $\epsilon > 0$ und $x \in [a,b]$ gibt es ein $x_0 \in A$ mit $|x - x_0| < \epsilon$. Es genügt Riemann'sche Summenfolgen zu betrachten, für die die Zwischenpunkte ξ_k^n in A liegen:

$$\begin{aligned} \sigma(Z_n, \xi_k^n, f) &= \sum_{k=1}^n f(\xi_k^n) \cdot (x_k - x_{k-1}) \\ &= \sum_{k=1}^n g(\xi_k^n) \cdot (x_k - x_{k-1}) \\ &= \sigma(Z_n, \xi_k^n, g) \end{aligned}$$

Für $n \to \infty$ folgt die Behauptung.

iii) Es wird das Riemann'sche Integrabilitätskriterium angewandt. Sei dazu $\epsilon > 0$.

φ sei lipschitz-stetig mit Lipschitzkonstante $L > 0$, d.h. es gilt

$$|\varphi(f(x)) - \varphi(f(y))| \leq L \cdot |f(x) - f(y)| \quad \forall\, x, y \in [a, b]$$

Sei I_k ein Teilintervall von $[a, b]$. Dann gibt es $x, y \in I_k$ mit

$$\begin{aligned} \varphi(f(x)) &> \sup \varphi(f(I_k)) - \frac{\epsilon}{4(b-a)} \\ \varphi(f(y)) &< \inf \varphi(f(I_k)) + \frac{\epsilon}{4(b-a)} \end{aligned}$$

Dann gilt

$$\begin{aligned} \sup \varphi(f(I_k)) - \inf \varphi(f(I_k)) &< \varphi(f(x)) - \varphi(f(y)) + \frac{\epsilon}{2(b-a)} \\ &\leq L \cdot |f(x) - f(y)| + \frac{\epsilon}{2(b-a)} \\ &\leq L \cdot (\sup f(I_k) - \inf f(I_k)) + \frac{\epsilon}{2(b-a)} \qquad (1) \end{aligned}$$

Da f Riemann-integrierbar ist, gibt es nach Riemann'schen Integrabilitätskriterium eine Zerlegung Z von $[a, b]$ mit

$$S(Z, f) - s(Z, f) < \frac{\epsilon}{2L} \qquad (2)$$

Für diese Zerlegung gilt nun

$$\begin{aligned} S(Z, \varphi \circ f) - s(Z, \varphi \circ f) &= \sum_k \sup(\varphi \circ f)(I_k)\, |I_k| - \sum_k \inf(\varphi \circ f)(I_k)\, |I_k| \\ &= \sum_k (\sup(\varphi \circ f)(I_k) - \inf(\varphi \circ f)(I_k))\, |I_k| \\ &\overset{(1)}{\leq} \sum_k \left(L \cdot (\sup f(I_k) - \inf f(I_k)) + \frac{\epsilon}{2(b-a)} \right) |I_k| \\ &= L \cdot (S(Z, f) - s(Z, f)) + \frac{\epsilon}{2(b-a)} \cdot (b-a) \\ &\overset{(2)}{<} L \cdot \frac{\epsilon}{2L} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

iv) Es reicht sich zu überlegen, dass $\varphi:\ f([a, b]) \to \mathbb{R}$ definiert durch

$$\varphi(y) = |y|\,,\ \varphi(y) = y^+,\ \varphi(y) = y^- \text{ bzw. } \varphi(y) = y^2 \quad \forall\, y \in f([a, b])$$

lipschitz-stetig ist. Gleiches gilt für die Funktion

$$\begin{aligned} \varphi:\ ]-\infty, \delta] \cup [\delta, +\infty[&\to \mathbb{R} \\ y &\mapsto \frac{1}{y} \end{aligned}$$

mit $L = \frac{1}{\delta^2}$ als Lipschitz-Konstante.

v) Es gilt

$$4 \cdot f \cdot g = (f+g)^2 - (f-g)^2 \Leftrightarrow f \cdot g = \frac{(f+g)^2 - (f-g)^2}{4}$$

und

$$\max(f,g) = f + (g-f)^+, \ \min(f,g) = -\max(-f,-g)$$

Mit Lemma 5.2.1 und Punkt iv) folgt, dass die Funktionen integrierbar sind.

□

5.2.3 Monotonie des Integrals

Lemma 5.2.3. Seien $f, g \in [a,b] \to \mathbb{R}$ beschränkt. Wenn $f(x) \leq g(x)$ für fast alle $x \in [a,b]$, dann gilt

$$J_*(f) \leq J_*(g) \text{ und } J^*(f) \leq J^*(g)$$

insbesondere

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x$$

wenn $f, g \in R([a,b])$.

Beweis. für das untere Riemann-Integral: Die Behauptung folgt sofort aus

$$s(Z,f) = \sum_k \inf f(I_k)\,|I_k| \leq \sum_k \inf g(I_k)\,|I_k| = s(Z,g)$$

für alle Zerlegungen Z von $[a,b]$, geht man mit $Z = Z_n$ für $(Z_n)_{n\in\mathbb{N}}$ Zerlegungsnullfolge zum Grenzwert für $n \to \infty$ über. □

5.2.4 Dreiecksungleichung für Integrale

Lemma 5.2.4. Für $f \in R([a,b]$ ist

$$\left| \int_a^b f(x)\,\mathrm{d}x \right| \leq \int_a^b |f(x)|\,\mathrm{d}x$$

Speziell gilt: Wenn $|f(x)| \leq K$ für alle $x \in [a,b]$, dann gilt

$$\int_a^b f(x)\,\mathrm{d}x \leq K \cdot (b-a)$$

Beweis. Nach Satz 5.2.2 folgt aus $f \in R([a,b])$, dass $|f| \in R([a,b])$, und für jede Riemann'sche Summenfolge gilt

$$\begin{aligned} |\sigma(Z_n, \xi_k^n, f)| &= \left| \sum_k f(\xi_k^n)(x_k^{(n)} - x_{k-1}^{(n)}) \right| \\ &\leq \sum_k |f(\xi_k^n)|\,(x_k^{(n)} - x_{k-1}^{(n)}) \\ &= \sigma(Z_n, \xi_k^n, |f|) \end{aligned}$$

Für $n \to \infty$ folgt die gewünschte Ungleichung.

Der Zusatz folgt trivialerweise schon aus (5.1.2) und den Definitionen

$$\int_a^b |f(x)| \,\mathrm{d}x \leq S(Z, |f|) \leq S((a,b), |f|) = \sup |f([a,b])| \cdot (b-a) = K \cdot (b-a)$$

□

5.2.5 Mittelwertsatz der Integralrechnung

Satz 5.2.5. Es sei $f \in R([a,b])$ und

$$\mu(f) := \frac{1}{b-a} \int_a^b f(x) \,\mathrm{d}x$$

der **Integralmittelwert von** f **über** $[a,b]$. Nun gilt

a) der Mittelwersatz der Integralrechnung

$$\inf f([a,b]) \leq \mu(f) \leq \sup f([a,b])$$

Insbesondere: Ist f stetig auf $[a,b]$, dann existiert ein $\xi \in \,]a,b[$ mit $f(\xi) = \mu(f)$.

b) und der erweiterte Mittelwertsatz der Integralrechnung:

Ist $p \in R([a,b])$, $p \geq 0$, und gilt $m \leq f(x) \leq M$ für alle $x \in [a,b]$, dann gilt

$$m \cdot \int_a^b p(x) \,\mathrm{d}x \leq \int_a^b f(x)p(x) \,\mathrm{d}x \leq M \int_a^b p(x) \,\mathrm{d}x$$

Insbesondere gilt: Ist f stetig auf $[a,b]$, dann existiert ein $\xi \in \,]a,b[$ mit

$$\int_a^b f(x)p(x) \,\mathrm{d}x = f(\xi) \cdot \int_a^b p(x) \,\mathrm{d}x$$

Beweis.

a) Es gilt

$$\inf f([a,b]) \leq f(x) \leq \sup f([a,b]) \quad \forall\, x \in [a,b]$$

Mit Lemma 5.2.3 folgt

$$(b-a) \cdot \inf f([a,b]) \leq \int_a^b f(x) \,\mathrm{d}x \leq (b-a) \sup f([a,b])$$

Division durch $(b-a)$ ergibt die Behauptung. Der Zusatz folgt aus dem Zwischenwertsatz für stetige Funktionen.

b) Es gilt also

$$mp(x) \leq f(x)p(x) \leq Mp(x) \quad \forall\, x \in [a,b]$$

Aus der Monotonie und Linearität des Integrals folgt erneut die Behauptung. Den Zusatz erhält man wieder durch den Zwischenwertsatz.

□

5.2.6 Vertauschung von Limes und Integration

Es stellt sich wieder die Frage: Ist $(f_n)_{n\in\mathbb{N}} \subset R([a,b])$ und konvergiert $(f_n)_n$ punktweise bzw. gleichmäßig gegen eine Grenzfunktion f auf $[a,b]$, ist dann $f \in R([a,b])$? Wenn ja, gilt dann

$$\lim_{n\to\infty} \int\limits_a^b f_n(x)\,\mathrm{d}x = \int\limits_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \int\limits_a^b f(x)\,\mathrm{d}x \text{ ?}$$

Für punktweise Konvergenz sind die Fragen mit nein zu beantworten. Aber es gilt der folgende

Satz 5.2.6. Seien $(f_n)_{n\in\mathbb{N}} \subset R([a,b])$, $(f_n)_n$ auf $[a,b]$ gleichmäßig konvergent gegen eine Funktion f. Dann folgt $f \in R([a,b])$ und

$$\lim_{n\to\infty} \int\limits_a^b f_n(x)\,\mathrm{d}x = \int\limits_a^b f(x)\,\mathrm{d}x$$

Beweis. Konvergiert $(f_n)_n$ gegen f gleichmäßig, so gilt

$$\forall\,\epsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall\, n \geq n_0 \,:\, |f_n(x) - f(x)| < \epsilon \quad \forall\, x \in [a,b]$$

Sei also $\epsilon > 0$, und n_0 gemäß dieser Aussage gewählt. Es gilt somit

$$f_n(x) - \epsilon < f(x) < f_n(x) + \epsilon \quad \forall\, x \in [a,b]\,,\ \forall\, n \geq n_0$$

Mittels der Monotonie der Unterintegrale laut Lemma 5.2.3 folgt

$$J_*(f_n(x) - \epsilon) \leq J_*(f) \leq J_*(f_n(x) + \epsilon)$$

Da $f_n(x) \pm \epsilon$ integrierbar ist, folgt

$$\int\limits_a^b f_n(x)\,\mathrm{d}x - \epsilon(b-a) \leq J_*(f) \leq \int\limits_a^b f_n(x)\,\mathrm{d}x + \epsilon(b-a) \qquad \Rightarrow$$

$$\left| \int\limits_a^b f_n(x)\,\mathrm{d}x - J_*(f) \right| < \epsilon(b-a) \quad \forall\, n \geq n_0 \qquad \Rightarrow$$

$$\lim_{n\to\infty} \int\limits_a^b f_n(x)\,\mathrm{d}x = J_*(f)$$

Dieselbe Argumentation für obere Riemann-Integrale liefert

$$\lim_{n\to\infty} \int\limits_a^b f_n(x)\,\mathrm{d}x = J^*(f)$$

und damit

$$J_*(f) = J^*(f) \Rightarrow f \in R([a,b]) \text{ mit } \lim_{n\to\infty} \int\limits_a^b f_n(x)\,\mathrm{d}x = \int\limits_a^b f(x)\,\mathrm{d}x$$

□

Korollar 5.2.7. Sei $(f_n)_{n\in\mathbb{N}} \subset R([a,b])$, $S(x) = \sum_{n=0}^{\infty} f_n(x)$ gleichmäßig konvergent auf $[a,b]$. Dann ist $S \in R([a,b])$ und

$$\int\limits_a^b S(x)\,\mathrm{d}x = \lim_{N\to\infty} \int\limits_a^b \sum_{n=0}^{N} f_n(x)\,\mathrm{d}x = \sum_{n=0}^{\infty} \int\limits_a^b f_n(x)\,\mathrm{d}x$$

5.2.7 Beispiel

Die Exponentialreihe ist gleichmäßig konvergent auf $[a,b]$, weshalb gilt

$$\begin{aligned}
\int\limits_a^b \mathrm{e}^x\,\mathrm{d}x &= \int\limits_a^b \sum_{n=0}^{\infty} \frac{x^n}{n!}\,\mathrm{d}x = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot \int\limits_a^b x^n\,\mathrm{d}x \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \frac{b^{n+1} - a^{n+1}}{n+1} \\
&= \sum_{n=0}^{\infty} \frac{b^{n+1}}{(n+1)!} - \sum_{n=0}^{\infty} \frac{a^{n+1}}{(n+1)!} \\
&= \sum_{n=1}^{\infty} \frac{b^n}{n!} - \sum_{n=1}^{\infty} \frac{a^n}{n!} \\
&= \left(\mathrm{e}^b - 1\right) - (\mathrm{e}^a - 1) = \mathrm{e}^b - \mathrm{e}^a
\end{aligned}$$

d.h.

$$\int\limits_a^b \mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^b - \mathrm{e}^a$$

5.2.8 Integrale über Teilintervallen

Satz 5.2.8. Sei $f:\ [a,b] \to \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ beschränkt, $a < c < b$. Dann gilt:

$$f \in R([a,b]) \iff f \in R([a,c]) \text{ und } f \in R([c,b])$$

In diesem Fall gilt

$$\int\limits_a^b f(x)\,\mathrm{d}x = \int\limits_a^c f(x)\,\mathrm{d}x + \int\limits_c^b f(x)\,\mathrm{d}x$$

Beweis. Sei $(Z_n)_{n\in\mathbb{N}}$ Zerlegungsnullfolge, sodass jedes Z_n den Stützpunkt c enthält. Seien $\left(Z_n^1\right)_{n\in\mathbb{N}}$ und $\left(Z_n^2\right)_{n\in\mathbb{N}}$ die induzierten Zerlegungsfolgen von $[a,c]$ und $[c,b]$. Dann gilt

$$\begin{aligned}
s(Z_n, f) &= s(Z_n^1, f) + s(Z_n^2, f) \\
S(Z_n, f) &= S(Z_n^1, f) + S(Z_n^2, f)
\end{aligned}$$

also

$$\underbrace{S(Z_n, f) - s(Z_n, f)}_{O(Z_n, f)} = \underbrace{(S(Z_n^1, f) - s(Z_n^1, f))}_{O(Z_n^1, f)} + \underbrace{(S(Z_n^2, f) - s(Z_n^2, f))}_{O(Z_n^2, f)}$$

wobei

$$O(Z_n, f) \overset{n\to\infty}{\longrightarrow} 0 \iff \left[O(Z_n^1, f) \overset{n\to\infty}{\longrightarrow} 0 \text{ und } O(Z_n^2, f) \overset{n\to\infty}{\longrightarrow} 0\right]$$

□

5.2.9 Schreibkonvention für Integralgrenzen

Definition 5.2.9. Für $a < b$ und $f \in R([a, b])$ setze

$$\int\limits_a^b f(x)\,\mathrm{d}x = -\int\limits_b^a f(x), \quad \int\limits_c^c f(x)\,\mathrm{d}x = 0 \quad \forall\, c \in [a, b]$$

Somit ist für $f \in R([a, b])$ und $\alpha, \beta, \gamma \in [a, b]$

$$\int\limits_\alpha^\beta f(x)\,\mathrm{d}x + \int\limits_\beta^\gamma f(x)\,\mathrm{d}x + \int\limits_\gamma^\alpha f(x)\,\mathrm{d}x = 0$$

5.2.10 Integralfunktion

Satz 5.2.10. Sei $f : [a, b] \to \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ beschränkt und R-integrierbar, $c \in [a, b]$.

$$\begin{aligned} F :\ [a, b] &\to \mathbb{R} \\ x &\mapsto \int\limits_c^x f(t)\,\mathrm{d}t \end{aligned}$$

Diese **Integralfunktion** genannte Funktion ist lipschitz-stetig. Genauer:

Gilt $|f(x)| < K$ für alle $x \in [a, b]$ mit einem $K > 0$, dann gilt

$$|F(x) - F(y)| \leq K\,|x - y| \quad \forall\, x, y \in [a, b]$$

Beweis. Es gilt für $y \leq x$

$$|F(x) - F(y)| = \left|\int\limits_y^x f(t)\,\mathrm{d}t\right| \leq \int\limits_y^x |f(t)|\,\mathrm{d}t \leq \int\limits_y^x K\,\mathrm{d}t = K\,|x - y|$$

□

5.2.11 Hauptsatz der Differential- und Integralrechnung

Satz 5.2.11.

a) Sei $f \in R([a, b])$ und stetig in $x_0 \in [a, b]$. Dann ist $F(x) = \int\limits_c^x f(t)\,\mathrm{d}t$ für alle $c \in [a, b]$ differenzierbar in x_0 mit $F'(x_0) = f(x_0)$.
Es gilt also $f \in C^0([a, b]) \Rightarrow F \in C^1([a, b])$ mit $F' = f$.

b) Sei $F: [a,b] \to \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ differenzierbar und $F' \in R([a,b])$. Dann ist:

$$F(b) - F(a) = \int\limits_a^b F'(t)\,\mathrm{d}t$$

und entsprechend für $x, c \in [a,b]$

$$F(x) = F(c) + \int\limits_c^x F'(t)\,\mathrm{d}t$$

Beweis.

a) Sei zur Abkürzung $I_x := [\min\{x, x_0\}, \max\{x, x_0\}]$ das Intervall zwischen x_0 und x. Es gilt

$$\begin{aligned} |F(x) - F(x_0) - f(x_0)(x - x_0)| &= \left| \int\limits_{x_0}^x f(t)\,\mathrm{d}t - \int\limits_{x_0}^x f(x_0)\,\mathrm{d}t \right| \\ &= \left| \int\limits_{x_0}^x (f(t) - f(x_0))\,\mathrm{d}t \right| \\ &\leq \sup_{t \in I_x} |f(t) - f(x_0)| \cdot |x - x_0| \qquad \Rightarrow \\ \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| &\leq \sup_{t \in I_x} |f(t) - f(x_0)| \end{aligned}$$

Sei $\epsilon > 0$ gegeben. Da f stetig ist, gibt es ein $\delta > 0$, sodass für alle $t \in [a,b]$ mit $|t - x_0| < \delta$ gilt: $|f(t) - f(x_0)| < \epsilon$. Folglich gilt

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \epsilon \quad \forall\ |x - x_0| < \delta$$

und damit die Behauptung.

b) Sei $(Z_n)_{n \in \mathbb{N}}$ Zerlegungsnullfolge mit $Z_n = (a = x_0,\ x_1,\ \ldots,\ x_n = b)$. Dann gilt für jedes $n \in \mathbb{N}$

$$F(b) - F(a) = \sum_{k=1}^n (F(x_k) - F(x_{k-1}))$$

Nach Mittelwertsatz der Integralrechnung gibt es also $\xi_k \in\ ]x_{k-1}, x_k[$ mit

$$F(b) - F(a) = \sum_{k=1}^n (x_k - x_{k-1}) \cdot F'(\xi_k) \overset{n \to \infty}{\longrightarrow} \int\limits_a^b F'(t)\,\mathrm{d}t$$

□

5.2.12 Stammfunktion

Definition 5.2.12. Sei $J \subseteq \mathbb{R}$ ein nicht-leeres Intervall, $f, F : J \to \mathbb{R}$. F heißt **Stammfunktion** von f, wenn F auf J differenzierbar ist mit $F' = f$.

Bemerkung.

1. Nach Satz 5.2.11 folgt also:

   $\boxed{\text{Jede stetige Funktion } f : J \to \mathbb{R} \text{ besitzt eine Stammfunktion.}}$

2. Eine Funktion $G : J \to \mathbb{R}$ ist Stammfunktion von $f : J \to \mathbb{R}$ genau dann, wenn

   $$G(x) = \int_c^x f(t)\,\mathrm{d}t, \quad x, c \in J$$

   Somit gilt: Integrale von Funktionen $f \in R([a,b])$, die Stammfunktionen besitzen, können über diese berechnet werden:

   $$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) =: F(x)\Big|_a^b$$

Notation Ist G Stammfunktion von f, so schreiben wir

$$G(x) = \underbrace{\int f(x)\,\mathrm{d}x}_{\text{unbestimmtes Integral}} + \text{const}$$

Achtung. Stammfunktionen sind nur bis auf eine Konstante eindeutig bestimmt.

Bemerkung.

1. Eine Stammfunktion kann nicht immer explizit angegeben werden, betrachte zum Beispiel $f(x) = \mathrm{e}^{-x^2} \quad \forall x \in \mathbb{R}$.

2. Eine nicht-stetige Funktion besitzt im Allgemeinen keine Stammfunktion, betrachte zum Beispiel die Heaviside-Funktion

   $$H(x) := \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

   Sie ist R-integrierbar auf jedem Intervall $[a,b] \subseteq \mathbb{R}$, besitzt aber keine Stammfunktion, was man leicht durch den Zwischenwertsatz für Ableitungen einsieht.

5.2.13 Allgemeines Taylor-Restglied

Satz 5.2.13. Sei $J \subseteq \mathbb{R}$ ein Intervall, $f \in C^{n+1}(J)$ für ein $n \in \mathbb{N}_0$ und $a, x \in J$. Dann ist

$$\begin{aligned} f(x) &= T_n(f,a) + \int_a^x \frac{(x-t)^n}{n!} \cdot f^{(n+1)}(t)\,\mathrm{d}t \\ &\overset{\text{MWS der Integralrechnung}}{=} T_n(f,a) + \frac{f^{(n+1)}(\xi)}{n!} \int_a^x (x-t)^n\,\mathrm{d}t \end{aligned}$$

für ein $\xi \in \,]\min\{a,x\}, \max\{a,x\}[$.

Beweis. durch Induktion über n.

$n = 0$:

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t$$

gilt nach Satz 5.2.11 b).

$n \to (n+1)$: Sei $x \in J$ fest. Betrachte

$$F(t) := \frac{(x-t)^{n+1}}{(n+1)!} \cdot f^{(n+1)}(t)$$

Dann ist

$$F'(t) = \frac{(x-t)^{n+1}}{(n+1)!} \cdot f^{(n+2)}(t) - \frac{(x-t)^n}{n!} \cdot f^{(n+1)}(t)$$

Integriert man auf beiden Seiten von a bis x

$$\begin{aligned} \underbrace{F(x)}_{=0} - F(a) &= R_{n+1}(x) - R_n(x) && \Rightarrow \\ R_n(x) &= F(a) + R_{n+1}(x) && \stackrel{\text{IV}}{\Rightarrow} \\ f(x) - T_n(f,a) &= F(a) + R_{n+1}(x) && \Rightarrow \\ f(x) &= T_n(f,a) + F(a) + R_{n+1}(x) && \Rightarrow \\ f(x) &= T_{n+1}(f,a) + R_{n+1}(x) \end{aligned}$$

□

5.3 Integrationstechniken

5.3.1 Partielle Integration

Satz 5.3.1. Sei $I \subseteq \mathbb{R}$ ein Intervall, $f, g : I \to \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ differenzierbar auf I. Dann gilt:

- $$\int f(x)g'(x)\,\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\,\mathrm{d}x$$

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

Beweis. Folgt direkt aus der Produktregel: $(f \cdot g)' = f' \cdot g + g' \cdot f$. □

5.3.2 Substitutionsregel

Satz 5.3.2. Seien $I, J \subseteq \mathbb{R}$ Intervalle, $\varphi : I \to J$ differenzierbar, $f : J \to \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ stetig und sei F Stammfunktion von f. Dann gilt:

i)

$$\frac{\mathrm{d}}{\mathrm{d}t}F(\varphi(t)) = f(\varphi(t)) \cdot \varphi'(t) \Rightarrow \int f(x)\,\mathrm{d}x\Big|_{x=\varphi(t)} = \int f(\varphi(t)) \cdot \varphi'(t)\,\mathrm{d}t$$

Ist zusätzlich φ streng monoton in I, dann existiert die Umkehrfunktion φ^{-1} und für $t = \varphi^{-1}(x)$ gilt dann

$$\int f(x)\,\mathrm{d}x = \int f(\varphi(t)) \cdot \varphi'(t)\,\mathrm{d}t\Big|_{t=\varphi^{-1}(x)}$$

ii)

$$\int_a^b f(x)\,\mathrm{d}x = \int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)\,\mathrm{d}t$$

mit $\alpha = \varphi^{-1}(a)$ und $\beta = \varphi^{-1}(b)$, oder

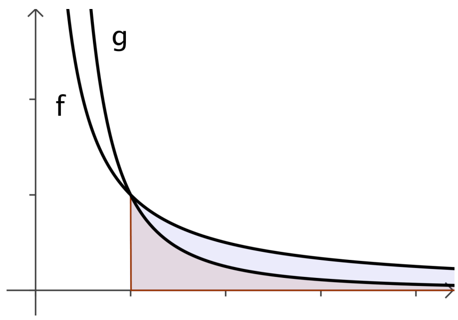
$$\int_a^b f(\varphi(x)) \cdot \varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(u)\,\mathrm{d}u$$

5.4 Uneigentliche Integrale

5.4.1 Motivation

Betrachte

$$f(x) = \frac{1}{x} \qquad\qquad g(x) = \frac{1}{x^2}$$



Was ist $\int_1^\infty f(x)\,\mathrm{d}x$ und $\int_1^\infty g(x)\,\mathrm{d}x$?

$$\int_1^\infty f(x)\,\mathrm{d}x \overset{?}{:=} \lim_{c\to\infty} \int_1^c \frac{1}{x}\,\mathrm{d}x = \lim_{c\to\infty} \ln|c| = \infty$$

$$\int_1^\infty g(x)\,\mathrm{d}x \overset{?}{:=} \lim_{c\to\infty} \int_1^c \frac{1}{x^2}\,\mathrm{d}x = \lim_{c\to\infty} \left(1 - \frac{1}{c}\right) = 1$$

Definition 5.4.1. Sei $a \in \mathbb{R}$, $f : [a, +\infty[\to \mathbb{R}$ mit $f \in R([a, c])$ für alle $c \in \mathbb{R}$, $c > a$.

a) Wenn $\lim\limits_{c\to\infty} \int\limits_a^c f(x)\,\mathrm{d}x$ in $\mathbb{R}$ existiert, dann existiert (oder konvergiert) das sogenannte **uneigentliche Integral**

$$\int_a^\infty f(x)\,\mathrm{d}x := \lim_{c\to\infty} \int_a^c f(x)\,\mathrm{d}x$$

In diesen Fall heißt f uneigentlich R-integrierbar auf $[a, +\infty[$.

Falls $\lim\limits_{c\to\infty} \int\limits_a^c f(x)\,\mathrm{d}x$ nicht existiert, dann heißt $\int\limits_a^\infty f(x)\,\mathrm{d}x$ divergent.

b) Analog für $\int\limits_{-\infty}^a f(x)\,\mathrm{d}x$ für eine Funktion $f :]-\infty, a] \to \mathbb{R}$ mit $f \in R([c, a])$ für alle $c < a$.

c) Für $f :]-\infty, +\infty[\to \mathbb{R}$ mit $f \in R([c, d])$ für alle $c, d \in \mathbb{R}$, $c < d$, heißt $\int\limits_{-\infty}^\infty f(x)\,\mathrm{d}x$ konvergent, falls $\int\limits_{-\infty}^a f(x)\,\mathrm{d}x$ und $\int\limits_a^\infty f(x)\,\mathrm{d}x$ konvergent sind mit $a \in \mathbb{R}$, d.h. falls $\lim\limits_{c\to-\infty} \int\limits_c^a f(x)\,\mathrm{d}x$ und $\lim\limits_{d\to\infty} \int\limits_a^d f(x)\,\mathrm{d}x$ in $\mathbb{R}$ existieren für ein $a \in \mathbb{R}$. In diesem Fall gilt

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x := \int_{-\infty}^a f(x)\,\mathrm{d}x + \int_a^\infty f(x)\,\mathrm{d}x$$

5.4.2 Beispiele

1.
$$\int_1^\infty \frac{1}{x}\,\mathrm{d}x = \lim_{c\to\infty} \int_1^c \frac{1}{x}\,\mathrm{d}x = \lim_{c\to\infty} \ln(x)\big|_1^c = \lim_{c\to\infty} (\ln(c) - \ln(1)) = +\infty$$

2.
$$\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x = \lim_{c\to\infty} \int_1^c \frac{1}{x^2}\,\mathrm{d}x = \lim_{c\to\infty} -\frac{1}{x}\bigg|_1^c = \lim_{c\to\infty} \left(-\frac{1}{c} + 1\right) = 1$$

3. $\int\limits_{-\infty}^\infty x\,\mathrm{d}x$ existiert nicht (ist divergent), denn

$$\int_{-\infty}^0 x\,\mathrm{d}x = \lim_{c\to-\infty} \int_c^0 x\,\mathrm{d}x = \lim_{c\to-\infty} \frac{x^2}{2}\bigg|_c^0 = \lim_{c\to-\infty} \left(0 - \frac{c^2}{2}\right) = +\infty$$

analog

$$\int_0^\infty x\,\mathrm{d}x = +\infty$$

Achtung.

$$\lim_{c\to\infty}\int\limits_{-c}^{c} x\,\mathrm{d}x = \lim_{c\to\infty}\left.\frac{x^2}{2}\right|_{-c}^{c} = 0 \quad \forall\, c>0$$

Dies reicht nicht für die Konvergenz des uneigentlichen Integrals $\int\limits_{-\infty}^{\infty} x\,\mathrm{d}x$ aus!

Bemerkung. Die bekannten Eigenschaften wie Linearität, Ordnungserhaltung etc. übertragen sich auf das uneigentliche Riemann-Integral. Sind zum Beispiel f, g auf $[a,+\infty[$ uneigentlich R-integrierbar und $f(x)\le g(x)$ für alle $x\in[a,+\infty[$, dann gilt

$$\int\limits_a^\infty f(x)\,\mathrm{d}x \le \int\limits_a^\infty g(x)\,\mathrm{d}x$$

Klar ist auch: Wenn $f:\ [a,+\infty[\ \to\mathbb{R}$ mit $f\in R([a,c])\quad\forall\, c\in\mathbb{R},\ c>a$, dann konvergiert $\int\limits_a^\infty f(x)\,\mathrm{d}x$ genau dann, wenn $\int\limits_c^\infty f(x)\,\mathrm{d}x$ konvergiert für ein $c\in\mathbb{R},\ c\ge a$. und in diesem Fall ist

$$\int\limits_a^\infty f(x)\,\mathrm{d}x = \int\limits_a^c f(x)\,\mathrm{d}x + \int\limits_c^\infty f(x)\,\mathrm{d}x \quad \forall\, c\in\mathbb{R},\ c\ge a$$

5.4.3 Cauchy-Kriterium

Satz 5.4.2. Sei $f:\ [a,+\infty[\ \to\mathbb{R}$ mit $f\in R([a,c])\quad\forall\, c\in\mathbb{R},\ c\ge a$. Dann gilt

$$\int\limits_a^\infty f(x)\,\mathrm{d}x \text{ konvergiert} \iff \forall\,\epsilon>0\quad \exists\, A\ge a\ :\ \left|\int\limits_c^d f(x)\,\mathrm{d}x\right| < \epsilon \quad \forall\, A\le c\le d$$

5.4.4 Beispiel

Sei $\alpha>0$. Behauptung:

$$\int\limits_1^\infty \frac{\sin(t)}{t^\alpha}\,\mathrm{d}t \text{ konvergiert}$$

Der Integrand ist stetig auf $[1,\infty[$, also R-integrierbar auf $[1,c]\quad\forall\, c>1$. Betrachte für $d\ge c\ge 1$

$$\int\limits_c^d \frac{\sin(t)}{t^\alpha}\,\mathrm{d}t = \left.\frac{-\cos(t)}{t^\alpha}\right|_c^d - \int\limits_c^d \frac{-\cos(t)\cdot(-\alpha)}{t^{\alpha+1}}\,\mathrm{d}x \qquad \Rightarrow$$

$$\left|\int\limits_c^d \frac{\sin(t)}{t^\alpha}\,\mathrm{d}x\right| \le \frac{1}{c^\alpha}+\frac{1}{d^\alpha}+\int\limits_c^d \frac{\alpha}{t^{\alpha+1}}\,\mathrm{d}x = \frac{1}{c^\alpha}+\frac{1}{d^\alpha}-\frac{1}{d^\alpha}+\frac{1}{c^\alpha} = \frac{2}{c^\alpha}$$

Sei nun $\epsilon > 0$. Wähle $A \geq 1$ mit $\frac{2}{A^\alpha} < \epsilon$, so folgt für alle $d \geq c \geq A$:

$$\left| \int_c^d \frac{\sin(t)}{t^\alpha} \, \mathrm{d}x \right| \leq \frac{2}{c^\alpha} \leq \frac{2}{A^\alpha} < \epsilon$$

5.4.5 Absolute Konvergenz

Definition 5.4.3. Sei $f : [a, \infty[\to \mathbb{R}$ mit $f \in R([a, c]) \quad c \geq a$. $\int_a^\infty f(x) \, \mathrm{d}x$ heißt **absolut konvergent**, wenn $\int_a^\infty |f(x)| \, \mathrm{d}x$ konvergiert.

Satz 5.4.4. Ist $f : [a, \infty[\to \mathbb{R}$ mit $f \in R([a, c]) \quad c \geq a$, so gilt

i) Ist $\int_a^\infty |f(x)| \, \mathrm{d}x$ absolut konvergent, dann ist $\int_a^\infty f(x) \, \mathrm{d}x$ konvergent.

ii) Wenn $|f(x)| \leq g(x) \quad \forall x \in [a, \infty[$ und $\int_a^\infty g(x) \, \mathrm{d}x$ konvergiert, dann konvergiert auch $\int_a^\infty f(x) \, \mathrm{d}x$ (absolut) (**Majorantenkriterium**).

Beweis. mit Cauchy-Kriterium: Es gilt

i) $\left| \int_a^\infty f(x) \, \mathrm{d}x \right| \leq \int_a^\infty |f(x)| \, \mathrm{d}x$ und

ii) $\left| \int_a^\infty |f(x)| \, \mathrm{d}x \right| \leq \int_a^\infty |f(x)| \, \mathrm{d}x \leq \int_a^\infty g(x) \, \mathrm{d}x$.

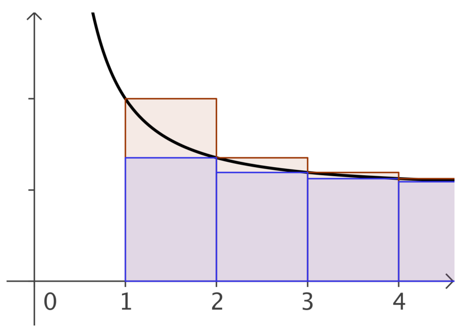
□

5.4.6 Integralvergleichskriterium für Reihen

Satz 5.4.5. Sei $f : [a, +\infty[\to \mathbb{R}$ monoton fallend mit $f \in R([a, c]) \quad \forall c > a$. Dann gilt

$$\sum_{k=n+1}^{\infty} f(k) \leq \int_n^\infty f(x) \, \mathrm{d}x \leq \sum_{k=n}^{\infty} f(k) \quad \forall n \in \mathbb{N},\, n \geq a$$

Mit anderen Worten: Die unendliche Reihe und das uneigentliche Integral haben das gleiche Konvergenzverhalten.



5.4.7 Vertauschbarkeit von Limes und Integration

Satz 5.4.6. Seien $f_n, g : [a, +\infty[\to \mathbb{R}$ mit $f_n, g \in R([a,c]) \quad \forall\, c > a,\ n \in \mathbb{N}$. Es gelte zudem

$$|f(x)| \leq g(x) \quad \forall\, x \in [a, +\infty[\,,\ n \in \mathbb{N} \quad (*)$$

Wenn außerdem $\int\limits_a^\infty g(x)\,\mathrm{d}x$ konvergiert und $f_n \to f$ gleichmäßig auf $[a, c] \quad \forall\, c > a$, dann existieren die uneigentlichen Integrale $\int\limits_a^\infty f_n(x)\,\mathrm{d}x$ und $\int\limits_a^\infty f(x)\,\mathrm{d}x$ und es gilt

$$\lim_{n\to\infty} \int_a^\infty f_n(x)\,\mathrm{d}x = \int_a^\infty f(x)\,\mathrm{d}x$$

Beweis. Da g alle f_n majorisiert, folgt aus der Konvergenz von $\int\limits_a^\infty g(x)\,\mathrm{d}x$ die (absolute) Konvergenz von $\int\limits_a^\infty f_n(x)\,\mathrm{d}x$ für alle $n \in \mathbb{N}$. Mit der punktweisen Konvergenz $f_n \to f$ auf $[a, +\infty[$ folgt aus $(*)$ auch $|f(x)| \leq g(x) \quad \forall\, x \in [a, +\infty[$ und somit nach dem Majorantenkriterium auch die Konvergenz von $\int\limits_a^\infty f(x)\,\mathrm{d}x$. Weiter ist

$$\begin{aligned}
\left| \int_a^\infty f_n(x)\,\mathrm{d}x - \int_a^\infty f(x)\,\mathrm{d}x \right| &= \left| \int_a^\infty f_n(x) - f(x)\,\mathrm{d}x \right| \\
&\overset{\forall\, c \geq a}{\leq} \left| \int_a^c f_n(x) - f(x)\,\mathrm{d}x \right| + \int_c^\infty |f_n(x) - f(x)|\,\mathrm{d}x \\
&\leq \left| \int_a^c f_n(x) - f(x)\,\mathrm{d}x \right| + 2\int_c^\infty g(x)\,\mathrm{d}x
\end{aligned}$$

Da $\int\limits_a^\infty g(x)\,\mathrm{d}x$ konvergiert, gilt

$$\lim_{c\to\infty} \int_c^\infty g(x)\,\mathrm{d}x = 0$$

Sei nun $\epsilon > 0$. Wähle $c > a$ mit

$$2\int_c^\infty g(x)\,\mathrm{d}x < \frac{\epsilon}{2}$$

c ist nun fest gewählt. Nach Satz 5.2.6 über die Vertauschbarkeit von Limes und Integration für eigentliche R-Integrale gilt

$$\lim_{n\to\infty} \int_a^c f_n(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x$$

Also existiert ein $n_0 \in \mathbb{N}$ mit

$$\left| \int_a^c f_n(x) - f(x) \,\mathrm{d}x \right| < \frac{\epsilon}{2} \quad \forall\, n \geq n_0$$

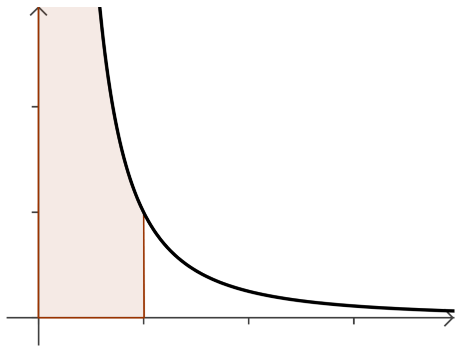
Folglich gilt

$$\left| \int_a^\infty f_n(x) \,\mathrm{d}x - \int_a^\infty f(x) \,\mathrm{d}x \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad \forall\, n \geq n_0$$

□

5.4.8 Uneigentliche Integrale unbeschränkter Funktionen

Wie sieht es mit dem Flächeninhalt unter dem Graphen unseres Eingangsbeispiels von 0 bis 1 aus?



$$\int_0^1 \frac{1}{x^2} \,\mathrm{d}x \overset{?}{:=} \lim_{c \searrow 0} \int_c^1 \frac{1}{x^2} \,\mathrm{d}x = \lim_{c \searrow 0} \left(-1 + \frac{1}{c} \right) = +\infty$$

Definition 5.4.7. Seien $a, b \in \mathbb{R}$, $a < b$.

i) Sei f eventuell unbeschränkt bei a, R-integrierbar auf $[c, b] \quad \forall\, a < c < b$. Dann heißt das uneigentliche Integral $\int_a^b f(x) \,\mathrm{d}x$ konvergent, wenn

$$\lim_{c \searrow a} \int_c^b f(x) \,\mathrm{d}x$$

in $\mathbb{R}$ existiert. In diesem Fall definiert man

$$\int_a^b f(x) \,\mathrm{d}x := \lim_{c \searrow a} \int_c^b f(x) \,\mathrm{d}x$$

Andernfalls heißt $\int_a^b f(x) \,\mathrm{d}x$ divergent.

ii) Enstprechend definiert man das uneigentliche Integral $\int\limits_a^b f(x)\,\mathrm{d}x$, wenn f bei b unbeschränkt ist:

$$\int\limits_a^b f(x)\,\mathrm{d}x := \lim_{c\nearrow b}\int\limits_a^c f(x)\,\mathrm{d}x$$

falls die rechte Seite existiert.

iii) Falls f bei a und b unbeschränkt ist, $f \in R([c,d]) \quad \forall\, a < c < d < b$, dann heißt $\int\limits_a^b f(x)\,\mathrm{d}x$ konvergent, falls für ein $c \in\,]a,b[$ $\int\limits_a^c f(x)\,\mathrm{d}x$ und $\int\limits_c^b f(x)\,\mathrm{d}x$ konvergieren; in diesem Fall definiert man

$$\int\limits_a^b f(x)\,\mathrm{d}x := \int\limits_a^c f(x)\,\mathrm{d}x + \int\limits_c^b f(x)\,\mathrm{d}x$$

iv) Ist f auf $]a,c[$ mit eventueller Ausnahme eines Punktes $c \in\,]a,b[$ definiert, integrierbar auf jedem abgeschlossenen Intervall $[d,e] \subset (]a,c[\setminus \{c\})$ und bei c unbeschränkt, dann heißt $\int\limits_a^b f(x)\,\mathrm{d}x$ konvergent, falls $\int\limits_a^c f(x)\,\mathrm{d}x$ und $\int\limits_c^b f(x)\,\mathrm{d}x$ konvergieren. Dann ist

$$\int\limits_a^b f(x)\,\mathrm{d}x := \int\limits_a^c f(x)\,\mathrm{d}x + \int\limits_c^b f(x)\,\mathrm{d}x$$

5.4.9 Beispiele

1.
$$\int\limits_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{c\nearrow 1}\int\limits_0^c \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x$$

Substituiere $x = \varphi(t) = \sin(t)$ $(\varphi'(t) = \cos(t))$:

$$\begin{aligned}\int\limits_0^c \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x &= \int\limits_{\arcsin(0)}^{\arcsin(c)} \frac{1}{\sqrt{1-\sin^2(t)}}\cdot\cos(t)\,\mathrm{d}t\\ &= \int\limits_0^{\arcsin(c)} \frac{1}{|\cos(t)|}\cdot\cos(t)\,\mathrm{d}t = \int\limits_0^{\arcsin(c)} 1\,\mathrm{d}t = \arcsin(c)\end{aligned}$$

Damit folgt

$$\int\limits_0^1 \frac{1}{\sqrt{1-x^2}} = \lim_{c\nearrow 1}\arcsin(c) = \frac{\pi}{2}$$

Das uneigentliche Integral ist also konvergent und gleich $\pi/2$.

2. Sei $\alpha > 0$.

$$\int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x = \lim_{c\searrow 0}\int_c^1 \frac{1}{x^\alpha}\,\mathrm{d}x$$

Für $0 < c < 1$ hat man

$$\int_c^1 \frac{1}{x^\alpha}\,\mathrm{d}x = \begin{cases} \ln|x|\big|_c^1, & \alpha = 1 \\ \frac{x^{-\alpha+1}}{1-\alpha}\Big|_c^1, & \alpha \neq 1 \end{cases}$$

Folglich ist

$$\lim_{c\searrow 0}\int_c^1 \frac{1}{x^\alpha}\,\mathrm{d}x = \begin{cases} +\infty, & \alpha = 1 \\ \frac{1}{1-\alpha}, & 0 < \alpha < 1 \\ +\infty, & \alpha > 1 \end{cases}$$

Das uneigentliche Integral ist divergent für $a \geq 1$, konvergent und gleich $\frac{1}{1-\alpha}$ für $0 < \alpha < 1$.

3. Sei $f : D_f = [0,4]\setminus\{3\} \to \mathbb{R}$, $f(x) := \frac{1}{(x-3)^2} \quad \forall\, x \in D_f$, so ist f stetig, in jedem abgeschlossen Intervall $\subset D_f$ R-integrierbar und in 3 unbeschränkt. Betrachte

$$\int_0^3 \frac{1}{(x-3)^2}\,\mathrm{d}x = \lim_{c\nearrow 3}\int_0^c \frac{1}{(x-3)^2}\,\mathrm{d}x = \lim_{c\nearrow 3}\left(-\frac{1}{x-3}\right)\Bigg|_0^c = +\infty$$

Daraus folgt, dass $\int_0^4 \frac{1}{(x-3)^2}\,\mathrm{d}x$ divergent ist.

Achtung.

$$\int_0^4 \frac{1}{(x-3)^2}\,\mathrm{d}x = \left(-\frac{1}{x-3}\right)\Bigg|_0^4 = -\frac{1}{4-3} + \frac{1}{0-3} = -1 - \frac{1}{3} = -\frac{4}{3}$$

ist **falsch**!

5.4.10 Integrand und Integrationsbereich unbeschränkt

Bemerkung. Es sollte nun klar sein, wie die Konvergenz eines uneigentlichen Integrals bei gleichzeitigem Auftreten von unbeschränktem Integranden und unbeschränktem Integrationsbereich definiert sind.

Beispiel

$$\int_0^\infty \frac{1}{x^\alpha}\,\mathrm{d}x \text{ konvergent} \iff \int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x \text{ und } \int_1^\infty \frac{1}{x^\alpha}\,\mathrm{d}x \text{ konvergent}$$

$\int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x$ konvergiert für alle $0 < \alpha < 1$, aber es ist

$$\int_1^\infty \frac{1}{x^\alpha}\,\mathrm{d}x = \lim_{c\to\infty} \frac{x^{-\alpha+1}}{1-\alpha}\Bigg|_1^c = \lim_{c\to\infty}\left(\frac{c^{-\alpha+1}}{1-\alpha}\right) = +\infty \quad \forall\, 0 < \alpha < 1$$

also ist $\int\limits_0^\infty \frac{1}{x^\alpha}\,\mathrm{d}x$ divergent für alle $\alpha > 0$.

6 Topologische Grundlagen

6.1 Metrische und normierte Räume

6.1.1 Normierte lineare Räume

Definition 6.1.1. Sei X ein $\{\mathbb{R}, \mathbb{C}\} \ni \mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : X \to \mathbb{R}$ mit den Eigenschaften

i) $\|x\| \geq 0 \quad \forall\, x \in X$ und $\|x\| = 0 \Leftrightarrow x = 0$ (Definitheit)

ii) $\|\alpha \cdot x\| = |\alpha|\, \|x\| \quad \forall\, \alpha \in \mathbb{R},\, x \in X$ (Homogenität)

iii) $\|x + y\| \leq \|x\| + \|y\| \quad \forall,\, x, y \in X$ (Dreiecksungleichung)

heißt **Norm** auf X; $(X, \|\cdot\|)$ heißt **normierter linearer Raum**.

6.1.2 Beispiele

1. $X = \mathbb{R}^n$, $n \in \mathbb{N}$. Für alle $1 \leq p < \infty$ definiert

$$\|x\|_p = \|(x_1, \dots, x_n)\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

eine Norm auf $\mathbb{R}^n$, die sogenannte p**-Norm** (die Dreiecksungleichung folgt direkt aus der Minkowski-Ungleichung). Im Spezialfall $p = 2$ erhält man

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

also die Standardnorm auf $\mathbb{R}^n$, genannt die **Euklidische Norm**.

Ebenso definiert

$$\|x\|_\infty = \|(x_1, \dots, x_n)\|_\infty := \max_{i=1,\dots,n} |x_i|, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

eine Norm auf $\mathbb{R}^n$, die sogenannte **Supremums-** oder **Maximumsnorm**.

2. $X = C([a, b]) = \{f :\ [a, b] \to \mathbb{R} \mid f \text{ stetig}\}$ ist ein $\mathbb{R}$-Vektorraum und

$$\|f\|_\infty = \max_{t \in [a,b]} |f(t)|, \quad f \in X$$

ist Norm auf X, die sogenannte **Supremums-** oder **Maximumsnorm**. Auch

$$\|f\|_p = \left(\int_a^b |f(t)| \,\mathrm{d}t \right)^{\frac{1}{p}}, \quad f \in X$$

definiert für $1 \leq p < \infty$ eine Norm auf X.

6.1.3 Metrische Räume

Definition 6.1.2. Sei X eine nicht-leere Menge. Eine Abbildung $d: X \times X \to \mathbb{R}$ mit den folgenden Eigenschaften

i) $d(x,y) \geq 0 \quad \forall\, x,y \in X$ und $d(x,y) = 0 \Leftrightarrow x = y$ (Definitheit)

ii) $d(x,y) = d(y,x) \quad \forall\, x,y \in X$ (Symmetrie)

iii) $d(x,y) \leq d(x,z) + d(z,y) \quad \forall\, x,y,z \in X$ (Dreiecksungleichung)

heißt **Metrik** auf X; (X,d) heißt **metrischer Raum**.

Wie im Skript zu Analysis I schon angemerkt, genügt es bei der Definition von Norm und Metrik jeweils in i) die Definitheit, d.h. $\|x\| = 0 \Leftrightarrow x = 0$ bzw. $d(x,y) = 0 \Leftrightarrow x = y$ zu fordern. Die Nichtnegativität folgt aus den drei Gesetzen.

6.1.4 Beispiele

1. $X = \mathbb{R}$, $d(x,y) := |x-y| \quad \forall\, x,y \in \mathbb{R}$ ist eine Metrik auf $\mathbb{R}$. Allgemeiner: Ist $(X, \|\cdot\|)$ ein normierter linearer Raum, dann definiert

$$\begin{aligned} d:\ & X \times X \to \mathbb{R} \\ & (x,y) \mapsto \|x-y\| \end{aligned}$$

eine Metrik auf X. D.h. jeder normierte lineare Raum ist in natürlicher Weise auch ein metrischer Raum.

Andersherum kann man aus einer Metrik auf einem $\mathbb{R}$-Vektorraum X eine Norm zurückgewinnen, falls gilt

- *d ist translationsinvariant, d.h. es gilt*

$$d(x-z, y-z) = d(x,y) \quad \forall\, x,y,z \in X$$

- *d ist homogen, d.h. es gilt*

$$d(\alpha x, \alpha y) = |\alpha|\, d(x,y) \quad \forall\, \alpha \in \mathbb{R},\ x,y \in X$$

2. Ist $\emptyset \neq S \subseteq X$, (X,d) ein metrischer Raum, dann definiert

$$\begin{aligned} d_S:\ & S \times S \to \mathbb{R} \\ & (x,y) \mapsto d(x,y) \end{aligned}$$

eine Metrik auf S. d_s heißt **von d auf S induzierte Metrik** (oder auch **Spurmetrik**).

3. Sei $\emptyset \neq X$ eine Menge. Dann definiert

$$\begin{aligned} d:\ & X \times X \to \mathbb{R} \\ & (x,y) \mapsto \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases} \end{aligned}$$

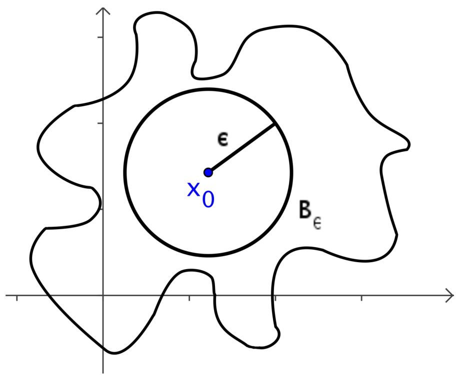
eine Metrik, die sogenannte **diskrete Metrik** auf X.

6.1.5 Kugeln, Umgebungen, offen, abgeschlossen

Definition 6.1.3. (X, d) sei ein metrischer Raum, $x_0 \in X$.

i) Für $\epsilon > 0$ heißt $B_\epsilon(x_0) : \{x \in X \mid d(x, x_0) < \epsilon\}$ **offene Kugel** mit Radius ϵ um x_0.

ii) $U \subseteq X$ heißt **Umgebung** von x_0, falls ein $\epsilon > 0$ existiert mit $B_\epsilon(x_0) \subseteq U$.



iii) $O \subseteq X$ heißt **offen**, wenn O Umgebung eines jeden Punktes $x \in O$ ist.

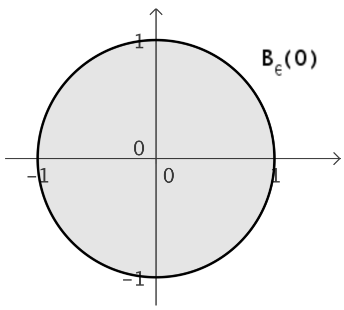
iv) $A \subseteq X$ heißt **abgeschlossen**, wenn $X \setminus A$ offen ist.

6.1.6 Beispiele

1. i) Betrachte $X = \mathbb{R}^2$ mit $d_2(x, y) = \|x - y\|_2 \quad \forall x, y \in \mathbb{R}$, also der durch die Euklidische Norm induzierte Euklidische Metrik.

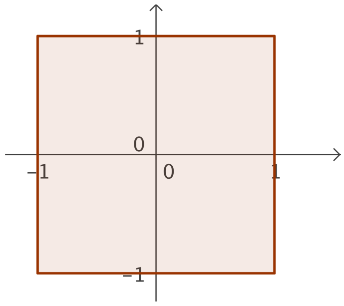
$$B_1(0) = \left\{ x = (x_1, x_2) \in \mathbb{R}^2 \mid \sqrt{x_1^2 + x_2^2} < 1 \right\}$$

ist die **Einheitskugel** in $X = \mathbb{R}^2$.



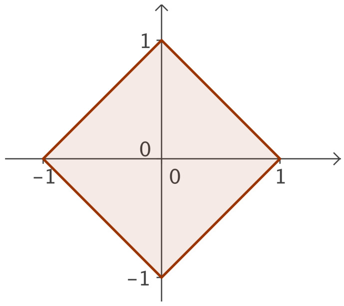
ii) $d_\infty(x, y) := \|x - y\|_\infty$

$$B_1(0) = \left\{ x = (x_1, x_2) \in \mathbb{R}^2 \mid \max\{x_1, x_2\} < 1 \right\}$$



iii) $d_1(x,y) := \|x-y\|_1$

$$B_1(0) = \left\{x = (x_1,x_2) \in \mathbb{R}^2 \mid |x_1| + |x_2| < 1\right\}$$



(In den Abbildungen gehören die Ränder jeweils nicht zur Kugel.)

2. Die offenen ϵ-Kugeln sind offen. Denn sei $y \in B_\epsilon(x_0)$, dann definiere $\epsilon_0 := d(y,x_0) < \epsilon$. Mit $\delta := \epsilon - \epsilon_0$ gilt für $z \in B_\delta(y)$

$$d(z,x_0) \leq d(z,y) + d(y,x_0) < \epsilon - \epsilon_0 + \epsilon_0 = \epsilon$$

also $B_\delta(y) \subseteq B_\epsilon(x_0)$.

3. Bezüglich der diskreten Metrik auf einer Menge X sind alle Teilmengen von X sowohl offen als auch abgeschlossen.

4. In jedem metrischen Raum (X,d) sind $\emptyset$ und X sowohl offen als auch abgeschlossen.

6.1.7 Metriken induzieren Topologien

Satz. Sei (X,d) ein metrischer Raum.

i) Die Vereinigung von beliebig vielen offenen Teilmengen und der Durchschnitt von endlich vielen offene Teilmengen von X ist offen.

ii) Die Vereinigung von endlich vielen abgeschlossenen Teilmengen und der Durchschnitt von beliebig vielen abgeschlossenen Teilmengen von X ist wieder abgeschlossen.

Beweis.

i) – Seien O_i offene Teilmengen von X mit I beliebige Indexmenge. Zu zeigen ist

$$\bigcup_{i\in I} O_i = \bigcup O_i \text{ offen}$$

Sei dazu $x \in \bigcup O_i$, dann existiert ein $i_0 \in I$ mit $x \in O_{i_0}$. Da O_{i_0} offen, existiert ein $\epsilon > 0$, sodass $B_\epsilon(x) \subseteq O_{i_0} \subseteq \bigcup O_i$, also ist $\bigcup O_i$ offen.

– Seien $O_1, \dots, O_n$ endlich viele offene Teilmengen von X. Zu zeigen ist

$$\bigcap_{i=1}^{n} O_i = \bigcap O_i \text{ offen}$$

Sei dazu $x \in \bigcap O_i$, es ist dann $x \in O_i \quad \forall\, i = 1, \dots, n$. Da jedes O_i offen ist, existieren

$$B_{\epsilon_i}(x) \subseteq O_i \quad \forall\, i = 1, \dots, n$$

Mit $\epsilon := \min\limits_{i=1,\dots,n} \epsilon_i$ folgt

$$\begin{aligned} &B_\epsilon(x) \subseteq B_{\epsilon_i} \subseteq O_i \quad \forall\, i = 1, \dots, n && \Rightarrow \\ &B_\epsilon(x) \subseteq \bigcap_{i=1}^{n} O_i && \Rightarrow \\ &\bigcap_{i=1}^{n} O_i \text{ offen} \end{aligned}$$

ii) Folgt aus i) durch Komplementbildung:

$$\bigcup_{i=1}^{n} A_i = X \setminus \left(\bigcap_{i=1}^{n} \underbrace{(X \setminus A_i)}_{\text{offen}} \right) \qquad \bigcap_{i\in I} A_i = X \setminus \left(\bigcup_{i\in I} \underbrace{(X \setminus A_i)}_{\text{offen}} \right)$$

□

6.1.8 Beispiele

Der Durchschnitt von beliebig vielen offenen Mengen ist im Allgemeinen nicht offen. Die Vereinigung von beliebig vielen abgeschlossenen Teilmengen ist im Allgemeinen nicht abgeschlossen.

- $O_n := \left]-\frac{1}{n}, \frac{1}{n}\right[$ ist für jedes $n \in \mathbb{N}$ offen, aber $\bigcap\limits_{n\in\mathbb{N}} O_n = \{0\}$ ist nicht offen.
- $A_n := \left[\frac{1}{n}, 1\right]$ ist für jedes $n \in \mathbb{N}$ abgeschlossen, aber $\bigcup\limits_{n\in\mathbb{N}} A_n =]0,1]$ ist nicht abgeschlossen.

6.1.9 Metrische Räume sind hausdorffsch

Lemma 6.1.4. Sei (X,d) ein metrischer Raum. Sind $x, y \in X$, $x \neq y$, dann gibt es Umgebungen U von x und V von y mit $U \cap V = \emptyset$.

Man sagt: „Verschiedene Punkte besitzen disjunkte Umgebungen oder lassen sich durch disjunkte Umgebungen trennen.“

Beweis. Seien $x, y \in X$, $x \neq y$. Definiere $d_0 := d(x,y) > 0$, so sind $U := B_\epsilon(x), V := B_\epsilon(y)$ für $0 < \epsilon < \frac{d_0}{2}$ disjunkte Umgebungen. $\square$

6.1.10 Topologie

Bemerkung. Wenn X eine beliebige Menge ist, $\mathcal{T} \subseteq \mathcal{P}(X)$ ein System von Teilmengen mit

i) $\emptyset, X \in \mathcal{T}$,

ii) Endliche Durschnitte und beliebige Vereinigungen von Teilmengen in $\mathcal{T}$ liegen wieder in $\mathcal{T}$.

dann nennt man $\mathcal{T}$ eine **Topologie** auf X $(X, \mathcal{T})$ einen **topologischen Raum**. Die Mengen in $\mathcal{T}$ nennt man die **offenen (Teil-)Mengen** in X.

Eine Teilmenge U von X heißt **Umgebung** eines Punktes $x \in X$, falls es eine offene Teilmenge $O \in \mathcal{T}$ gibt mit $x \in O \subseteq U$.

$(X, \mathcal{T})$ heißt **hausdorffsch** , wenn verschiedene Punkte disjunkte Umgebungen besitzen. Ein metrischer Raum ist also ein hausdorffscher topologischer Raum.

6.1.11 Offene Mengen in Teilräumen

Satz 6.1.5. Sei (S, d_S) ein Teilraum des metrischen Raums (X,d). $O \subseteq S$ ist offen in $(S, d_S) \iff \exists O'$ offen in (X,d) mit $O = O' \cap S$. Eine solche Teilmenge O von X heißt **relativ S-offen**.

Analoges für abgeschlossene Mengen.

Beweis. Sei $O' \subseteq X$ offen mit $O = O' \cap S$. Zu zeigen ist, dass O in (S, d_S) offen ist. Sei dazu $x \in O = O' \cap S$; da O' in X offen ist, existiert ein $\epsilon > 0$ mit

$$\begin{aligned} B_\epsilon^X(x) = \{x \in X \mid d(x,y) < \epsilon\} \subseteq O' \quad & \Rightarrow \\ B_\epsilon^X(x) \cap S \subseteq O' \cap S = O \quad & \Leftrightarrow \\ \{y \in S \mid d(x,y) < \epsilon\} \subseteq O \quad & \Leftrightarrow \\ \{y \in S \mid d_S(x,y) < \epsilon\} \subseteq O \quad & \Leftrightarrow \\ B_\epsilon^S(x) \subseteq O \quad & \end{aligned}$$

D.h. O ist offen in S.

Sei nun umgekehrt O offene Teilmenge von (S, d_S). Dann gilt

$$\forall\, x \in O \quad \exists\, \epsilon_x > 0 \;:\; B_{\epsilon_x}^S(x) \subseteq O \qquad \Rightarrow$$

$$O = \bigcup_{x \in O} B_{\epsilon_x}^S(x) = \bigcup_{x \in O} \left(B_{\epsilon_x}^X(x) \cap S\right) = \underbrace{\left(\bigcup_{x \in O} B_{\epsilon_x}^X(x)\right)}_{=:O'} \cap S$$

O' ist eine offene Teilmenge von X als Vereinigung offener Teilmengen. $\square$

6.1.12 Beispiele

Betrachte $X = \mathbb{R}$ ausgestattet mit der Standardmetrik $d(x,y) = |x-y|$, $x,y \in \mathbb{R}$ und $S :=]0,1[$ ausgestattet mit der Spurmetrik d_S. Dann ist etwa

- $A = \left[\frac{1}{2}, 1\right[$ abgeschlossene Teilmenge in (S, d_S) und
- $O = \left]0, \frac{1}{4}\right[$ offene Teilmenge in (S, d_S).

Achtung. A ist nicht abgeschlossen in (X,d), nur relativ S-abgeschlossen.

6.1.13 Inneres, Abschluss, Rand; Häufungs- und Berührpunkte

Definition 6.1.6. Sei (X,d) ein metrischer Raum, $S \subseteq X$.

i) Die Vereinigung aller offenen Teilmengen von S:

$$\mathring{S} = \operatorname{int}(S) := \bigcup_{O \subseteq S \text{ offen}} O$$

heißt **Inneres** von S ($\mathring{S}$ ist offensichtlich die größte offene Teilmenge von S).

ii) Der Durchschnitt aller abgeschlossenen Obermengen von S:

$$\overline{S} = \operatorname{cl}(S) := \bigcap_{A \supseteq S \text{ abgeschlossen}} A$$

heißt **abgeschlossene Hülle** oder **Abschluss** von S ($\overline{S}$ ist offensichtlich die kleinste abgeschlossene Obermenge von S).

iii) Der **Rand** von S ist definiert als $\partial S := \overline{S} \setminus \mathring{S}$. Ein Punkt $s \in \partial S$ **Randpunkt** von S.

iv) $s \in S$ heißt **innerer Punkt** von S, falls ein $\epsilon > 0$ existiert mit $B_\epsilon(x) \subseteq S$.

v) $x \in X$ heißt **Häufungspunkt** von S, wenn

$$\forall \epsilon > 0 \, : \, B_\epsilon(x) \cap (S \setminus \{x\}) \neq \emptyset$$

vi) $x \in X$ heißt **Berührpunkt** von S, wenn

$$\forall \epsilon > 0 \, : \, B_\epsilon(x) \cap S \neq \emptyset$$

Lemma 6.1.7. Sei (X,d) ein metrischer Raum, $S \subseteq X$. Es gilt

i) $\mathring{S} = \{s \in s \mid s \text{ innerer Punkt}\}$

ii) $\overline{S} = \{s \in s \mid s \text{ Berührpunkt von } S\}$

iii) $\partial S = \{s \in s \mid s \text{ Berührpunkt von } S \text{ und von } X \setminus S\}$ und ∂S ist abgeschlossen.

Beweis.

i) ist trivial.

ii) Sei $M := \{s \in s \mid s \text{ Berührpunkt von } S\}$. Zeige zunächst $M \subseteq \overline{S}$:

Falls $\overline{S} = X$ ist nichts zu zeigen. Sei also $x \in X \setminus \overline{S}$. Da $\overline{S}$ abgeschlossen ist, ist $X \setminus \overline{S}$ offen, d.h.

$$\exists \epsilon > 0 \, : \, B_\epsilon(x) \subseteq X \setminus \overline{S} \subseteq X \setminus S$$

x ist also kein Berührpunkt von S, d.h. $x \in X \setminus M$. Damit ist $X \setminus \overline{S} \subseteq X \setminus M$, also $M \subseteq \overline{S}$ gezeigt.

Zeige $\overline{S} \subseteq M$: O.B.d.A. sei $M \neq X$. Sei $x \in \overline{S}$, d.h. $x \in A$ für alle abgeschlossenen Mengen $A \supseteq S$. Sei $\epsilon > 0$ gegeben. Angenommen x sei kein Berührpunkt, d.h. $B_\epsilon(x) \cap S = \emptyset$, dann ist $B_\epsilon(x) \subseteq X \setminus S$ bzw. $S \subseteq X \setminus B_\epsilon(x)$. $X \setminus B_\epsilon(x)$ ist damit aber abgeschlossene Obermenge von S, die x enthalten müsste, Widerspruch.

iii) Es ist $\partial S = \overline{S} \setminus \mathring{S} = \overline{S} \cap \left(X \setminus \mathring{S} \right)$ und damit also Schnitt abgeschlossener Mengen abgeschlossen. Der Rest der Behauptung folgt ebenso unmittelbar aus dieser Darstellung.

$\square$

6.2 Folgen und Konvergenz in metrischen Räumen

6.2.1 Konvergenz

Definition 6.2.1. Sei (X, d) ein metrischer Raum.

i) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in X und sei $x \in X$. Dann heißt $(x_n)_{n\in\mathbb{N}}$ konvergent gegen x in (X, d), wenn

$$\lim_{n\to\infty} d(x_n, x) = 0 \qquad \Leftrightarrow$$
$$\forall\, \epsilon > 0 \quad \exists\, n_0 \in \mathbb{N} \, : \, d(x_n, x) < \epsilon \quad \forall\, n \geq n_0$$

Notation:

$$x_n \to x, \; x_n \underset{d}{\to} x, \; (\text{d}-)\lim_{n\to\infty} x_n = x$$

$(x_n)_{n\in\mathbb{N}}$ heißt **konvergent**, wenn $x \in X$ existiert mit $x_n \to x$. Andernfalls heißt $(x_n)_{n\in\mathbb{N}}$ **divergent**.

ii) Eine Folge $(x_n)_{n\in\mathbb{N}}$ in X heißt **Cauchy-Folge** in (X, d), falls

$$\forall\, \epsilon > 0 \quad \exists\, n_0 \in \mathbb{N} \quad \forall\, n, m \geq n_0 \, : \, d(x_n, x_m) < \epsilon$$

Bemerkung.

1. Wenn $(X, \|\cdot\|)$ ein normierter linearer Raum ist, $(x_n)_{n\in\mathbb{N}} \subseteq X$, $x \in X$, dann konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen x in $(X, \|\cdot\|)$ genau dann, wenn $x_n \underset{d}{\to} x$ in X bezüglich der von der Norm induzierten Metrik d. Notation: $x_n \underset{\|\cdot\|}{\to} x$.

2. In metrischen Räumen (X, d) kann Konvergenz von Folgen äquivalent nur mit Hilfe des Umgebungsbegriffs definiert werden:

Ist $(x_n)_{n\in\mathbb{N}}$ Folge in X und $x \in X$, dann

$$x_n \underset{d}{\rightarrow} x \quad \Leftrightarrow \quad \forall\, U(x) \quad \exists\, n_0 \in \mathbb{N} \,:\, x_n \in U(x) \quad \forall\, n \geq n_0$$

wobei $U(x)$ eine Umgebung von x bezeichnet.

Beweis. Es gelte zunächst $x_n \underset{d}{\rightarrow} x$. Sei U Umgegbung von x, dann existiert ein $\epsilon > 0$ mit $B_\epsilon(x) \subseteq U$. Da $x_n \underset{d}{\rightarrow} x$, existiert ein $n_0 \in \mathbb{N}$ mit $x_n \in B_\epsilon(x)$ für alle $n \geq n_0$, also $x_n \in U \quad \forall\, n \geq n_0$.

Die andere Richtung ist trivial, da $B_\epsilon(x)$ Umgebung von x für alle $\epsilon > 0$. □

Die Definition der Konvergenz mit Hilfe des Umgebungsbegriffs lässt sich auf allgemeine topologische Räume erweitern.

3. Wie in Analysis I zeigt man, dass konvergente Folgen in (X, d) Cauchy-Folgen sind:

$$x_n \underset{d}{\rightarrow} x \Rightarrow d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < 2\epsilon \quad \forall\, n, m \geq n_0$$

für in Abhängigkeit von ϵ hinreichend groß gewähltes $n_0 \in \mathbb{N}$.

Aber: Im Allgemeinen sind Cauchy-Folgen in einem metrischen Raum (X, d) nicht notwendigerweise konvergent.

6.2.2 Vollständigkeit

Definition 6.2.2.

i) Ein metrischer Raum (X, d) heißt **vollständig**, wenn jede Cauchy-Folge in X konvergiert.

ii) Ist $(X, \|\cdot\|)$ ein normierter linearer Raum und X bezüglich der von der Norm induzierten Metrik d vollständig, dann heißt $(X, \|\cdot\|)$ **Banachraum**.

iii) Eine Teilmenge N eines metrischen Raums (X, d) heißt vollständig, falls (N, d_N) vollständig ist.

6.2.3 Beispiele

1. $(\mathbb{R}, |\cdot|)$ ist ein Banachraum, da nach Analysis I Cauchy-Folgen in $\mathbb{R}$ konvergieren.

2. $(C([a, b]), \|\cdot\|_\infty)$ $(a, b \in \mathbb{R}, a < b)$ ist ein Banachraum.

Beweis. Es ist zu zeigen: $(C([a, b]), d_\infty)$ ist vollständig.

Sei dazu $(f_n)_{n\in\mathbb{N}}$ Cauchy-Folge bezüglich der d_∞-Metrik, d.h.

$$\forall\, \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \,:\, d_\infty(f_n, f_m) = \|f_n - f_m\|_\infty < \epsilon \quad \forall\, n, m \geq n_0 \Leftrightarrow$$
$$\forall\, \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \,:\, \max_{x\in[a,b]} |f_n(x) - f_m(x)| < \epsilon \quad \forall\, n, m \geq n_0 \Leftrightarrow$$
$$\forall\, \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \,:\, |f_n(x) - f_m(x)| < \epsilon \quad \forall\, n, m \geq n_0,\, x \in [a, b]$$

Dies ist aber das Cauchy-Kriterium für gleichmäßige Konvergenz der Funktionfolge $(f_n)_{n\in\mathbb{N}}$ aus Analysis I, d.h. $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen eine stetige Funktion f auf $[a,b]$, also

$$\forall\,\epsilon>0\quad \exists\,n_0\in\mathbb{N}\,:\,|f_n(x)-f(x)|<\epsilon\quad \forall\,n\geq n_0,\,x\in[a,b]\Leftrightarrow$$
$$\forall\,\epsilon>0\quad \exists\,n_0\in\mathbb{N}\,:\,\|f_n-f\|_\infty<\epsilon\quad \forall\,n\geq n_0$$

und somit $f_n \underset{\|\cdot\|_\infty}{\to} f$ in $C([a,b])$. □

6.2.4 Konvergenz und Vollständigkeit von R bezüglich p-Normen

Was ist mit $(\mathbb{R}^N,\|\cdot\|_p)$?

Lemma 6.2.3. Sei $X=\mathbb{R}^N$, $N\geq 1$ ausgestattet mit der p-Norm $\|\cdot\|_p$ $(1\leq p\leq\infty)$, $(x_n)_{n\in\mathbb{N}}\subseteq X$, $x\in X$. Dann gilt

i) $(x_n)_{n\in\mathbb{N}}$ ist Cauchy-Folge in $(\mathbb{R}^N,\|\cdot\|_p)$ $\iff$ $(x_{ni})_{n\in\mathbb{N}}$ Cauchy-Folge in $(\mathbb{R},|\cdot|)$ für alle $i=1,\ldots,N$.

ii) $x_n \underset{\|\cdot\|_p}{\to} x$ in $\mathbb{R}^N$ $\iff$ $x_{ni}\underset{|\cdot|}{\to} x_i$ in $\mathbb{R}$ für alle $i=1,\ldots,N$.

(Dabei bezeichnet $x=(x_1,\ldots,x_N)$ und $x_n=(x_{n1},\ldots,x_{nN})$.)

Beweis. Vorüberlegung: Ist $x\in\mathbb{R}^N$, $1\leq p<\infty$, dann gilt für alle $i=1,\ldots,N$:

$$\begin{aligned}|x_i| &\leq \left(\sum_{i=1}^N |x_i|^p\right)^{\frac{1}{p}} = \|x\|_p\\ &\leq \left(\sum_{i=1}^N\left(\max_{i=1,\ldots,N}|x_i|\right)^p\right)^{\frac{1}{p}} = \left(N\cdot\left(\max_{i=1,\ldots,N}|x_i|\right)^p\right)^{\frac{1}{p}} = N^{\frac{1}{p}}\cdot\|x\|_\infty\end{aligned}$$

Da $i=1,\ldots,N$ beliebig, folgt

$$\|x\|_\infty\leq\|x\|_p\leq N^{\frac{1}{p}}\cdot\|x\|_\infty$$

Zum eigentlichen Beweis:

i) Sei zunächst Cauchy-Folge in $(\mathbb{R}^N,\|x\|_p)$. Es gilt für alle $i=1,\ldots,N$

$$|x_{ni}-x_{mi}|\leq\|x_n-x_m\|_\infty\leq\|x_n-x_m\|_p\quad\forall\,n,m\in\mathbb{N}$$

wonach sofort folgt, dass $(x_{ni})_{n\in\mathbb{N}}$ Cauchy-Folge in $(\mathbb{R},|\cdot|)$ ist.
Sei $(x_{ni})_{n\in\mathbb{N}}$ für jedes $i=1,\ldots,N$ Cauchy-Folge, d.h. es gilt

$$\forall\,i=1,\ldots,N\quad\forall\,\epsilon>0\quad\exists\,n_0(i)\in\mathbb{N}\,:\,|x_{ni}-x_{mi}|<\epsilon\quad\forall\,n,m\geq n_0(i)$$

Bei gegebenen $\epsilon>0$ folgt für $n_0:=\max_{i=1,\ldots,N}n_0(i)$

$$|x_{ni}-x_{mi}|<\epsilon\quad\forall\,n,m\geq n_0,\,i=1,\ldots,N$$

d.h. $\|x_n-x_m\|_\infty<\epsilon$ für alle $n,m\geq n_0$, also ist $(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge in $\mathbb{R}^N$ bezüglich $\|x\|_\infty$. Aus der Vorüberlegung folgt die Betrachtung auch für $1\leq p<\infty$.

ii) verläuft analog.

$\square$

Korollar 6.2.4. Für $N \geq 1$, $1 \leq p \leq \infty$ ist $(\mathbb{R}^N, \|\cdot\|_p)$ ein Banachraum.

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge in $(\mathbb{R}^N, \|\cdot\|_p)$. Nach Lemma ist $(x_{ni})_{n\in\mathbb{N}}$ Cauchy-Folge in $\mathbb{R}$ für jede Koordinate $i = 1, \dots, N$, also wegen der Vollständigkeit von $(\mathbb{R}, |\cdot|)$ konvergent. Wieder nach Lemma konvergiert also $(x_n)_{n\in\mathbb{N}}$. $\square$

Merke: Konvergenz in $\mathbb{R}^N$ ist „koordinatenweise Konvergenz"..

Bemerkung. Zwei Normen $\|\cdot\|_1, \|\cdot\|_2$ auf einem Vektorraum X bezeichnet man als **äquivalent**, falls $c, d > 0$ existieren mit

$$c \cdot \|x\|_1 \leq \|x\|_2 \leq d \cdot \|x\|_1 \quad \forall\, x \in X$$

Sind $\|\cdot\|_1, \|\cdot\|_2$ äquivalente Normen auf X, dann besitzen $(X, \|x\|_1)$ und $(X, \|x\|_2)$ die gleichen Cauchy-Folgen und die gleichen konvergenten Folgen, d.h. $(x_n)_{n\in\mathbb{N}}$ ist Cauchy-Folge/konvergent in $(X, \|\cdot\|_1) \iff (x_n)_{n\in\mathbb{N}}$ Cauchy-Folge/konvergent in $(X, \|\cdot\|_2)$.

Wir haben also gesehen: Alle $\|\cdot\|_p$-Normen auf $\mathbb{R}^N$ $1 \leq p \leq \infty$ sind äquivalent.

6.2.5 Charakterisierung topologischer Eigenschaften über Folgen

In metrischen Räumen können topologische Begriffe wie Häufungspunkt, Berührpunkt, abgeschlossene Hülle etc. mit Hilfe von Folgen beschrieben werden.

Lemma 6.2.5. Sei (X, d) ein metrischer Raum, $S \subseteq X$.

i) $x \in X$ Häufungspunkt von $S \iff$ es gibt eine Folge $(x_n)_{n\in\mathbb{N}} \subseteq (S \setminus \{x\})$ mit $x_n \underset{d}{\to} x$.

ii) $x \in X$ Berührpunkt von $S \iff$ es gibt eine Folge $(x_n)_{n\in\mathbb{N}} \subseteq S$ mit $x_n \underset{d}{\to} x$.

iii) $\overline{S} = \left\{ x \in X \mid \exists\, (x_n)_{n\in\mathbb{N}} \subseteq S \,:\, x_n \underset{d}{\to} x \right\}$. Insbesondere gilt:

$$S \text{ ist abgeschlossen} \iff \forall\, (x_n)_{n\in\mathbb{N}} \subseteq S \,:\, x_n \underset{d}{\to} x \in X \Rightarrow x \in S$$

D.h. S ist **folgen-abgeschlossen**.

Beweis.

i) Sei $x \in X$ Häufungspunkt von S, d.h.

$$\begin{aligned} \forall\, \epsilon > 0 \,:\, B_\epsilon(x) \cap (S \setminus \{x\}) \neq \emptyset \qquad & \Rightarrow \\ \forall\, n \in \mathbb{N} \,:\, B_{\frac{1}{n}}(x) \cap (S \setminus \{x\}) \neq \emptyset \end{aligned}$$

Wähle also für jedes $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(x) \cap (S \setminus \{x\})$. Dann ist $(x_n)_{n\in\mathbb{N}} \subseteq (S \setminus \{x\})$ und $d(x_n, x) < \frac{1}{n} \quad \forall n \in \mathbb{N}$, also $x_n \underset{d}{\to} x$.

Sei nun $(x_n)_{n\in\mathbb{N}} \subseteq (S \setminus \{x\})$ mit $x_n \underset{d}{\to} x$. Sei dann $\epsilon > 0$. Da $x_n \underset{d}{\to} d$, existiert zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $x_n \in B_\epsilon(x)$ für alle $n \geq n_0$. Da $x_n \in (S \setminus \{x\})$ für alle $n \in \mathbb{N}$, folgt $B_\epsilon(x) \cap S \neq \emptyset$ für alle $n \geq n_0$.

ii) verläuft analog.

iii) Die Mengengleichheit folgt aus Lemma 6.1.7 und ii), der Zusatz aus S abgeschlossen $\iff S = \overline{S}$ und der Mengengleichheit.

□

6.2.6 Abgeschlossene Teilräume vollständiger Räume sind vollständig

Satz 6.2.6. Sei (X, d) vollständiger metrischer Raum und $S \subseteq X$. Dann gilt:
S ist vollständig (d.h. (S, d_S) ist vollständig) $\iff$ S ist abgeschlossen.

Beweis.

⇒ Es genügt nach Lemma 6.2.5 iii) zu zeigen, dass S folgen-abgeschlossen ist. Sei also $(x_n)_{n\in\mathbb{N}} \subseteq S$ mit $x_n \underset{d}{\to} x$. Als konvergente Folge ist $(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge in (X, d). Da $(x_n)_{n\in\mathbb{N}} \subseteq S$, folgt, dass $(x_n)_{n\in\mathbb{N}}$ auch in (S, d_S) Cauchy-Folge ist. Nach Voraussetzung ist (S, d_S) vollständig, also $x_n \underset{d_S}{\to} y \in S$. Damit gilt aber auch $x_n \underset{d}{\to} y$ und da der Grenzwert eindeutig ist also $x = y$, d.h. $x \in S$. S ist folglich folgen-abgeschlossen.

⇐ Sei $(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge in (S, d_S). Damit ist sie auch Cauchy-Folge in (X, d). Da (X, d) vollständig ist, konvergiert $(x_n)_{n\in\mathbb{N}}$ gegen ein $x \in X$. Und weil S folgen-abgeschlossen ist, folgt $x \in S$, womit sofort $x_n \underset{d_S}{\to} x \in S$ abfällt.

□

6.2.7 Banach'scher Fixpunktsatz

Satz 6.2.7. Sei (X, d) ein vollständiger metrischer Raum ($X \neq \emptyset$), $T : X \to X$ eine kontrahierende Abbildung, d.h.

$$\exists\, 0 < q < 1 \;:\; d(T(x), T(y)) \leq q \cdot d(x, y) \quad \forall\, x, y \in X$$

Dann gilt:

i) Es existiert genau ein Fixpunkt $\hat{x} \in X \;:\; T(\hat{x}) = \hat{x}$.

ii) Für alle $x_0 \in X$ gilt: Die Folge $(x_n)_{n\in\mathbb{N}}$, rekursiv definiert durch $x_{n+1} := T(x_n) \quad \forall n \geq 0$, konvergiert gegen $\hat{x}$ in (X, d) und es gilt die folgende Fehlerabschätzung

$$d(x_n, \hat{x}) \leq \frac{q^n}{1-q} \cdot d(x_0, x_1) \quad \forall\, n \geq 2$$

Beweis. Eindeutigkeit des Fixpunkts: Angenommen, $\hat{x}$ und $\hat{y} \in X$ sind Fixpunkte von T. Dann gilt

$$d(\hat{x}, \hat{y}) = d(T(\hat{x}), T(\hat{x})) \leq q \cdot d(\hat{x}, \hat{y}) \Leftrightarrow d(\hat{x}, \hat{y}) = 0 \Leftrightarrow \hat{x} = \hat{y}$$

Zur Existenz: Sei $x_0 \in X$, $x_{n+1} := T(x_n) \quad \forall\, n \in \mathbb{N}_0$. Dann gilt

$$d(x_n, x_{n+1}) \leq q^n \cdot d(x_0, x_1) \quad n \geq 1$$

Beweise dies duch vollständige Induktion: Für $n = 1$ gilt

$$d(x_1, x_2) = d(T(x_0), T(x_1)) \leq q \cdot d(x_0, x_1)$$

Induktionsschritt $n \to n + 1$:

$$\begin{aligned} d(x_{n+1}, x_{n+2}) &= d(T(x_n), T(x_{n+1})) \\ &\leq q \cdot d(x_n, x_{n+1}) \\ &\overset{\text{IV}}{\leq} q \cdot q^n \cdot d(x_0, x_1) = q^{n+1} \cdot d(x_0, x_1) \end{aligned}$$

Somit folgt für beliebige $n, k \geq 1$

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + \ldots + d(x_{n+k-1}, x_{n+k}) \\ &\leq q^n \cdot d(x_0, x_1) + \ldots + q^{n+k-1} \cdot d(x_0, x_1) \\ &= q^n \cdot \left(1 + \ldots + q^{k-1}\right) \cdot d(x_0, x_1) \\ &\leq q^n \cdot \sum_{i=0}^{\infty} q^i \cdot d(x_0, x_1) \\ &= q^n \cdot \frac{1}{1-q} \cdot d(x_0, x_1) \overset{n \to \infty}{\longrightarrow} 0 \end{aligned}$$

Somit ist gezeigt, dass $(x_n)_{n \in \mathbb{N}}$ Cauchy-Folge ist. Da (X, d) vollständig ist, konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen ein $\hat{x} \in X$.

Durch Übergang zum Limes für $k \to \infty$ in obiger Ungleichung erhält man

$$d(x_n, \hat{x}) \leq \frac{q^n}{1-q} \cdot d(x_0, x_1) \quad \forall\, n \geq 1$$

Hierbei wurde benutzt, dass $d(x_n, x_{n+k}) \overset{n \to \infty}{\longrightarrow} d(x_n, \hat{x})$; dies folgt sofort aus der Dreiecksungleichung:

$$|d(x_n, x_{n+k}) - d(x_n, \hat{x})| \leq d(x_{n+k}, \hat{x})$$

Es bleibt zu zeigen, dass $\hat{x}$ Fixpunkt von T ist. Betrachte dazu

$$\begin{aligned} d(\hat{x}, T(\hat{x})) &\leq d(\hat{x}, x_{n+1}) + d(x_{n+1}, T(\hat{x})) \\ &= d(\hat{x}, x_{n+1}) + d(T(x_n), T(\hat{x})) \\ &\leq d(\hat{x}, x_{n+1}) + q \cdot d(x_n, \hat{x}) \overset{n \to \infty}{\longrightarrow} 0 \end{aligned}$$

Also ist $d(\hat{x}, T(\hat{x})) = 0$, d.h. $\hat{x} = T(\hat{x})$. $\square$

Bemerkung.

1. Auf die Voraussetzungen, dass $0 < q < 1$ kann nicht verzichtet werden. Ist $q = 1$, so besitzt $f : X \to X$ im Allgemeinen keinen Fixpunkt. Betrachte zum Beispiel

$$\begin{aligned} f:\ \mathbb{R} \to \mathbb{R}, \quad x \mapsto x+1 & & \Rightarrow \\ |f(x) - f(y)| = |x+1-y-1| = 1 \cdot |x-y| & & \forall\, x, y \in \mathbb{R} \end{aligned}$$

aber f besitzt natürlich keinen Fixpunkt.

2. Auch die Vollständigkeit ist essentiell. Betrachte beispielsweise $X = \mathbb{R}\setminus\{0\}$ ausgestattet mit der von $\mathbb{R}$ induzierten Standardmetrik, und

$$\begin{aligned} f:\ \mathbb{R}\setminus\{0\} \to \mathbb{R}\setminus\{0\}, \quad x \mapsto \frac{x}{2} & & \Rightarrow \\ |f(x) - f(y)| = \left|\frac{x}{2} - \frac{y}{2}\right| = \frac{1}{2} \cdot |x-y| & & \forall\, x, y \in \mathbb{R}\setminus\{0\} \end{aligned}$$

f ist also kontrahierend, besitzt aber keinen Fixpunkt.

3. Die Voraussetzung, dass T kontrahierend ist, kann aber durch die allgemeinere Voraussetzung ersetzt werden, dass T^k kontrahierend ist für ein $k \in \mathbb{N}$. Denn dann gibt es nach Fixpunktsatz genau ein $\hat{x} \in X$ mit $f^m(\hat{x}) = \hat{x}$. Damit gilt

$$\begin{aligned} d(f(\hat{x}), \hat{x}) &= d(f(f^m(\hat{x})), f^m(\hat{x})) \\ &= d(f^m(f(\hat{x})), f^m(\hat{x})) \\ &\leq q \cdot d(f(\hat{x}), \hat{x}) \end{aligned}$$

woraus wegen $0 < q < 1$ folgt: $d(f(\hat{x}), \hat{x}) = 0 \Leftrightarrow f(\hat{x}) = \hat{x}$. Zur Eindeutigkeit betrachte man $x' \in X$ mit $f(x') = x'$. Dann gilt

$$d(x', \hat{x}) = d(f(x'), f(\hat{x})) = \ldots = d(f^m(x'), f^m(\hat{x})) \leq q \cdot d(x', \hat{x})$$

also $d(x', \hat{x}) = 0$ und damit $x' = \hat{x}$.

6.3 Kompaktheit

6.3.1 Definition

Definition 6.3.1. Sei (X, d) ein metrischer Raum, $S \subseteq X$.

i) Eine Familie $\mathcal{U} = (U_i)_{i\in I}$ von offenen Teilmengen von X, die S überdecken, d.h. $S \subseteq \bigcup_{i\in I} U_i$, heißt **offene Überdeckung** von S.

   Eine Familie $\mathcal{V}$ heißt **Teilüberdeckung** von S bezüglich $\mathcal{U}$, wenn $\mathcal{V} \subseteq \mathcal{U}$ und $S \subseteq \bigcup_{U_i\in\mathcal{V}} U_i$, d.h. es existiert $K \subseteq I$, sodass $S \subseteq \bigcup_{i\in K} U_i$.

   Die Teilüberdeckung $\mathcal{V}$ heißt endlich, wenn K endlich ist.

ii) Die Teilmenge S heißt **kompakt**, wenn jede offene Überdeckung von S eine endliche Teilüberdeckung besitzt.

iii) S heißt **relativ kompakt**, wenn $\overline{S}$ kompakt ist.

6.3.2 Beispiele

1. Sei (X,d) ein metrischer Raum, $(x_n)_{n\in\mathbb{N}} \subseteq X$, $x \in X$ mit $x_n \underset{d}{\to} x$. Dann ist $S := \{x_n \mid n \in \mathbb{N}\} \cup \{x\}$ kompakt.

   Beweis. Sei $\mathcal{U} = (U_i)_{i\in I}$ eine offene Überdeckung von S, also $S \subseteq \bigcup_{i\in I} U_i$. Insbesondere gilt
   $$x \in \bigcup_{i\in I} U_i$$
   d.h. $\exists\, i_0 \in I \,:\, x \in U_{i_0}$. Da $x_n \underset{d}{\to} x$ und U_{i_0} offen gibt es ein $n_0 \in \mathbb{N}$
   $$x_n \in U_{i_0} \quad \forall\, n \geq n_0$$
   Seien $U_{i_1}, \dots, U_{i_{n_0}}$ die offenen Mengen, die die endliche vielen Folgenglieder $x_1, \dots, x_{n_0-1}$ enthalten. Dann ist $\mathcal{V} = \left\{U_{i_1}, \dots, U_{i_{n_0}}, U_{i_0}\right\}$ eine endliche Teilüberdeckung von S. □

2. Sei $X = \mathbb{R}$, $S = \mathbb{N}$ ausgestattet mit der Standardmetrik. S ist nicht kompakt, denn
   $$\mathcal{U} = \left( ]n - \frac{1}{2}, n + \frac{1}{2}[\right)_{n\in\mathbb{N}}$$
   ist offene Überdeckung von S, die keine endliche Teilüberdeckung von S enthält.

6.3.3 Kompaktheit und Abgeschlossenheit

Satz 6.3.2. Sei (X,d) eine metrischer Raum, $S \subseteq X$.

i) Wenn S kompakt, dann ist S abgeschlossen und beschränkt, wobei $S \subseteq X$ **beschränkt** heißt, wenn der **Diameter** oder **Durchmesser** von S
   $$\operatorname{diam}(S) = \sup\{d(x,y) \mid x, y \in S\} < \infty \quad (S \neq \emptyset)$$
   Für $S = \emptyset$ definiert man $\operatorname{diam}(\emptyset) = 0$.

ii) Abgeschlossene Teilmengen eines kompakten metrischen Raums sind kompakt.

Beweis.

i) Sei S kompakt.

   Zur Abgeschlossenheit: Zeige $X \setminus S$ ist offen. O.B.d.A. ist $X \setminus S \neq \emptyset$. Sei dann $x \in X \setminus S$. Bestimme nun $\epsilon > 0$ mit $B_\epsilon(x) \subseteq X \setminus S$.

   Bemerke dazu, dass zu jedem $k \in S$ ein $\epsilon_k > 0$ existiert, sodass $B_\epsilon(k) \cap B_{\epsilon_k}(x) = \emptyset$ (denn (X,d) ist hausdorffsch, d.h. verschiedene Punkte können durch disjunkte ϵ-Kugeln getrennt werden). Dann gilt
   $$S \subseteq \bigcup_{k\in S} B_{\epsilon_k}(k)$$

d.h. $\mathcal{U} = (B_{\epsilon_k}(k))_{k\in S}$ ist eine offene Überdeckung von S. Da S kompakt ist, existieren $k_1, \dots, k_n \in S$, sodass

$$S \subseteq \bigcup_{i=1}^{n} B_{\epsilon_{k_i}}(k_i)$$

Setze dann $\epsilon := \min_{i=1,\dots,n} \epsilon_{k_i}$. Klar ist dann:

$$B_{\epsilon_{k_i}}(k_i) \cap B_\epsilon(x) = \emptyset \quad \forall\, i = 1, \dots, n$$

also $S \cap B_\epsilon(x) = \emptyset$, d.h. $B_\epsilon(x) \subseteq X \setminus S$. Da $x \in X \setminus S$ beliebig war, ist also gezeigt, dass $X \setminus s$ offen ist.

Zur Beschränktheit: O.B.d.A. sei $S \neq \emptyset$. Sei $x_0 \in S$, dann gilt

$$S \subseteq \bigcup_{n\in\mathbb{N}} B_n(x_0)$$

d.h. $\mathcal{U} = (B_n(x_0))_{n\in\mathbb{N}}$ ist offene Überdeckung von S. Da S kompakt ist, besitzt $\mathcal{U}$ eine endliche Teilüberdeckung, d.h. aber gerade, dass ein $n_0 \in \mathbb{N}$ existiert mit $S \subseteq B_{\epsilon_0}(x)$. Somit folgt

$$\operatorname{diam}(S) = \sup\{d(x,y) \mid x, y \in S\} \leq 2n_0$$

d.h. aber gerade, S ist beschränkt.

ii) Sei (X, d) ein kompakter metrischer Raum. Sei S eine abgeschlossene Teilmenge von X. Zeige: S ist kompakt.

Sei dazu $\mathcal{U} = (U_i)_{i\in I}$ eine offene Überdeckung von S, d.h. $S \subseteq \bigcup_{i\in I} U_i$. Da S abgeschlossen ist, ist $X \setminus S$ offen, also ist

$$\tilde{\mathcal{U}} := (U_i)_{i\in I} \cup \{X \setminus S\}$$

eine offene Überdeckung von X. Da X kompakt ist, besitzt $\tilde{\mathcal{U}}$ eine endliche Teilüberdeckung

$$\{U_{i_1}, \dots, U_{i_n}, X \setminus S\} \;:\; X \subseteq \bigcup_{k=1}^{n} U_{i_k} \cup \{X \setminus S\}$$

Dann ist aber $\{U_{i_1}, \dots, U_{i_n}\}$ eine endliche Teilüberdeckung von S.

□

6.3.4 Folgen-Kompaktheit

Satz 6.3.3. Sei (X, d) eine metrischer Raum. Dann sind äquivalent:

i) S ist kompakt.

ii) Jede Folge in S hat einen Häufungspunkt in S, wobei $x \in X$ **Häufungspunkt** von $(x_n)_{n\in\mathbb{N}} \subseteq X$

$$:\Longleftrightarrow\; \forall\,\epsilon > 0 \quad \forall\, m \in \mathbb{N} \quad \exists\, n \geq m \;:\; d(x_n, x) < \epsilon$$

($\Longleftrightarrow$ in jeder ϵ-Kugel liegen unendlich viele Folgenglieder von $(x_n)_{n\in\mathbb{N}}$

$\Longleftrightarrow$ es existiert eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$, die gegen x konvergiert.)

iii) S ist folgen-kompakt, d.h. jede Folge in S besitzt eine in S konvergente Teilfolge.

Beweis. Die Äquivalenz zwischen ii) und iii) zeigt man wie in Analysis I.

i) $\Rightarrow$ ii) Angenommen, ii) gilt nicht, dann existiert eine Folge $(y_n)_{n\in\mathbb{N}}$ in S, die keinen Häufungspunkt in S besitzt, d.h.

$$\forall\, s \in S \quad \exists\, \epsilon_s > 0 \quad \exists\, m \in \mathbb{N} \quad \forall\, n \geq m \,:\, y_n \notin B_\epsilon(s)$$

Da $S \subseteq \bigcup_{s\in S} B_{\epsilon_s}(s)$ und S kompakt, existieren $s_1, \ldots, s_n \in s$ mit $S \subseteq \bigcup_{i=1}^n B_{\epsilon_i}(s_i)$. S enthält unendliche viele Folgenglieder von $(y_n)_{n\in\mathbb{N}}$, die Vereinigung auf der rechten Seite aber nur endliche viele, Widerspruch.

ii) $\Rightarrow$ i) Zu zeigen ist: Jede offene Überdeckung $\mathcal{U} = (U_i)_{i\in I}$ von S besitzt eine endliche Teilüberdeckung.

1. Schritt: Zeige zunächst, dass S durch endlich viele ϵ-Kugeln überdeckt werden kann

$$\forall\, \epsilon > 0 \quad \exists\, x_0, \ldots, x_m \in S \,:\, S \subseteq \bigcup_{i=0}^{m} B_\epsilon(x_i)$$

Teilmengen $S \subseteq (X, d)$ mit dieser Eigenschaft heißen **totalbeschränkt**. Klar ist:

$$S \text{ kompakt} \Rightarrow S \text{ totalbeschränkt} \Rightarrow S \text{ beschränkt.}$$

Die umgekehrten Implikationen gelten im Allgemeinen nicht.

Zeige also die Totalbeschränktheit von S. Angenommen, S sei nicht totalbeschränkt. Dann existiert ein $\epsilon > 0$, sodass S nicht durch endlich viele ϵ-Kugeln überdeckt werden kann. Sei $x_0 \in S$. Insbesondere existiert dann $x_1 \in S \setminus B_\epsilon(x_0)$. Wiederum überdeckt $B_\epsilon(x_0) \cup B_\epsilon(x_1)$ nicht S, d.h. es existiert $x_2 \in S \setminus (B_\epsilon(x_0) \cup B_\epsilon(x_1))$ usw.

So finden wir induktiv eine Folge $(x_n)_{n\in\mathbb{N}} \subseteq S$ mit der Eigenschaft $x_{n+1} \in S \setminus \bigcup_{i=0}^n B_\epsilon(x_i)$. Nach Voraussetzung besitzt $(x_n)_{n\in\mathbb{N}}$ einen Häufungspunkt $x \in S$. Insbesondere liegen in der Kugel $B_{\frac{\epsilon}{2}}(x)$ unendlich viele Folgenglieder, etwa $(x_{n_k})_k$. Folglich gilt

$$d(x_{n_k}, x_{n_l}) < \epsilon \quad \forall\, l, k \in \mathbb{N}$$

Andererseits gilt nach Konstruktion der Folge $(x_n)_{n\in\mathbb{N}}$

$$x_{n_l} \notin B_\epsilon(x_{n_k}) \quad \forall\, l > k$$

also $d(x_{n_k}, x_{n_l}) \geq \epsilon \quad \forall\, l > k$, Widerspruch. Die Menge ist also totalbeschränkt.

2. Schritt: Angenommen, S ist nicht kompakt, d.h. es existiert eine offene Überdeckung $\mathcal{U} = (U_i)_{i\in I}$ von S, die keine endliche Teilüberdeckung besitzt.

Da S totalbeschränkt ist, wissen wir, dass $\forall\, k \in \mathbb{N}$ S durch endliche viele offene Kugeln mit Radius $\frac{1}{k}$ überdeckt werden kann. Für jedes $k \in \mathbb{N}$ gibt es also eine Kugel $B_{\frac{1}{k}}(x_k)$, $x_k \in S$, die nicht durch endlich viele der Mengen U_i überdeckt werden kann. Nach Voraussetzung besitzt die so gefundene Folge $(x_k)_{k\in\mathbb{N}} \subseteq S$ einen Häufungspunkt $\hat{x} \in S$.

Da $\mathcal{U}$ Überdeckung von S, existiert ein $i_0 \in I$ mit $\hat{x} \in U_{i_0}$. U_{i_0} ist offen, also existiert ein $\epsilon > 0$ mit $B_\epsilon(\hat{x}) \subseteq U_{i_0}$. Da $\hat{x}$ Häufungspunkt von $(x_k)_{k\in\mathbb{N}}$ ist, gibt es ein $n \in \mathbb{N}$, $n \geq \frac{2}{\epsilon}$, sodass $x_n \in B_{\frac{\epsilon}{2}}(\hat{x})$. Für dieses n gilt

$$B_{\frac{1}{n}}(x_n) \subseteq B_\epsilon(\hat{x}) \subseteq U_{i_0}$$

weil für alle $x \in B_{\frac{1}{n}}(x_n)$ gilt

$$d(x,\hat{x}) \leq \underbrace{d(x,x_n)}_{<\frac{1}{n}\leq\frac{\epsilon}{2}} + \underbrace{d(x_n,\hat{x})}_{<\frac{\epsilon}{2}} < \epsilon$$

D.h. aber, dass $B_{\frac{1}{n}}(x_n)$ durch endlich viele der offenen Mengen U_i überdeckt werden kann, Widerspruch. □

6.3.5 Satz von Heine-Borel

Satz 6.3.4. Sei $X = \mathbb{R}^N$ für $N \geq 1$ mit d als euklidischen Metrik auf X (oder einer der äquivalenten Metriken d_p für $1 \leq p \leq \infty$) und $A \subseteq X$. Dann gilt

$$\boxed{A \text{ kompakt} \iff A \text{ beschränkt und abgeschlossen.}}$$

Beweis. Die Hinrichtung gilt nach Satz 6.3.2 in jedem metrischen Raum.

Sei also A beschränkt und abgeschlossen. Zeige: A ist folgen-kompakt. Sei dazu $(x_n)_{n\in\mathbb{N}} \subseteq A$; A ist beschränkt, also ist $(x_n)_{n\in\mathbb{N}}$ beschränkt, d.h. $\exists\, M > 0$:

$$|x_{ni} - x_{1i}| \overset{\forall\, i=1,\ldots,N}{\leq} \left(\sum_{i=0}^{N} |x_{ni} - x_{1i}|^2\right)^{\frac{1}{2}} = d_2(x_n, x_1) \leq M \quad \forall\, n \in \mathbb{N}$$

Also ist jede Komponentenfolge $(x_{ni})_{n\in\mathbb{N}}$ beschränkt in $\mathbb{R}$. Nach dem Satz von Bolzano-Weierstraß besitzt jede Komponentenfolge also eine in $\mathbb{R}$ gegen ein $\hat{x_i} \in \mathbb{R}$ konvergente Teilfolge. Nach Lemma 6.2.3 folgt also

$$x_{n_k i} \underset{d}{\rightarrow} \hat{x} = (\hat{x_1}, \ldots, \hat{x_N}) \in \mathbb{R}^N$$

Da A abgeschlossen ist und $(x_{n_k})_{k\in\mathbb{N}} \subseteq A$, folgt $\hat{x} \in A$. Somit ist gezeigt, dass $(x_n)_{n\in\mathbb{N}}$ eine in A konvergente Teilfolge besitzt. Also ist A folgen-kompakt, und damit kompakt. □

6.3.6 Beispiele

1. Abgeschlossene Quader der der Form

   $$[a_1, b_2] \times [a_2, b_2] \times \ldots \times [a_N, b_N] = \left\{x \in \mathbb{R}^N \mid a_1 \leq x_i \leq b_i \quad \forall\, i = 1, \ldots, N\right\}$$

   für $a_i, b_i \in \mathbb{R}\, \forall\, i = 1, \ldots, N$ sind kompakte Teilmengen in $\left(\mathbb{R}^N, d_2\right)$ mit d_2 euklidische Metrik.

2. Sei $X = \mathbb{R}$ mit d definiert durch $d(x,y) = \frac{|x-y|}{1+|x-y|} \quad \forall\, x, y \in X$ als Metrik ist beschränkt, da $d(x,y) \leq 1 \quad \forall\, x, y \in \mathbb{R}$. Aber X ist auch abgeschlossen, trotzdem ist x nicht kompakt.

6.3.7 Kompaktheit impliziert Vollständigkeit

Satz 6.3.5. Sei (X, d) ein metrischer Raum, $S \subseteq X$. Ist S kompakt, so ist S vollständig.

Beweis. Sei $(x_n)_{n\in\mathbb{N}} \subseteq S$ eine Cauchy-Folge. Da S kompakt ist, existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ die in S gegen ein $x \in S$ konvergiert. Wie in Analysis I folgt dann, dass die gesamte Folge gegen x konvergiert:

Sei $\epsilon > 0$. Da $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, existiert ein $n_0 \in \mathbb{N}$, sodass

$$d(x_n, x_m) < \frac{\epsilon}{2} \quad \forall\, n, m \geq n_0$$

Da $x_{n_k} \overset{k\to\infty}{\underset{d}{\longrightarrow}} x$, existiert ein $k_0 \in \mathbb{N}$, sodass

$$d(x_{n_k}, x) < \frac{\epsilon}{2} \quad \forall\, k \geq k_0$$

Für alle $n \geq \max\{n_0, n_{k_0}\}$ gilt also

$$d(x_n, x) \leq \underbrace{d(x_n, x_{n_k})}_{\substack{<\frac{\epsilon}{2} \\ \forall\, k\geq k_0, n\geq n_0}} + \underbrace{d(x_{n_k}, x)}_{<\frac{\epsilon}{2}} < \epsilon$$

□

6.3.8 Verallgemeinerung von Heine-Borel

Satz 6.3.6. Sei (X, d) ein metrischer Raum, $\emptyset \neq S \subseteq X$. Dann gilt

$$\boxed{S \text{ ist kompakt} \iff S \text{ ist vollständig und totalbeschränkt.}}$$

Beweis. Die Hinrichtung wurde im vorherigen Satz gezeigt.

Sei S also vollständig und totalbeschränkt. Zeige, dass S folgen-kompakt ist. Sei dazu $(y_n)_{n\in\mathbb{N}}$ eine beliebige Folge in S. Da S totalbeschränkt ist, kann S für jede $k \in \mathbb{N}$ durch endlich viele offene Kugeln mit dem Radius $\frac{1}{k}$ überdeckt werden.

Zu jedem $k \in \mathbb{N}$ existiert dann eine Kugel $B_{\frac{1}{k}}(x_k)$, $x_k \in S$, die unendlich viele Folgenglieder von $(y_n)_{n\in\mathbb{N}}$ besitzt. Genauer: Für $k = 1$ existiert eine Kugel $B_1(x_1)$, $x_1 \in S$, sodass unendlich viele Glieder, also etwa die Teilfolge

$$(y_{1_n})_{n\in\mathbb{N}} = (y_{1_1}, y_{1_2}, \ldots, y_{1_n}, \ldots)$$

in $B_1(x_1)$ liegt. Für $k = 2$ gibt es eine Kugel $B_{\frac{1}{2}}(x_2)$, sodass unendliche viele Glieder von $(y_{1_n})_{n\in\mathbb{N}}$, also etwa die Teilfolge

$$(y_{2_n})_{n\in\mathbb{N}} = (y_{2_1}, y_{2_2}, \ldots, y_{2_n}, \ldots)$$

in $B_{\frac{1}{2}}(x_2)$ liegt usw.

So erhalten wir induktiv Kugeln $B_{\frac{1}{k}}(x_k)$ und Teilfolgen $(y_{n_k})_{k\in\mathbb{N}}$ mit $(y_{n_k})_n \subseteq B_{\frac{1}{k}}(x_k)$. Betrachte nun die Diagonalfolge $(y_{n_n})_{n\in\mathbb{N}}$. Dann gilt

$$y_{n_n} \in B_{\frac{1}{k}}(x_k) \quad \forall\, n \geq k$$

für alle $k \in \mathbb{N}$ und $(y_{n_n})_{n\in\mathbb{N}}$ ist eine Cauchy-Folge in S, denn zu $\epsilon > 0$ wähle $k \in \mathbb{N}$ mit $\frac{1}{k} < \frac{\epsilon}{2}$, damit gilt

$$\forall\, n, m \geq k \;:\; d(y_{n_n}, y_{m_m}) \leq \underbrace{d(y_{n_n}, x_k)}_{<\frac{1}{k}} + \underbrace{d(x_k, y_{m_m})}_{<\frac{1}{k}} < \epsilon$$

Klar ist auch, dass nach Konstruktion $(y_{n_n})_{n\in\mathbb{N}}$ Teilfolge von $(y_n)_{n\in\mathbb{N}}$ ist. Da S nach Voraussetzung vollständig ist, ist $(y_{n_n})_{n\in\mathbb{N}}$ in S konvergent. Damit ist gezeigt, dass $(y_n)_{n\in\mathbb{N}}$ eine in S konvergente Teilfolge besitzt, also ist S folgenkompakt. $\square$

6.4 Zusammenhang

6.4.1 Erinnerung: Intervalle

$I \subseteq \mathbb{R}$ heißt Intervall, wenn gilt

$$\forall\, x, z \in I,\ y \in \mathbb{R} \,:\, x < y < z \Rightarrow y \in I$$

Beispiele von Intervallen in $\mathbb{R}$:

$$\begin{array}{ll} \emptyset,\ \mathbb{R},\ \{x\}\ \ (x \in \mathbb{R}), & \\ ]a,b[= \{x \in \mathbb{R} \mid a < x < b\}\,, & [a,b] = \{x \in \mathbb{R} \mid a \le x \le b\} \\ [a,b[= \{x \in \mathbb{R} \mid a \le x < b\}\,, &]a,b] = \{x \in \mathbb{R} \mid a < x \le b\} \end{array}$$

6.4.2 Topologische Charakterisierung von Intervallen

Für $I \subseteq \mathbb{R}$ ist äquivalent:

i) I ist ein Intervall.

ii) $\emptyset$ und I sind die einzigen Teilmengen von I, die in (I, d_I) sowohl offen als auch abgeschlossen sind.

Beweis.

$\Leftarrow$ Beweis durch Kontraposition: Sei I sei kein Intervall, d.h.

$$\exists\, x, z \in I,\ y \in \mathbb{R} \,:\, x < y < z \text{ und } y \notin I$$

Setze $B := \{a \in I \mid a < y\}$. Dann gilt $B \neq \emptyset$, weil $x \in B$, $B \subsetneq I$, $z \notin B$. Weiterhin gilt:

$$B = I \cap \,]-\infty, y[\, = I \cap \,]-\infty, y]$$

also ist B offen und abgeschlossen in I gleichzeitig, also gilt i) nicht.

$\Rightarrow$ Sei I ein Intervall, $\emptyset \neq B \subsetneq I$. Zeige: B ist nicht zugleich offen und abgeschlossen. Da $B \neq \emptyset$, existiert ein $x \in B$; da $B \neq I$, existiert ein $z \in I \setminus B$, o.B.d.A. $x < z$. Setze

$$y := \sup\{a \in B \mid a < z\}$$

Dann gilt: $x \le y \le z$. Da $x, z \in I$ und I ein Intervall ist, folgt $y \in I$.

Ist $y \in B$, so ist $y < z$ und $]y, z[\, \cap B = \emptyset$, also ist B nicht offen.

Ist $y \notin B$, so kann B nicht abgeschlossen sein.

$\square$

6.4.3 Definition

Definition 6.4.1. Sei (X, d) ein metrischer Raum.

i) X heißt **zusammenhängend**, wenn $\emptyset$ und X die einzigen Teilmengen von X, die zugleich offen und abgeschlossen sind.

ii) $Y \subseteq X$ heißt zusammenhängend, falls (Y, d_Y) zusammenhängend.

Bemerkung. Ist (X, d) ein metrischer Raum, so sind äquivalent:

i) X ist zusammenhängend.

ii) X kann nicht als Vereinigung disjunkter, nicht-trivialer offener Teilmengen von X geschrieben werden. Das heißt, wenn A, B offene Teilmengen von X sind mit $A \cap B = \emptyset$ und $A \cup B = X$, so ist $A = \emptyset$ oder $B = \emptyset$.

iii) Für alle Teilmengen $\emptyset \neq A \subsetneq X$ gilt $\partial A \neq \emptyset$.

Beweis.

i) $\Rightarrow$ ii) Angenommen X kann als Vereinigung disjunkter, nicht-trivialer offener Teilmengen A, B geschrieben werden, dann ist $A = X \setminus B$ zugleich offen - nach Voraussetzung - und abgeschlossen, weil B offen ist. Also ist X nicht zusammenhängend.

ii) $\Rightarrow$ i) Angenommen X ist nicht zusammenhängend, dann gibt es eine offene und abgeschlossene Teilmenge $\emptyset \neq A \subsetneq X$. Aber dann ist auch $\emptyset \neq (X \setminus A) \neq X$ offen und abgeschlossen, und natürlich $A \dot{\cup} (X \setminus A) = X$.

i) $\Leftrightarrow$ iii) X ist zusammenhängend genau dann, wenn $\emptyset$ und X die einzigen Teilmengen von X sind, die zugleich offen und abgschlossen sind. Die ist äquivalent zu: Für alle Teilmengen $\emptyset \neq A \subsetneq X$ ist $\mathring{A}$ verschieden von $\overline{A}$, also $\partial A = \overline{A} \setminus \mathring{A} \neq \emptyset$.

□

6.4.4 Vereinigung zusammenhängender Mengen

Satz 6.4.2. Sei (X, d) ein metrischer Raum.

i) Ist $(A_i)_{i \in I}$ eine Familie zusammenhängender Teilmengen von X und ist $\bigcap_{i \in I} A_i \neq \emptyset$, dann ist auch $\bigcup_{i \in I} A_i$ zusammenhängend.

ii) Wenn A zusammenhängende Teilmenge von X, dann ist auch $\overline{A}$ zusammenhängend.

Beweis.

i) Angenommen $\bigcup_{i \in I} A_i$ sei nicht zusammenhängend, d.h. es gibt eine offene und abgeschlossene Teilmenge $\emptyset \subsetneq A \subsetneq \bigcup_{i \in I} A_i$. Dann gibt es ein $i \in I$ mit $\emptyset \subsetneq (A \cap A_i) \subsetneq A_i$, womit eine nicht-triviale offene und abgeschlossene Teilmenge von A_i gefunden ist. Denn wäre $(A \cap A_i) = A_i$ oder $(A \cap A_i) = \emptyset$ für alle $i \in I$, dann wäre entweder $A = \bigcup_{i \in I} A_i$, was ein Widerspruch ist, oder es gälte $A \cap A_i = A_i$ und $A \cap A_j = \emptyset$ gleichzeitig für bestimmte $i, j \in I$, also $A_i \cap A_j \subseteq A \cap A_j = \emptyset$, womit dann folgt

$$\bigcap_{i \in I} A_i = \emptyset$$

ii) Sei $\emptyset \neq B \subseteq \overline{A}$ eine Teilmenge von $\overline{A}$, die zugleich offen und abgeschlossen ist. Wäre $B \cap A = \emptyset$, so wäre A echte Teilmenge von $\overline{A} \setminus B$. Da B offen ist, ist $\overline{A} \setminus B = \overline{A} \cap (X \setminus B)$ als Durchschnitt abgeschlossener Mengen abgeschlossen. $\overline{A}$ ist aber die kleinste abgeschlossene Obermenge von A, also müsste $\overline{A} \subseteq \overline{A} \setminus B$, d.h. $B = \emptyset$ gelten.

Somit ist dann $B \cap A$ eine nicht-leere offene und zugleich abgeschlossene Teilmenge von A. A ist zusammenhängend, also folgt

$$A \cap B = A \;\Rightarrow\; A \subseteq B \subseteq \overline{A} \overset{B \text{ abgeschlossen}}{\Rightarrow} \overline{A} \subseteq B \subseteq \overline{A} \;\Rightarrow\; \overline{A} = B$$

□

6.5 Separabilität

6.5.1 Definition

Definition 6.5.1. Sei (X, d) ein metrischer Raum.

i) Eine Teilmenge $M \subseteq X$ heißt **dicht** in X (oder dichte Teilmenge von X), wenn $\overline{M} = X$, d.h.

$$\forall\, x \in X \quad \exists\, (x_n)_{n \in \mathbb{N}} \subset M \,:\, x_n \underset{d}{\to} x$$

ii) (X, d) heißt **separabel**, wenn es eine abzählbare Teilmenge von X gibt, die dicht in X ist. X heißt dann **abzählbar dichte** Teilmenge von X

6.5.2 Beispiele

1. $(\mathbb{R}, d_{|\cdot|})$ ist separabel, denn $\mathbb{Q}$ ist abzählbar dicht.

2. $\mathbb{R}^N$ für $N \geq 1$ ausgestattet mit einer der äquivalenten Metriken d_p, $1 \leq p < \infty$, ist separabel, da $\mathbb{Q}^N$ abzählbar dichte Teilmenge ist.

3. $C([a, b])$ ausgestattet mit der Supremumsnorm ist separabel, denn die Menge aller Polynome mit rationalen Koeffizienten ist abzählbar dichte Teilmenge von $(C([a, b]), d_\infty)$ (Satz von Weierstraß).

4. Die beschränkten reellen Zahlenfolgen

$$\ell^\infty(\mathbb{N}) := \left\{ (x_n)_{n \in \mathbb{N}} \subset \mathbb{R} \mid \sup_{n \in \mathbb{N}} |x_n| < \infty \right\}$$

ist ein Vektorraum. $\left\| (x_n)_{n \in \mathbb{N}} \right\| := \sup_{n \in \mathbb{N}} |x_n|$ ist eine Norm auf $\ell^\infty(\mathbb{N})$ (Supremumsnorm). d_∞ bezeichne die dann induzierte Norm. Es gilt:

$$\boxed{(\ell^\infty(\mathbb{N}), d_\infty) \text{ ist nicht separabel.}}$$

Beweis. Angenommen es existierte ein abzählbares $A \subseteq \ell^\infty(\mathbb{N})$ mit $\overline{A} = \ell^\infty(\mathbb{N})$.

Sei $M \subseteq \mathbb{N}$, definiere $(\chi_{M\,n})_{n\in\mathbb{N}}$ durch

$$\chi_{M\,n} := \begin{cases} 1, & n \in M \\ 0, & n \notin M \end{cases}$$

Es gilt $\|\chi_{M\,n}\| = \sup_{n\in\mathbb{N}} |\chi_{M\,n}| = 1 < \infty$, also ist $(\chi_{M\,n})_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$. Für $\mathbb{N} \supseteq \hat{M} \neq M$ ist

$$d_\infty \left((\chi_{M\,n})_n , (\chi_{\hat{M}\,n})_n \right) = \sup_{n\in\mathbb{N}} |\chi_{M\,n} - \chi_{\hat{M}\,n}| = 1$$

Die Menge

$$\triangle = \left\{ (\chi_{M\,n})_{n\in\mathbb{N}} \mid M \subseteq \mathbb{N} \right\} \subseteq \ell^\infty$$

hat folglich die Kardinalität der Potenzmenge $\mathcal{P}(\mathbb{N})$, ist also überabzählbar. Wegen $\overline{A} = \ell^\infty(\mathbb{N})$ gilt

$$\forall\, (x_n)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N}) \quad \forall\, \epsilon > 0 \quad \exists\, (a_n)_{n\in\mathbb{N}} \in A \mid d_\infty \left((x_n)_n , (a_n)_n \right) < \epsilon$$

Insbesondere gibt es also für jedes $(\chi_{M\,n})_n \in \triangle$ ein $(a_n)_n \in A$ mit

$$d_\infty((\chi_{M\,n})_n , (a_n)_n) < \frac{1}{4}$$

Da verschiedene Elemente aus $\triangle$ gemessen mit der d_∞-Metrik genau 1 voneinander entfernt liegen, liegt für jedes $(a_n)_n \in A$ höchstens ein $(\chi_{M\,n})_n$ in $B_{\frac{1}{4}}((a_n)_n)$. Damit muss es aber mindestens so viele $(a_n)_n$ geben wie $(\chi_{M\,n})_n$, d.h. A kann nicht abzählbar sein. □

6.5.3 Kompakte Räume sind separabel

Satz 6.5.2.

i) Jeder kompakte metrischer Raum (X, d) ist separabel.

ii) Eine Teilmenge eines separablen metrischen Raums ist separabel.

Beweis.

i) Wenn (X, d) kompakt ist, dann ist (X, d) insbesondere total beschränkt und somit kann X für jedes $n \in \mathbb{N}$ durch endliche viele Kugeln mit Radius $\frac{1}{n}$ überdeckt werden:

$$X \subseteq B_{\frac{1}{n}}(x_1^{(n)}) \cup \ldots \cup B_{\frac{1}{n}}(x_{k_n}^{(n)})$$

mit gewissen Punkten $x_1^{(n)}, \ldots, x_{k_n}^{(n)}$, $k_n \in \mathbb{N}$.

Dann ist $M := \left\{ x_i^{(n)} \mid n \in \mathbb{N},\, i = 1, \ldots, k_n \right\}$ abzählbare Teilmenge von X. Sei nun $x \in X$, dann gibt es $b_n \in \mathbb{N}$ mit $1 \leq b_n \leq k_n$, sodass $x \in B_{\frac{1}{n}}(x_{b_n})$ für jedes $n \in \mathbb{N}$. Für die Folge $\left(x_{b_n}^{(n)} \right)_{n\in\mathbb{N}} \subset M$ gilt dann

$$\forall\, n \in \mathbb{N} \,:\, d(x, x_{b_n}^{(n)}) < \frac{1}{n}$$

also $x_{b_n}^{(n)} \xrightarrow[d]{} x$.

ii) Wenn (X,d) separabel ist, $M \subseteq X$ abzählbar dichte Teilmenge von X und $Y \subseteq X$, dann ist zu zeigen, dass Y separabel ist.

Achtung. $Y \cap M$ ist im Allgemeinen nicht dicht in Y, betrachte etwa $Y := X \setminus M$.

Für jedes $m \in M$ und jedes $n \in \mathbb{N}$, sodass $B_{\frac{1}{n}}(m) \cap Y \neq \emptyset$ wähle $y_{mn} \in B_{\frac{1}{n}}(m) \cap Y$. Dann folgt: $N := \{y_{mn} \text{ so gewählt}\}$ ist abzählbar dichte Teilmenge von Y. Denn sei $y \in Y$, so gibt es wegen $\overline{M} = X$ für jedes $n \in \mathbb{N}$ ein $m \in M$ mit $m \in B_{\frac{1}{2n}}(y)$. Dann ist $B_{\frac{1}{n}}(m) \cap Y \neq \emptyset$, also gibt es ein $z_n := y_{mn} \in B_{\frac{1}{n}}(m)$ mit

$$d(y, z_n) = d(y, y_{mn}) \leq d(y,m) + d(m, y_{mn}) \leq \frac{1}{2n} + \frac{1}{n} = \frac{3}{2n}$$

Für dies so definierte Folge $(z_n)_{n\in\mathbb{N}} \subset N$ gilt $z_n \underset{d}{\to} y$.

□

6.6 Stetige Abbildungen in metrischen Räumen

6.6.1 Definition

Definition 6.6.1. Seien $(X,d), (X',d')$ metrische Räume.

i) Eine Abbildung $f : X \to X'$ heißt **stetig** in $p \in X$, wenn für jede Umgebung U von $f(p)$ in X' gilt, dass $f^{-1}(U)$ Umgebung von p in X ist.

ii) $f : X \to X$ heißt stetig auf X, wenn f in jedem Punkt $p \in X$ stetig ist.

Bemerkung. Die obige Definition lässt sich verallgemeinern auf den Fall, wo $(X,\tau), (X',\tau')$ beliebige topologische Räume sind. In metrischen Räumen ist die obige Definition der Stetigkeit zu der aus Analysis I bekannten ϵ-δ-Stetigkeit und zur Folgenstetigkeit äquivalent.

6.6.2 Äquivalente Charakterisierungen

Satz 6.6.2. Sind $(X,d), (X',d')$ metrische Räume, $f : X \to X'$, $p \in X$, dann ist äquivalent:

i) f ist stetig in p.

ii)
$$\forall \epsilon > 0 \quad \exists \delta > 0 \,:\, d'(f(x), f(p)) < \epsilon \quad \forall x \in X,\, d(x,p) < \delta$$
oder
$$\forall \epsilon > 0 \quad \exists \delta > 0 \,:\, f(B_\delta(p)) \subseteq B_\epsilon(f(p))$$

iii) f ist **folgenstetig** in p, d.h.
$$\exists\, (x_n)_{n\in\mathbb{N}} \subset X \,:\, x_n \underset{d}{\to} p \;\Rightarrow\; f(x_n) \underset{d'}{\to} f(p)$$

Beweis.

ii) $\Leftrightarrow$ iii) zeigt man wie in Analysis I.

i) $\Rightarrow$ ii) : Sei $\epsilon > 0$. Dann ist $B_\epsilon(f(p))$ eine Umgebung von $f(p)$. Nach i) ist dann $f^{-1}(B_\epsilon(f(p)))$ eine Umgebung von p in X. D.h. aber gerade, dass ein $\delta > 0$ existiert mit $B_\delta(p) \subseteq f^{-1}(B_\epsilon(f(p)))$, also ist $f(B_\delta(p)) \subseteq B_\epsilon(f(p))$.

ii) $\Rightarrow$ i) Sei U eine Umgebung von $f(p)$ in X'. Dann existiert ein $\epsilon > 0$ mit $B_\epsilon(f(p)) \subseteq V$. Nach Voraussetzung ii) gibt es ein $\delta > 0$ mit $f(B_\delta(p)) \subseteq B_\epsilon(f(p))$. Es folgt $p \in B_\delta(p) \subseteq f^{-1}(V)$. Somit ist $f^{-1}(V)$ Umgebung von p in X.

□

6.6.3 Charakterisierung globaler Stetigkeit

Äquivalent charakterisiert man globale Stetigkeit.

Satz 6.6.3. Seien wieder $(X, d), (X', d')$ metrische Räume, $f : X \to X'$, dann ist äquivalent:

i) f ist stetig auf X.

ii) Für alle offenen Mengen $O \subseteq X'$ ist $f^{-1}(O)$ offen in X.

iii) Für alle abgeschlossenen Mengen $A \subseteq X'$ ist $f^{-1}(A)$ abgeschlossen in X.

Beweis.

i) $\Rightarrow$ ii) Sei O eine offene Teilmenge von X'. Sei weiterhin $p \in f^{-1}(O)$. Da O offen in X' und $f(p) \in O$, existiert ein $\epsilon > 0$ mit $B_\epsilon(f(p)) \subseteq O$. Da f stetig ist nach Voraussetzung, existiert nach Satz 6.6.2 ein $\delta > 0$ mit $f(B_\delta(p)) \subseteq B_\epsilon(f(p))$, d.h. aber dann

$$B_\delta(p) \subseteq f^{-1}(B_\epsilon(f(O))) \subseteq f^{-1}(O)$$

also ist $f^{-1}(O)$ offen.

ii) $\Rightarrow$ i) Sei $p \in X$, $\epsilon > 0$. Dann ist $B_\epsilon(f(p))$ offene Menge in X'. Nach Voraussetzung ist $f^{-1}(B_\epsilon(f(p)))$ offen in X. Es folgt

$$\exists \delta > 0 \; : \; B_\delta(p) \subseteq f^{-1}(B_\epsilon(f(p)))$$

also $f(B_\delta(p)) \subseteq B_\epsilon(f(p))$, womit gezeigt ist, dass f in p stetig ist. Da $p \in X$ beliebig war, ist f stetig auf X.

ii) $\Leftrightarrow$ ii) folgt mit $f^{-1}(A) = X \setminus f^{-1}(X' \setminus A)$.

□

6.6.4 Beispiele

1. Sei (X, d) ein metrischer Raum, $p \in X$.

$$\begin{aligned} d(\cdot, p) : X &\Rightarrow \mathbb{R}, \\ x &\mapsto d(x, p) \end{aligned}$$

ist stetig (der Beweis findet sich in dem des Banach'schen Fixpunktsatzes).

2. Wenn $(X, \|\cdot\|_X)$ ein normierter Raum ist, so ist die Norm $\|\cdot\|_X$ stetig auf X.

6.6.5 Verknüpfung stetiger Funktion ist stetig

Wie in Analysis I zeigt man

Satz 6.6.4. Sind $(X, d_X), (Y, d_Y), (Z, d_Z)$ metrische Räume und $f : X \to Y$ stetig in $p \in X$, $g : Y \to Z$ stetig in $f(p)$, so ist $g \circ f$ stetig in a.

Bemerkung. Sind $f, g : (X, d_X) \to (Y, d_Y)$ Abbildungen und Y ein normierter (reeller oder komplexer) Vektorraum, so sind Linearkombinationen von f und g stetig, wenn f und g selbst stetig sind. D.h.

$$\begin{aligned} C(X, Y) &:= \{f : X \to Y \mid f \text{ stetig}\} \\ C_b(X, Y) &:= \{f : X \to Y \mid f \text{ stetig und beschränkt}\} \end{aligned}$$

sind Vektorräume.

Sei $\|\cdot\|_Y$ eine Norm auf Y und d_Y die durch $\|\cdot\|_Y$ induzierte Metrik, dann gilt

$$C_b(X, Y) = \left\{ f : X \to Y \mid f \text{ stetig und } \sup_{x \in X} \|f(x)\|_Y < \infty \right\}$$

$\|f\|_\infty := \sup_{x \in X} \|f(x)\|_Y$ heißt Supremumsnorm auf $C_b(X, Y)$, womit $C_b(X, Y)$ ein normierter Vektorraum ist. Ist $(Y, \|\cdot\|_Y)$ vollständig, so ist auch $C_b(X, Y)$ vollständig, was man analog wie beim Cauchy-Kriterium für gleichmäßige Konvergenz aus Analysis I zeigt. Insbesondere ist $(C_b(X, (\mathbb{R}^N, \|\cdot\|_p)), \|\cdot\|_\infty)$ ein Banachraum, da $(\mathbb{R}^N, \|\cdot\|_p)$ vollständig ist.

6.6.6 Verallgemeinerung des Zwischenwertsatzes

Satz 6.6.5. Sei (X, d) ein zusammenhängender metrischer Raum, (X', d') ein metrischer Raum. Sei $f : X \to X'$ stetig, dann ist $f(X)$ ein zusammenhängender Teilraum von X'.

Beweis. Sei $\emptyset \neq B \subseteq f(X)$ offen und abgeschlossen. Dann gilt

$$\begin{aligned} \exists O \subseteq X' \text{ offen } &: B = f(X) \cap O \\ \exists A \subseteq X' \text{ abgeschlossen } &: B = f(X) \cap A \end{aligned}$$

womit folgt $f^{-1}(B) = f^{-1}(O)$ und $f^{-1}(B) = f^{-1}(A)$. Nun ist aber $f^{-1}(O)$ offen und $f^{-1}(A)$ abgeschlossen, weil f stetig ist, d.h. $f^{-1}(B)$ ist offen und abgeschlossene Teilmenge von X. Da X zusammenhängend ist, muss $f^{-1}(B) = X$, also $B = f(X)$ sein, was zu zeigen war. □

6.6.7 Wegzusammenhang

Mit Hilfe der Definition stetiger Abbildungen können wir nun einen stärkeren Zusammenhangsbegriff definieren.

Definition. Sei (X, d) ein metrischer Raum. Eine Menge $M \subseteq X$ heißt **wegzusammenhängend**, wenn es für je zwei Punkte $x_0, x_1 \in X$ eine stetige Abbildung $\varphi : [0, 1] \to X$ gibt mit $\varphi(0) = x_0$ und $\varphi(1) = x_1$.

Wegzusammenhang ist stärker als Zusammenhang, was die Aussage des folgenden Satzes ist.

Satz. Ist $M \subseteq X$ in (X, d) eine wegzusammenhängend, dann ist M insbesondere zusammenhängend.

Beweis. Seien $\emptyset \neq U, V \subsetneq M$ zwei disjunkte offene Mengen, $u \in U, v \in V$. Da M wegzusammenhängend ist, gibt es eine stetige Abbildung $\varphi : [0,1] \to M$ mit $\varphi(0) = u$ und $\varphi(1) = v$. Da φ stetig ist, sind $\varphi^{-1}(U), \varphi^{-1}(V)$ offen. Das Intervall $[0,1]$ ist zusammenhängend, also gilt $\varphi^{-1}(U) \cup \varphi^{-1}(V) \subsetneq [0,1]$ oder $\varphi^{-1}(U) \cap \varphi^{-1}(V) \neq \emptyset$, denn keine der beiden Mengen ist leer und keine der beiden ist M (da U, V disjunkt sind). Also muss $U \cup V \subsetneq M$ oder $U \cap V \neq \emptyset$ gelten. $\square$

6.6.8 Kompaktheitssatz

Satz 6.6.6. $(X, d), (X', d')$ seien metrische Räume, $f : X \to X'$ stetig. Sei $A \subseteq X$ kompakt, dann ist $f(A) \subseteq X'$ kompakt.

Beweis. Seien $(U_i)_{i \in I}$ offene Überdeckungen von $f(A)$, d.h.

$$f(A) \subseteq \bigcup_{i \in I} U_i \Rightarrow A \subseteq \bigcup_{i \in I} f^{-1}(U_i)$$

Für jedes $i \in I$ ist $f^{-1}(U_i)$ offen, da U_i offen und f stetig. Da A kompakt ist, gibt es eine endliche Teilüberdeckung, sodass

$$A \subseteq \bigcup_{k=0}^{n} f^{-1}(U_i) \Rightarrow f(A) \subseteq \bigcup_{k=0}^{n} U_i$$

womit eine endliche Teilüberdeckung von $f(A)$ gefunden ist. $f(A)$ ist also kompakt. $\square$

6.6.9 Satz von Maximum und Minimum für reellwertige Funktionen

Korollar 6.6.7. Sei (X, d) metrischer Raum, $f : X \to \mathbb{R}$ stetig. Ist $A \subseteq X$ kompakt, so nimmt f auf A Minimum und Maximum an.

Beweis. Nach Satz 6.6.6 ist $f(A)$ kompakt, also insbesondere beschränkt, d.h. $\sup f(A) < \infty$, und es existiert eine Folge $(x_n)_{n \in \mathbb{N}} \subset A$ mit

$$\lim_{n \to \infty} f(x_n) = \sup f(A)$$

Da A kompakt ist, besitzt $(x_n)_{n \in \mathbb{N}}$ eine gegen $x^* \in A$ konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$. Damit gilt wegen der Stetigkeit von f

$$f(x^*) = \lim_{k \to \infty} f(x_{n_k}) = \lim_{n \to \infty} f(x_n) = \sup f(A) = \max f(A)$$

$\square$

6.6.10 Gleichmäßige Stetigkeit

Definition 6.6.8. Seien $(X, d), (X', d')$ metrische Räume, $f : X \to x'$ eine Abbildung.

i) f ist **gleichmäßig stetig** auf $X :\Longleftrightarrow$

$$\forall \epsilon > 0 \quad \exists \delta > 0 \, : \, f(B_\delta(x)) \subseteq B_\epsilon(f(x)) \quad \forall x \in X, \text{ d.h.}$$
$$d(x, \tilde{x}) < \delta \Rightarrow d'(f(x), f(\tilde{x})) < \epsilon \quad \forall x \in X$$

ii) f ist gleichmäßig stetig auf $M \subseteq X :\Longleftrightarrow$ die Einschränkung von f auf (M, d_M) ist gleichmäßig stetig.

6.6.11 Gleichmäßige Stetigkeit auf kompakten Mengen

Satz 6.6.9. Seien $(X, d), (X', d')$ metrische Räume, $f : X \to X'$ stetig. Sei $A \subseteq X$ kompakt, dann ist f gleichmäßig stetig auf A.

Beweis. Sei o.B.d.A. $A = X$ (sonst betrachte $f|_A$ in (A, d_A)). Sei $\epsilon > 0$. Da f stetig ist, existiert zu jedem $x \in X$ ein $\delta_x > 0$ mit

$$f(B_{\delta_x}(x)) \subseteq B_{\frac{\epsilon}{2}}(f(x)) \quad (*)$$

außerdem gilt

$$X \subseteq \bigcup_{x \in X} B_{\frac{\delta_x}{2}}(x)$$

Da X kompakt ist, gibt es $x_1, \ldots, x_n \in X$ mit

$$X \subseteq \bigcup_{i=1}^{n} B_{\frac{\delta_x}{2}}(x_i)$$

Sei $\delta := \min \left\{ \frac{\delta_{x_i}}{2} \mid i = 1, \ldots, n \right\} > 0$. Seien $x, \tilde{x} \in X$ mit $d(x, \tilde{x}) < \delta$.

$$\forall x \in X \quad \exists i \in \{1, \ldots, n\} \, : \, x \in B_{\frac{\delta_{x_i}}{2}}(x_i)$$

also ist $d(x, x_i) < \frac{\delta_{x_i}}{2}$, mit $(*)$ also $d(f(x), f(x_i)) < \frac{\epsilon}{2}$. Weiterhin ist

$$d(\tilde{x}, x_i) \leq d(\tilde{x}, x) + d(x, x_i) < \delta + \frac{\delta_{x_i}}{2} < \delta_{x_i}$$

also wieder $d(f(\tilde{x}), f(x_i)) < \frac{\epsilon}{2}$. Schließlich ist

$$d(f(x), f(\tilde{x})) \leq d(f(x), f(x_i)) + d(f(x_i), f(x)) < \epsilon$$

□

6.6.12 Gleichmäßig konvergente Funktionen sind stetig

Definition und Satz 6.6.10. Seien $(X, d), (X', d')$ metrische Räume und $f_n : x \to x'$ für $n \in \mathbb{N}$ stetige Funktionen. Wenn $(f_n)_{n \in \mathbb{N}}$ **gleichmäßig** auf X gegen eine Funktion $f : X \to X'$ **konvergiert**, d.h.

$$\forall \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \, : \, d'(f_n(x), f(x)) < \epsilon \quad \forall n \geq n_0, x \in X$$

dann ist f stetig.

Beweis. Zeige die Stetigkeit in einem beliebigen Punkt $p \in X$. Sei dazu $(x_k)_k \subset X$ mit $x_k \xrightarrow[d]{} p$. Sei $\epsilon > 0$. Wähle $n_0 \in \mathbb{N}$ gemäß der gleichmäßigen Konvergenz. Da f_n stetig ist in p, existiert ein $\delta > 0$ mit

$$d'(f_{n_0}(x), f_{n_0}(p)) < \epsilon \quad \forall\, x \in B_\delta(p)$$

Da $x_k \xrightarrow[d]{} p$, existiert ein $k_0 \in \mathbb{N}$ mit $d(x_k, p) < \delta$ für alle $k \geq k_0$. Dann gilt ab diesem k_0

$$d'(f(x_k), f(p)) \leq d'(f(x_k), f_{n_0}(x_k)) + d'(f_{n_0}(x_k), f_{n_0}(p)) + d'(f_{n_0}(p), f(p)) < 3\epsilon$$

□

6.6.13 Lineare Abbildungen

Betrachte im folgenden Abbildungen $f : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ zwischen normierten Vektorräumen. Begriffe wie Stetigkeit und entsprechende Ergebnisse lassen sich auf diesen Fall übertragen, indem die metrischen Räume $(X, d_X), (Y, d_Y)$ betrachtet werden, wobei d_X, d_Y die von $\|\cdot\|_X, \|\cdot\|_Y$ induzierten Metriken sind.

Definition 6.6.11. Seien X, Y $\mathbb{K}$-Vektorräume ($\mathbb{K} = \mathbb{R}, \mathbb{C}$). Eine Abbildung $T : X \to Y$ heißt linear, wenn

$$T(\alpha \cdot x + \beta \cdot y) = \alpha \cdot Tx + \beta \cdot Ty \quad \forall\, x, y \in X,\ \alpha, \beta \in \mathbb{K}$$

Notation für lineare Abbildungen: Tx statt $T(x)$.

6.6.14 Stetigkeit linearer Abbildungen

Satz 6.6.12. Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Vektorräume, $T : X \to Y$ eine lineare Abbildung. Dann sind äquivalent:

i) T ist stetig auf X.

ii) T ist stetig in 0.

iii) $\exists\, M > 0 : \|Tx\|_Y \leq M \cdot \|x\|_X \quad \forall\, x \in X$.

iv) T ist gleichmäßig stetig auf X.

Beweis. i) ⇒ ii) und iv) ⇒ i) sind trivial.

iii) ⇒ iv) Sei $\epsilon > 0$, wähle $\delta := \frac{\epsilon}{M} > 0$, dann gilt für alle $x, y \in X$

$$\|x - y\|_X < \delta \Rightarrow \|Tx - Ty\|_Y = \|T(x - y)\|_Y \leq M \cdot \|x - y\|_X < \epsilon$$

ii) ⇒ iii) Da T in 0 stetig ist, gibt es zu $\epsilon = 1$ ein $\delta > 0$ sodass

$$\|x\|_X < \delta \Rightarrow \|Tx - T0\|_Y < 1$$

Sei $0 \neq x \in X$ beliebig, dann ist

$$\frac{\delta}{2} \cdot \frac{x}{\|x\|_X} \in B_\delta(0)$$

also gilt

$$\left\| T\left(\frac{\delta}{2} \cdot \frac{x}{\|x\|_X} \right) \right\|_Y < 1 \Rightarrow \|Tx\|_Y < \underbrace{\frac{2}{\delta}}_{=:M} \|x\|_X$$

□

6.6.15 Lineare, stetige Operatoren

Definition. Sind $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ nomierte $\mathbb{K}$-Vektorräume, dann setzt man

$$\mathcal{L}(X,Y) := \{T : X \to Y \mid T \text{ linear und stetig}\}$$

$\mathcal{L}(X,Y)$ ist ein $\mathbb{K}$-Vektorraum. Für $T \in \mathcal{L}(X,Y)$ sei

$$\|T\|_{\mathcal{L}(X,Y)} := \sup_{\substack{x\in X\\ \|x\|_X \leq 1}} \|Tx\|_Y (< \infty)$$

Man rechnet leicht nach, dass dadurch auf $\mathcal{L}(X,Y)$ eine Norm definiert wird. Falls $(Y, \|\cdot\|_Y)$ vollständig ist, d.h. ein Banachraum ist, so ist auch $(\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)})$ ein Banachraum.

Abbildungen $T \in \mathcal{L}(X,Y)$ bezeichnet man auch als **Operatoren** (genauer: **stetige lineare Operatoren** von X nach Y). $\|T\|_{\mathcal{L}(X,Y)}$ heißt auch **Operatornorm** von T.

Spezialfall: Ist $Y = \mathbb{K}$ mit der Norm $|\cdot|$, so ist $\mathcal{L}(X,\mathbb{K}), \|\cdot\|_{\mathcal{L}(X,\mathbb{K})}$ ein Banachraum. Die Elemente $T \in \mathcal{L}(X,\mathbb{K})$ bezeichnet man auch als **(stetige, lineare) Funktionale** auf X.

Notation: $X^* := \mathcal{L}(X,\mathbb{K})$, $\|T\|_{X^*} := \|T\|_{\mathcal{L}(X,\mathbb{K})}$.

6.6.16 Topologische Isomorphismen, Isometrien

Definition 6.6.13.

i) Normierte Vektorräume $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ heißen **topologisch isomorph**, wenn es einen stetigen (linearen) Isomorphismus $\iota : X \to Y$ gibt, sodass auch $\iota^{-1} : Y \to X$ stetig ist. Solch eine Abbildung heißt dann **topologischer Isomorphismus** zwischen X und Y.

   Notation: $X \cong Y$.

ii) Existiert ein **isometrischer topologischer Isomorphismus** $\iota : X \to Y$, d.h. gilt zusätzlich
$$\|\iota(x)\|_Y = \|x\|_X \quad \forall\, x \in X$$
   dann heißen X und Y **isometrisch isomorph**.

Bemerkung. Topologisch isomorphe Räume können miteinander identifiziert werden, insbesondere bildet ein topologischer Isomorphismus Cauchy-Folgen auf Cauchy-Folgen, konvergente Folgen auf konvergente Folgen ab. Folglich gilt: Ist $(X, \|\cdot\|_X) \cong (Y, \|\cdot\|_Y)$, so ist X vollständig genau dann, wenn Y vollständig ist.

Satz 6.6.14. Sei $(X, \|\cdot\|_X)$ ein endlich-dimensionaler normierter $\mathbb{R}$-Vektorraum, $\{v_1, \ldots, v_n\}$ eine Basis von X. Dann ist die Koordinatenabbildung

$$T : X \to \mathbb{R}^n$$
$$x = \sum_{j=1}^{n} x_j v_j \mapsto \sum_{j=1}^{n} x_j e_j$$

ein topologischer Isomorphismus zwischen $(X, \|\cdot\|_X)$ und $(\mathbb{R}^n, \|\cdot\|)$, d.h. $X \cong \mathbb{R}^n$.

Beweis. Klar: T ist wohldefiniert, linear und bijektiv. Weiterhin gilt für $x = \sum x_j e_j \in \mathbb{R}^n$.

$$\begin{aligned}
\left\|T^{-1}x\right\|_X &= \left\|\sum_{j=1}^n x_j v_j\right\|_X \leq \sum_{j=1}^n |x_j| \cdot \|v_j\|_X \\
&\overset{\text{Hölder}}{\leq} \left(\sum_{j=1}^n |x_j|^2\right)^{\frac{1}{2}} \cdot \left(\sum_{j=1}^n \|v_j\|_X^2\right)^{\frac{1}{2}} = M \cdot \|x\|_2
\end{aligned}$$

d.h. es gilt $\left\|T^{-1}x\right\|_X \leq M \cdot \|x\|_2 \quad \forall\, x \in \mathbb{R}^n$, wonach T^{-1} stetig ist. Um zu zeigen, dass auch T stetig ist, definiert man

$$|x| := \left\|T^{-1}x\right\|_X \quad \forall\, x \in \mathbb{R}^n$$

Man rechnet leicht nach, dass $|\cdot|$ eine Norm auf $\mathbb{R}^n$ definiert. Setze

$$c := \inf_{x \in \mathbb{S}^{n-1}} |x|$$

wobei $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n,\ \|x\|_2 = 1\}$ die Einheitssphäre ist.

Angenommen $c = 0$. Dann existiert eine Folge $(x_k)_k \subset \mathbb{S}^{n-1}$ mit $|x_k| \to 0$. Nach dem Satz von Bolzano-Weierstraß im $\mathbb{R}^n$ besitzt die beschränkte Folge $(x_k)_k$ eine konvergente Teilfolge $(x_{k_l})_l$ mit $x_{k_l} \xrightarrow[\|\cdot\|_2]{l\to\infty} x$. Da $\|x_{k_l}\|_2 = 1$ für alle $l \in \mathbb{N}$, folgt auch $\|x\|_2 = 1$. Somit gilt

$$|x| \leq |x - x_{k_l}| + |x_{k_l}| \leq M \cdot \|x - x_{k_l}\|_2 + |x_{k_l}| \xrightarrow{l\to\infty} 0$$

also $|x| = 0$, d.h. $x = 0$, was $\|x\|_2 = 1$ widerspricht.

Also muss $c > 0$ gelten. Für dieses $c > 0$ gilt somit $c \leq |x| \quad \forall\, x \in \mathbb{S}^{n-1}$. Für $x \in \mathbb{R}^n \setminus \{0\}$ ist $y := \frac{x}{\|x\|_2} \in \mathbb{S}^{n-1}$, und somit

$$c \leq |y| = \left|\frac{x}{\|x\|_2}\right| = \left\|T^{-1}\frac{x}{\|x\|_2}\right\|_X = \frac{1}{\|x\|_2} \cdot \left\|T^{-1}x\right\|_X$$

d.h. $c \cdot \|x\|_2 \leq \left\|T^{-1}x\right\|_X$, bzw. $\|Ty\|_2 \leq \frac{1}{c} \cdot \|y\|_X \quad \forall\, y \in X$, woraus die Stetigkeit von T folgt. □

Korollar 6.6.15. Auf $\mathbb{R}^n$ sind alle Normen äquivalent.

6.6.17 Im Endlichdimensionalen sind lineare Operatoren stetig

Satz 6.6.16. Sei $(X, \|\cdot\|_X)$ ein endlich-dimensionaler normierter $\mathbb{R}$-Vektorraum, $(Y, \|\cdot\|_Y)$ ein beliebiger normierter $\mathbb{R}$-Vektorraum, $T:\ X \to Y$ linear. Dann ist T stetig.

Beweis. Wegen Satz 6.6.14 kann man o.B.d.A. annehmen, dass $(X, \|\cdot\|_X) = (\mathbb{R}^n, \|\cdot\|_2)$ für ein $n \in \mathbb{N}$.

Sei also $T : (\mathbb{R}^n, \|\cdot\|_2) \to (Y, \|\cdot\|_Y)$ linear, dann gilt für alle $x = \sum x_j e_j \in \mathbb{R}^n$

$$\begin{aligned} \|Tx\|_Y = \left\| T \sum_{j=1}^n x_j e_j \right\|_Y = \left\| \sum_{j=1}^n x_j \cdot Te_j \right\|_Y \leq \sum_{j=1}^n |x_j| \, \|Te_j\|_Y \\ \overset{\text{Hölder}}{\leq} \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n \|Te_j\|_Y^2 \right)^{\frac{1}{2}} = M \, \|x\|_2 \end{aligned}$$

für konstantes $M \in \mathbb{R}$, womit die Stetigkeit von T gezeigt ist. $\square$

7 Differentialrechnung in normierten Räumen

7.1 Definition der Differenzierbarkeit und elementare Eigenschaften

$(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ seien im gesamten Kapitel Banachräume. Weiterhin sei - sofern nicht anders gesagt - U eine offene, nicht-leere Teilmenge des jeweils angegebenen Banachraums.

7.1.1 Erinnerung: Reeller Fall

Ist $U \subseteq X$ eine offene, nicht-leere Teilmenge von X, $f : U \to Y$. Wann heißt f differenzierbar in einem Punkt $x_0 \in U$?

Im Fall $X = Y = \mathbb{R}$ war f differenzierbar in x_0 genau dann, wenn

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

in $\mathbb{R}$ existiert. In dem Fall heißt dieser Grenzwert Ableitung von f an der Stelle x_0.

Das Konzept ist in dieser Form nicht übertragbar, da durch Vektoren $x - x_0 \in X$ „nicht dividiert werden kann". Aber äquivalent gilt in $X = Y = \mathbb{R}$: f ist differenzierbar in x_0 genau dann, wenn ein $k \in \mathbb{R}$ und eine Abbildung $r : U \to \mathbb{R}$ existieren, sodass

$$f(x) = f(x_0) + k \cdot (x - x_0) + r(x)\,(x - x_0)\,, \quad \forall\, x \in U$$

und sodass $\lim_{x \to x_0} r(x) = 0$. D.h. f kann lokal in x_0 durch die affin lineare Funktion

$$x \mapsto f(x_0) + k \cdot (x - x_0)$$

approximiert werden, mit einem Fehler, der schneller als der Abstand von x zu x_0 gegen 0 strebt. In diesem Fall ist

$$k =: f'(x_0) \in \mathbb{R} \cong L(\mathbb{R}, \mathbb{R})$$

wobei $\iota :\ L(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$, $T \mapsto T(1)$ ein isometrischer Isomorphismus ist.

7.1.2 Definition

Dieses Konzept lässt sich auf allgemeine normierte Räume übertragen.

Definition 7.1.1. Sei $U \subseteq X$ offen. Eine Abbildung $f :\ U \subseteq X \to Y$ heißt differenzierbar in $x_0 \in U$, wenn es ein $A \in L(X, Y)$ gibt, sodass

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - A(x - x_0)}{\|x - x_0\|} = 0$$

d.h.

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - A(x - x_0)\|_Y}{\|x - x_0\|_X} = 0$$

7.1.3 Quotientenfreie Form

Lemma 7.1.2. Seien die Voraussetzung wie in der obigen Definition gegeben. f ist differenzierbar in x_0 genau dann, wenn $A \in L(X,Y)$ und $r : U \to Y$ existieren, sodass

$$f(x) = f(x_0) + A(x - x_0) + r(x) \cdot \|x - x_0\|_X , \quad \forall\, x \in U$$

und sodass

$$\lim_{x \to x_0} \|r(x)\|_Y = 0$$

Beweis. Aus der Bedingung im Lemma folgt trivialerweise die Definition der Differenzierbarkeit. Sei nun andersherum f in x_0 differenzierbar, so definiere

$$r(x) := \begin{cases} \frac{f(x)-f(x_0)-A(x-x_0)}{\|x-x_0\|_X}, & x \neq x_0 \\ 0, & x = x_0 \end{cases}$$

und es folgt sofort das obige Kriterium. □

7.1.4 Elementare Eigenschaften

Satz 7.1.3. Seien die Voraussetzungen wieder wie oben.

i) Ist f differenzierbar in $x_0 \in U$, dann ist f insbesondere stetig in x_0.

ii) Ist f differenzierbar in x_0, dann ist die Abbildung $A \in L(X,Y)$ aus der Definition eindeutig.

Beweis.

i) folgt unmittelbar aus dem Lemma:

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} \left(f(x_0) + A(x - x_0) + r(x) \cdot \|x - x_0\| \right) = f(x_0)$$

ii) Angenommen es existieren $A, B \in L(X,Y)$ mit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - A(x - x_0)}{\|x - x_0\|_X} = 0$$
$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - B(x - x_0)}{\|x - x_0\|_X} = 0$$

dann folgt

$$0 = \lim_{x \to x_0} \frac{A(x - x_0) - B(x - x_0)}{\|x - x_0\|_X} = \lim_{x \to x_0} (A - B) \frac{(x - x_0)}{\|x - x_0\|_X}$$

Wähle für jedes $n \in \mathbb{N}$: $x_n = x_0 + \frac{y}{n}$ für $y \in X \setminus \{0\}$. Dann folgt

$$\lim_{n \to \infty} (A - B) \frac{\frac{y}{n}}{\left\| \frac{y}{n} \right\|} = 0$$

d.h. aber gerade $(A-B)\left(\frac{y}{\|y\|}\right) = 0$ für alle $y \in X \setminus \{0\}$, also $(A - B)\, y = 0$ für alle $y \in X$, konkreter $Ay = By$, also insgesamt $A = B$.

□

7.1.5 Ableitung

Definition 7.1.4.

i) Sei $U \subseteq X$ offen, $f : U \to Y$. Ist f differenzierbar in $x_0 \in U$, dann heißt der eindeutig bestimmte Operator A aus der Definition 7.1.1 **Ableitung** (oder auch **Fréchet-Ableitung**, oder **(totales) Differential von f in x_0**).

   Notation: $A = f'(x_0)$, $Df(x_0)$ oder $\partial f(x_0)$.

ii) Ist f in jedem Punkt $x \in U$ differenzierbar, so heißt f differenzierbar auf U und die Abbildung

$$\begin{aligned} f' : U &\to \mathcal{L}(X, Y) \\ x &\mapsto f'(x) \end{aligned}$$

   heißt **Ableitung** von f auf U. Ist f' stetig in x_0 (bzw. auf U), dann heißt f **stetig differenzierbar** in x_0 (bzw. auf U) und man setzt

$$C^1(U, Y) := \{f : U \to Y \mid f \text{ stetig differenzierbar auf } U\}$$

Bemerkung.

1. Der besseren Unterscheidbarkeit von anderen Differenzierbarkeitsbegriffen wegen bezeichnet man eine im Sinn der Definition 7.1.4 differenzierbare Abbildung auch als **total differenzierbar** oder **Fréchet-differenzierbar**.

2. Die Begriffe „Differenzierbarkeit“ und „Ableitung“ sind unabhängig von der Wahl äquivalenter Normen in X und Y.

3. Da $\mathcal{L}(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}$ kann man also $\mathcal{L}(\mathbb{R}, \mathbb{R})$ und $\mathbb{R}$ identifizieren. Diese Indetifizierung zeigt, dass die obige Definition 7.1.4 der Differenzierbarkeit im Falle $X = Y = \mathbb{R}$ mit der aus Analysis I bekannten übereinstimmt.

7.1.6 Beispiele

1. Für $B \in \mathcal{L}(X, Y)$ betrachte

$$\begin{aligned} B : X &\to Y \\ X &\mapsto Bx \end{aligned}$$

   Dann ist für $x_0, x \in X$ also

$$Bx = Bx_0 + B(x - x_0)$$

   womit A das Lemma mit $A := B$ und $r(x) = 0 \quad \forall x \in X$ erfüllt. D.h. B ist differenzierbar mit $B' = B$ auf ganz X.

2. Betrachte für $y_0 \in Y$ die konstante Funktion

$$\begin{aligned} k_{y_0} : X &\to Y \\ x &\mapsto y_0 \end{aligned}$$

   dann ist für alle $x, x_0 \in X$

$$k_{y_0}(x) = k_{y_0}(x_0) + A(x - x_0) + r(x) \|x - x_0\|_X$$

   erfüllt mit $A = 0$, $r \equiv 0$. Also ist k_{y_0} differenzierbar auf X mit $k'_{y_0} \equiv 0$.

7.1.7 Spezialisierung auf endlich-dimensionale Banachräume

Schon nach der Betrachtung dieser beiden relativ einfachen Funktionenklassen erkennt man das Problem: Wie kann man im allgemeinen Fall Differenzierbarkeit bequem überprüfen und die Ableitung berechnen?

Im Falle $X = \mathbb{R}^n, Y = \mathbb{R}^m$ für $n, m \in \mathbb{N}$ ist für eine im Punkt $x_0 \in U$ differenzierbare Funktion $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ die Ableitung $f'(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Also kann bei Wahl von kanonischen Basen in $\mathbb{R}^n$ und $\mathbb{R}^m$ $f'(x_0)$ mit der diese linearen Abbildung darstellenden $m \times n$-Matrix identifiziert werden. Die Frage ist, wie die Einträge dieser Matrix berechnet werden können.

7.1.8 Richtungsableitung

Zur Vorbereitung folgende

Definition 7.1.5. Sei $f : U \subseteq X \to Y$, $x_0 \in U$, $v \in X \setminus \{0\}$. Existiert

$$D_v f(x_0) := \lim_{t \to 0} \frac{f(x_0 + t \cdot v) - f(x_0)}{t}$$

so heißt $D_v f(x_0)$ die **Richtungsableitung** von f in x_0 in Richtung von v.

Grafische Illustration im Falle $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$.

Satz 7.1.6. Sei $f : U \subseteq X \to Y$ differenzierbar in $x_0 \in U$. Dann existiert die Richtungsableitung von f in x_0 in jede Richtung $v \in X \setminus \{0\}$ und es gilt $D_v f(x_0) = f'(x_0)v$.

Bemerkung. Aus diesem Satz folgt im Falle $X = \mathbb{R}^n, Y = \mathbb{R}^m$, dass es genügt, die Richtungsableitungen von einer in einem Punkt x_0 differenzierbare Funktion in Richtung der kanonischen Einheitsvektoren $e_i = (0, \dots, 1, \dots, 0)$, $i = 1, \dots, n$ zu kennen, um die totale Ableitung zu berechnen.

Beweis. f ist differenzierbar in x_0; sei $v \in X \setminus \{0\}$. Dann existiert nach Lemma über die alternative Definition der Differenzierbarkeit eine Funktion $r : U \to Y$ mit

$$f(x_0 + t \cdot v) = f(x_0) + f'(x_0)(t \cdot v) + r(x_0 + t \cdot v) \|t \cdot v\|_X$$

für alle $t \in \mathbb{R}$ genügend klein, soll heißen: sodass $x_0 + t \cdot v \in U$. Daraus folgt

$$\begin{aligned} \frac{f(x_0 + t \cdot v) - f(x_0)}{t} &= f'(x_0)v + r(x_0 + t \cdot v) \cdot \frac{|t|}{t} \|v\|_X \qquad \Rightarrow \\ \lim_{t \to 0} \frac{f(x_0 + t \cdot v) - f(x_0)}{t} &= f'(x_0)v \end{aligned}$$

d.h. $D_v f(x_0)$ existiert und ist gleich $f'(x_0)v$. □

Bemerkung. Die Umkehrung des Satzes 7.1.6 gilt im Allgemeinen nicht, d.h. wenn $f : U \subseteq X \to Y$ in einem Punkt x_0 in jeder Richtung $v \in X \setminus \{0\}$ eine Richtungsableitung besitzt, dann ist f nicht notwendig auch (total) differenzierbar in x_0. Betrachte zum Beispiel

$$\begin{aligned} f &: \mathbb{R}^2 \to \mathbb{R} \\ f(x, y) &:= \begin{cases} \frac{x^2 \cdot y}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases} \end{aligned}$$

Dann ist für $v = (\xi, \nu) \in \mathbb{R}^2 \setminus \{(0,0)\}$

$$f(0 + t \cdot v) - f(0) = f(tv) = \frac{t^2 \cdot \xi^2 \cdot t \cdot \nu}{t^2 \cdot \xi^2 + t^2 \cdot \nu^2} = t \cdot \frac{\xi^2 \cdot \nu}{\xi^2 + \nu^2} = t \cdot f(v)$$

Folglich gilt

$$\frac{f(0 + t \cdot v) - f(0)}{t} = \frac{t \cdot f(v)}{t} = \underbrace{f(v)}_{=: D_v f(0)} \qquad \forall\, t \in \mathbb{R} \setminus \{0\}$$

D.h. f besitzt in $0 = (0,0)$ Richtungsableitungen in jede Richtung $v \neq 0$ mit $D_v f(0) = f(v)$. Angenommen f wäre sogar differenzierbar in 0, dann folgt aus dem obigen Satz 7.1.6

$$\underbrace{f'(0)}_{\text{linear}} v = D_v f(0) = \underbrace{f}_{\text{nicht linear}} (v) \quad \forall\, v \in \mathbb{R}^2 \setminus \{(0,0)\}$$

was offensichtlich ein Widerspruch ist.

7.1.9 Partielle Ableitung

Zurück zur Bemerkung: Wir stellen die besondere Bedeutung der Richtungsableitungen im Fall $X = \mathbb{R}^n, Y = \mathbb{R}^m$, in Richtung der kanonischen Basisvektoren fest und geben

Definition 7.1.7. Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $x_0 \in U$.

i) Dann heißt für $k = 1, \dots, n$

$$\frac{\partial f}{\partial x_n}(x_0) := D_{e_k} f(x_0) = \lim_{t \to 0} \frac{f(x_0 + t \cdot e_k) - f(x_0)}{t}$$

(sofern existent) die **k-te partielle Ableitung** von f in x_0. Existiert in allen Punkten $x \in U$ die k-te partielle Ableitung von f, so heißt

$$\begin{aligned} \frac{\partial f}{\partial x_k} : U &\to \mathbb{R} \\ x &\mapsto \frac{\partial f}{\partial x_k}(x) \end{aligned}$$

die k-te partielle Ableitung von f auf U.

ii) f heißt in x_0 **partiell differenzierbar**, wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0)$ existieren. Analog heißt f auf U partiell differenzierbar, wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ auf U existieren.

iii) Sind diese stetig in x_0 (auf U), so heißt f in x_0 (auf U) **stetig partiell differenzierbar**.

Bemerkung.

1. f ist genau dann nach der k-ten Koordinate partiell differenzierbar, wenn die Funktion

$$t \in \mathbb{R} \mapsto f(x_0^1, \dots, x_0^{k-1}, x_0^k + t, x_0^{k+1}, \dots, x_n)$$

als Funktion der reellen Variable t differenzierbar ist.

2. Ist f in x_0 total differenzierbar, dann ist f in x_0 partiell differenzierbar (klar, da f nach Satz 7.1.6 Richtungsableitungen in jede Richtung besitzt, falls f (total) differenzierbar ist).

3. Wenn f in x_0 partiell differenzierbar ist, dann folgt im Allgemeinen nicht, dass f in x_0 total differenzierbar ist.

4. Wenn f in x_0 partiell differenzierbar ist, dann folgt im Allgemeinen nicht einmal, dass f in x_0 stetig ist. Betrachte zum Beispiel

$$\begin{aligned} f:\ \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases} \end{aligned}$$

Für $t \neq 0$ gilt

$$\frac{f((0,0)+t\cdot(1,0)) - f(0,0)}{t} = \frac{f(t,0)-0}{t} = 0 \Rightarrow \frac{\partial f}{\partial x_1}(0,0) = 0$$

Völlig analog sieht man $\frac{\partial f}{\partial x_2}(0,0) = 0$ ein. Aber f ist nicht stetig in $(0,0)$, denn es gilt

$$f\left(\frac{1}{n},\frac{1}{n}\right) = \frac{\frac{1}{n}\cdot\frac{1}{n}}{\left(\frac{1}{n}\right)^2 + \left(\frac{1}{n}\right)^2} = \frac{1}{2} \not\to 0 = f(0,0)$$

7.1.10 Ableitung als Jacobi-Matrix

Wählt man in $\mathbb{R}^n$ und $\mathbb{R}^m$ die kanonische Basis und identifiziert man $A \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ mit ihrer darstellenden Matrix, so erhält man die Darstellung der Ableitung einer Funktion $f:\ U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ in einem Punkt x_0:

Satz 7.1.8. Wenn $f:\ U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ in $x_0 \in U$ differenzierbar ist, dann gilt

$$\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m) \ni f'(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \in \mathbb{R}^{m\times n}$$

Beweis. Wir bemerken zunächst, dass $f = (f_1,\dots,f_m)$ in x_0 differenzierbar ist genau dann, wenn alle Komponentenfunktionen $f_1,\dots,f_m$ in x_0 differenzierbar sind; in dem Fall gilt $f'(x_0) = (f_1'(x_0),\dots,f_m'(x_0))^T$ (da Konvergenz im $\mathbb{R}^m$ komponentenweiser Konvergenz entspricht). Folglich reicht es aus, den Fall $m = 1$ zu betrachten und zu zeigen:

$$f'(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \end{pmatrix}$$

Dies folgt, da für $h \in \mathbb{R}^n$, $h = \sum\limits_{j=1}^{n} h_j \cdot e_j$ mit $h_j \in \mathbb{R}$, e_j j-ter Einheitsvektor für jedes $j = 1, \dots, n$, gilt:

$$\begin{aligned}
f'(x_0)h &= f'(x_0)\left(\sum_{j=1}^{n} h_j \cdot e_j\right) \\
&= \sum_{j=1}^{n} h_j \cdot f'(x_0)e_j \\
&= \sum_{j=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j} \\
&= \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) & \dots & \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) & \dots & \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix} h
\end{aligned}$$

□

Definition 7.1.9.

i) Die im Satz 7.1.8 definiert Matrix heißt **Jacobimatrix** oder **Funktionalmatrix** von f in x_0 und wird mit $J_f(x_0)$ bezeichnet. Mit anderen Worten:

   Die Jacobimatrix ist die Darstellungsmatrix der Ableitung von f in x_0 bezüglich der kanonischen Basen.

ii) Ist $m = 1$, d.h. $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, dann heißt

$$\nabla f(x_0) := \operatorname{grad} f(x_0) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \end{pmatrix}$$

   der **Gradient** von f in x_0.

Bemerkung. Für $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ differenzierbar ist damit für alle $v \in \mathbb{R}^n$, $x_0 \in U$:

$$D_v f(x_0) = \nabla f(x_0) \cdot v$$

Wenn $x_0 \in U$ mit $\nabla f(x_0) \neq 0$, dann setze

$$h_0 := \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|_2} \in \mathbb{R}^n$$

Es gilt:

$$D_{h_0} f(x_0) = \nabla f(x_0) \cdot \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|_2} = \|\nabla f(x_0)\|_2$$

Für jeden beliebigen Vektor $v \in \mathbb{R}^n$ mit $\|v\|_2 = 1$ gilt nun:

$$D_v f(x_0) = \nabla f(x_0) \cdot v \overset{\text{C-S-Ungleichung}}{\leq} \|\nabla f(x_0)\|_2$$

D.h. die Richtung des steilsten Anstiegs von f in x_0 ist in Richtung des Gradienten $\nabla f(x_0)$. Analog kann man zeigen, dass die Richtung des steilsten Abstiegs gegeben ist durch $-\nabla f(x_0)$.

7.1.11 Aus stetiger partieller Differenzierbarkeit folgt totale Differenzierbarkeit

Erinnerung: Die Existenz aller partiellen Ableitungen in einem Punkt $x_0 \in U$ garantiert noch nicht die (totale) Differenzierbarkeit von f in x_0. Aber es gilt

Satz 7.1.10. Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $x_0 \in U$. Dann sind gilt:

i) Ist f in x_0 stetig differenzierbar, dann ist f in x_0 stetig partiell differenzierbar.

ii) Ist f in x_0 stetig partiell differenzierbar, dann ist f in x_0 total differenzierbar.

Beweis.

i) folgt leicht mit Satz 7.1.8: Existiert

$$f' = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

in einer Umgebung von x_0, dann existieren die Komponentenfunktionen in einer Umgebung von x_0. Ist die Ableitung dann in x_0 stetig, dann sind auch die Komponentenfunktionen in x_0 stetig.

ii) Bemerke zunächst, dass $f = (f_1, \ldots, f_m)$ stetig differenzierbar in x_0 genau dann, wenn f_i stetig differenzierbar in x_0 für jedes $i = 1, \ldots, m$. O.B.d.A. sei $m = 1$, also $f : \mathbb{R}^n \to \mathbb{R}$ stetig partiell differenzierbar in x_0.
Für $h = \sum h_j e_j$ mit $h_j \in \mathbb{R}$ für $j = 1, \ldots, n$ genügend klein ist

$$\begin{aligned} f(x_0 + h) - f(x_0) &= f\left(x_0 + \sum_{j=1}^{n} h_j e_j\right) - f\left(x_0 + \sum_{j=1}^{n-1} h_j e_j\right) \\ &\quad + f\left(x_0 + \sum_{j=1}^{n-1} h_j e_j\right) - f\left(x_0 + \sum_{j=1}^{n-2} h_j e_j\right) \\ &\quad + \ldots \\ &\quad + f(x_0 + h_1 e_1) - f(x_0) \\ &\overset{\text{MWS}}{=} \frac{\partial f}{\partial x_n}(\xi_n) h_n + \frac{\partial f}{\partial x_{n-1}}(\xi_{n-1}) h_{n-1} + \ldots + \frac{\partial f}{\partial x_1}(\xi_1) h_1 \end{aligned}$$

für Stellen $\xi_1, \ldots, x_n$ mit ξ_j zwischen x_0^j und $x_0^j + h_j$. Dabei wurde benutzt, dass die partiellen Ableitungen in einer Umgebung von x_0 existieren. Also gilt

$$f(x_0 + h) - f(x_0) = \nabla f(\xi) \cdot h$$

wobei $\xi := (\xi_1, \ldots, \xi_n)$. Damit folgt

$$\begin{aligned} |f(x_0 + h) - f(x_0) - \nabla f(x_0) \cdot h| &= |\nabla f(\xi) \cdot h - \nabla f(x_0) h| \\ &= |(\nabla f(\xi) - \nabla f(x_0)) \cdot h| \\ &\overset{\text{C-S}}{\leq} \|\nabla f(\xi) - \nabla f(x_0)\|_2 \cdot \|h\|_2 \end{aligned}$$

d.h.

$$\frac{|f(x_0+h)-f(x_0)-\nabla f(x_0)\cdot h|}{\|h\|_2} \leq \|\nabla f(\xi)-\nabla f(x_0)\|_2 \overset{h\to 0}{\longrightarrow} 0$$

da $\frac{\partial f}{\partial x_j}$ stetig in x_0 für jedes $k=1,\dots,n$ und $\xi=(\xi_1,\dots,\xi_n)\to x_0$, da ξ_k zwischen x_0^k und $x_0^k+h_k$ liegt.

D.h. aber gerade, dass f in x_0 total differenzierbar ist.

□

Insbesondere gilt das leichter einprägsame

Korollar. f ist auf U stetig differenzierbar $\iff$ f ist auf U stetig partiell differenzierbar.

Bemerkung. Aus der Ungleichung

$$|f(x_0+h)-f(x_0)-\nabla f(x_0)\cdot h| \leq \|\nabla f(\xi)-\nabla f(x_0)\|_2 \cdot \|h\|_2$$

aus dem obigen Beweis lässt sich zusätzlich folgendes Resultat ablesen:

Korollar. Ist $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ partiell differenzierbar in einer Umgebung von $x_0 \in U$ und sind die partiellen Ableitungen von f beschränkt auf dieser Umgebung, dann ist f stetig in x_0.

7.2 Ableitungsregeln und Mittelwertsätze

Seien weiterhin $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ Banachräume und U eine offene, nicht-leere Teilmenge des jeweils angegebenen Banachraums.

7.2.1 Differenzieren ist linear

Satz 7.2.1. Seien $f,g : U \subseteq X \to Y$ in $x_0 \in U$ differenzierbar, $\lambda,\mu \in \mathbb{R}$. Dann ist auch $\lambda f+\mu g : U \subseteq X \to Y$ differenzierbar in x_0 und es gilt

$$(\lambda f+\mu g)'(x_0) = \lambda f'(x_0)+\mu g'(x_0)$$

Mit anderen Worten: Differenzieren ist eine lineare Operation. Insbesondere ist $C^1(U,Y)$ ein Untervektorraum von $C(U,Y)$.

Beweis. Nach Voraussetzung gilt für alle $x \in U$

$$\begin{aligned} f(x) &= f(x_0)+f'(x_0)(x-x_0)+r_f(x)\cdot\|x-x_0\|_X \\ g(x) &= g(x_0)+g'(x_0)(x-x_0)+r_g(x)\cdot\|x-x_0\|_X \end{aligned}$$

wobei $\lim_{x\to x_0} r_f(x) = \lim_{x\to x_0} r_g(x) = 0$ gilt. Daraus folgt

$$\begin{aligned} \lambda f(x)+\mu g(x) = (\lambda f(x_0)+\mu g(x_0)) &+ \overbrace{(\lambda f'(x_0)+\mu g'(x_0))}^{\in\mathcal{L}(X,Y)}(x-x_0) \\ &+ \underbrace{(\lambda r_f(x)+\mu r_g(x))}_{=:r(x)}\cdot\|x-x_0\|_X \end{aligned}$$

und $\lim_{x\to x_0} r(x) = 0$, womit folgt: $\lambda f+\mu g$ ist differenzierbar in x_0 mit der Ableitung

$$(\lambda f+\mu g)'(x_0) = \lambda f'(x_0)+\mu g'(x_0)$$

□

7.2.2 Kettenregel

Satz 7.2.2. Sei $(Z, \|\cdot\|_Z)$ ein weiterer Banachraum, $\emptyset \neq V \subseteq Y$ offen. Sei $f : U \subseteq X \to Y$ differenzierbar in $x_0 \in U$ und $g : V \subseteq Y \to Z$ differenzierbar in $f(x_0)$ und es gilt $f(U) \subseteq V$. Dann ist $g \circ f : U \to Z$ differenzierbar in x_0 und es gilt

$$(g \circ f)'(x_0) = \underbrace{g'(f(x_0))}_{\in \mathcal{L}(Y,Z)} \circ \underbrace{f'(x_0)}_{\in \mathcal{L}(X,Y)}$$

Beweis. Es ist $g'(f(x_0)) \circ f'(x_0) \in L(X, Z)$. Weiter ist nach Voraussetzung

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r_f(x) \|x - x_0\|_X \quad \forall\, x \in U$$

mit $\lim_{x \to x_0} r_f(x) = 0$, und

$$g(y) = g(f(x_0)) + g'(f(x_0))(y - f(x_0)) + r_g(y) \|y - f(x_0)\|_Y \quad \forall\, y \in V$$

mit $\lim_{y \to f(x_0)} r_g(y) = 0$. Setzt man

$$r(x) := g'(f(x_0))r_f(x) + r_g(f(x)) \left\| f'(x_0) \frac{(x - x_0)}{\|x - x_0\|_X} + r_f(x) \right\|_Y \quad \forall\, x \in U \setminus \{x_0\}$$

so erhält man durch Einsetzen von $y = f(x)$ in die Darstellung von g:

$$g(f(x)) = g(f(x_0)) + g'(f(x_0))f'(x_0)(x - x_0) + r(x) \|x - x_0\|_X \quad \forall\, x \in U$$

Es ist klar, dass $\lim_{x \to x_0} r(x) = 0$:

$$r(x) = \underbrace{g'(f(x_0))r_f(x)}_{\to 0} + \underbrace{r_g(f(x))}_{\to 0} \left\| f'(x_0) \underbrace{\frac{(x - x_0)}{\|x - x_0\|_X}}_{\text{beschränkt}} + \underbrace{r_f(x)}_{\to 0} \right\|_Y$$

D.h. aber gerade, dass $g \circ f$ im Punkte x_0 differenzierbar ist mit

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$$

□

Bemerkung. Falls $X = \mathbb{R}^n, Y = \mathbb{R}^m, Z = \mathbb{R}^l$, dann ergibt sich die Kettenregel in Koordinatendarstellung und die Jacobimatrix von $g \circ f$ als Produkt der Jacobimatrizen von g und f:

$$(g \circ f)'(x_0) = \left(\sum_{i=1}^{m} \frac{\partial g_j}{\partial y_i}(f(x_0)) \cdot \frac{\partial f_i}{\partial x_k} \right)_{\substack{j=1,\ldots,l \\ k=1,\ldots,n}}$$

7.2.3 Produktregel, Ableitung bilinearer Funktionen

Für reellwertige Funktionen gilt auch eine Produktregel:

Satz 7.2.3. Seien $f, g : U \subseteq X \to \mathbb{R}$ in $x_0 \in U$ differenzierbar. dann ist auch $f \cdot g : U \subseteq X \to \mathbb{R}$ differenzierbar und es gilt

$$(f \cdot g)'(x_0) = g(x_0) \cdot f'(x_0) + f(x_0) \cdot g'(x_0)$$

Der Satz ist ein Spezialfall des Satzes über Ableitung bilinearer Funktionen

Satz. Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y), (Z, \|\cdot\|_Z)$ Banachräume, X, Y nur von endlicher Dimension. Sei $\beta : X \times Y \to Z$ eine bilineare Abbildung, $(x_0, y_0) \in X \times Y$, so ist β in (x_0, y_0) differenzierbar mit

$$\beta'(x_0, y_0)(a_1, a_2) = \beta(x_0, a_2) + \beta(a_1, y_0)$$

Beweis. O.B.d.A. kann man annehmen: $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$. Zeige zunächst, dass ein $c \in \mathbb{R}$ existiert, sodass

$$\|\beta(x,y)\|_Z \leq c \cdot \|(x,y)\|^2 \quad \forall\,(x,y) \in X \times Y$$

bezüglich einer Norm $\|\cdot\|$ auf $X \times Y$.

Es ist $X \times Y \cong \mathbb{R}^{n+m}$ ein endlich-dimensionaler Banachraum, und da auf solchen alle Normen äquivalent sind, können wir uns ein spezielle Normen auf $X \times Y$ wählen. Sei $(x,y) = (\sum \lambda_i e_i, \sum \mu_i e_i)$, wobei e_i die kanonischen Basisvektoren seien. Dann ist $\|\cdot\| : X \times Y \to \mathbb{R}$ definiert durch

$$\|(x,y)\| = \left\| \left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{i=1}^{m} \mu_i e_i \right) \right\| := \max\left\{|\lambda_1|, \ldots, |\lambda_n|, |\mu_1|, \ldots, |\mu_m|\right\}$$

eine Norm (klar, man kann sie als Supremumsnorm auf $\mathbb{R}^{n+m}$ auffassen). Damit gilt

$$\begin{aligned}
\|\beta(x,y)\|_Z &= \left\| \beta\left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{i=1}^{m} \mu_i e_i \right) \right\|_Z = \left\| \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j \beta\left(e_i e_i\right) \right\|_Z \\
&\leq \sum_{i=1}^{n} \sum_{j=1}^{m} |\lambda_i|\,|\mu_j|\, \|\beta\left(e_i e_i\right)\|_Z \\
&\leq \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\max\left\{|\lambda_1|, \ldots, |\lambda_n|, |\mu_1|, \ldots, |\mu_m|\right\}\right)^2 \|\beta\left(e_i e_i\right)\|_Z \\
&\leq \underbrace{\left(\sum_{i=1}^{n} \sum_{j=1}^{m} \|\beta\left(e_i e_i\right)\|_Z \right)}_{=\text{const}} \cdot \|(x,y)\|^2
\end{aligned}$$

Nun ist für festes $(x_0, y_0) \in X \times Y$ und $(h_1, h_2) \in X \times Y$

$$\beta(x_0 + h_1, y_0 + h_2) = \beta(x_0, y_0) + \beta(x_0, h_2) + \beta(h_1, y_0) + \beta(h_1, h_2)$$

Wie man leicht einsieht, ist $\tilde{\beta} : X \times Y \to Z$ definiert durch

$$\tilde{\beta}(h_1, h_2) := \beta(h_1, y_0) + \beta(x_0, h_2)$$

linear. Außerdem gilt

$$\frac{\|\beta(h_1,h_2)\|_Z}{\|(h_1,h_2)\|} \leq \frac{c \cdot \|(h_1,h_2)\|^2}{\|(h_1,h_2)\|} = c \cdot \|(h_1,h_2)\| \overset{\|(h_1,h_2)\| \to 0}{\longrightarrow} 0$$

Damit ist gezeigt, dass β in (x_0, y_0) differenzierbar ist mit

$$\beta'(x_0, y_0) = \tilde{\beta}$$

$\square$

7.2.4 Spezieller Schrankensatz

Erinnerung: Falls $f : [a, b] \to \mathbb{R}$ stetig und f auf $]a, b[$ differenzierbar ist, dann existiert ein $\xi \in]a, b[$ mit $f(b) - f(a) = f'(\xi)\,(b - a)$.

Für differenzierbare Funktionen $f : [a, b] \to Y$, wobei $(Y, \|\cdot\|_Y)$ ein Banachraum ist, gilt ein analoges Resultat nicht (auch für keinen endlich-dimensionalen Raum Y), wie wir weiter unten sehen werden. Aber aus dem Mittelwertsatz der reellen Analysis folgt

$$|f(b) - f(a)| \leq \sup_{\xi \in]a,b[} |f'(\xi)| \cdot (b - a)$$

Dieser Schrankensatz lässt sich auch für vektorwertige Funktionen beweisen

Satz 7.2.4. Sei $f \in C([a, b], Y)$, f differenzierbar auf $]a, b[$. Dann gilt

$$\|f(b) - f(a)\|_Y \leq \sup_{t \in]a,b[} \|f'(t)\|_{\mathcal{L}(\mathbb{R},Y)}(b - a)$$

Beweis. O.B.d.A. kann man

$$M := \sup_{t \in]a,b[} \|f'(t)\|_{\mathcal{L}(\mathbb{R},Y)} < \infty$$

annehmen (ansonsten ist nichts zu zeigen). Sei nun $\epsilon > 0$ (so klein, dass $a + \epsilon < b$). Wir zeigen

$$\|f(b) - f(a + \epsilon)\| \leq M\,(b - a - \epsilon)$$

Angenommen, dies gilt, dann folgt für $\epsilon \to 0$ wegen der Stetigkeit von f die Behauptung. Definiere

$$S := \{x \in [a, b] \mid \|f(x) - f(a + \epsilon)\|_Y \leq M(x - a - \epsilon)\}$$

Dann gilt $S \neq \emptyset$, denn $a + \epsilon \in S$. Da f stetig ist, ist S auch abgeschlossen. Folglich existiert

$$s := \max S \in [a + \epsilon, b]$$

Ist $s = b$, so folgt die obige Ungleichung. Angenommen es gelte $s < b$. Da f auf $]a, b[$, also wegen $a < s$ insbesondere in s differenzierbar ist, gilt

$$\frac{\|f(x) - f(s)\|}{x - s} \xrightarrow{x \searrow s} \|f'(s)\|$$

Da $\|f'(s)\| \leq M$, existiert ein $\delta > 0$ mit $s + \delta \leq b$, sodass

$$\|f(x) - f(s)\| \leq M(x - s) \quad \forall\, x \in]s, s + \delta[$$

Für ein solches $x \in]s, s + \delta[$ gilt

$$\begin{aligned} \|f(x) - f(a + \epsilon)\| &\leq \|f(x) - f(s)\| + \|f(s) - f(a + \epsilon)\| \\ &\leq M(x - s) + M(s - a - \epsilon) = M(x - a - \epsilon) \end{aligned}$$

was der Maximalität von s widerspricht. Also muss doch $s = b$ sein. $\square$

7.2.5 Allgemeiner Schrankensatz

Nun folgt leicht der allgemeine Schrankensatz

Satz 7.2.5. Sei $f : U \subseteq X \to Y$ differenzierbar. Sind $x, y \in U$, sodass auch die gesamte Verbindungsstrecke zwischen x und y noch in U liegt, d.h.

$$[[x,y]] := \{x + t \cdot (y-x) \mid t \in [0,1]\} \subseteq U$$

dann gilt

$$\|f(x) - f(y)\|_Y \leq \sup_{t \in]0,1[} \|f'(x + t(y-x))\|_{\mathcal{L}(X,Y)} \|x-y\|_X$$

Beweis. Definiere

$$\begin{aligned} \varphi : \ [0,1] &\to X \\ t &\mapsto x + t(y-x) \end{aligned}$$

Da $[[x,y]] \subseteq U$ gilt tatsächlich $\varphi([0,1]) \subseteq U$. Die damit wohldefinierte Komposition $f \circ \varphi$ der differenzierbaren Abbildungen φ und f ist stetig auf $[0,1]$ und nach Kettenregel differenzierbar mit

$$(f \circ \varphi)'(t) = f'(x + t(y-x))(y-x) \quad \forall\, t \in [0,1]$$

Somit folgt nach Satz 7.2.4

$$\begin{aligned} \|(f\circ\varphi)(1) - (f\circ\varphi)(0)\|_Y &\leq \sup_{t \in]0,1[} \left\|(f\circ\varphi)'(t)\right\|_{\mathcal{L}(\mathbb{R},Y)}(1-0) \\ &= \sup_{t \in]0,1[} \|f'(x+t(y-x))(y-x)\|_{\mathcal{L}(\mathbb{R},Y)} \\ &\leq \sup_{t \in]0,1[} \|f'(x+t(y-x))\|_{\mathcal{L}(X,Y)} \|y-x\|_X \end{aligned}$$

also wie gewünscht

$$\|f(y) - f(x)\|_Y \leq \sup_{t \in]0,1[} \|f'(x + t(y-x))\|_{\mathcal{L}(X,Y)} \|y-x\|_X$$

□

Hier anschließend folgt nun ein Gegenbeispiel für die allgemeine Ungültigkeit des Mittelwertsatzes. Betrachte

$$\begin{aligned} f : \ \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) &\mapsto (x^2, y^3) \end{aligned}$$

Es seien $(x,y) = (0,0)$ und $(\hat{x},\hat{y}) = (1,1)$. Zeige, dass es kein $\xi \in [[(x,y),(\hat{x},\hat{y})]]$ gibt, mit

$$(1,1) - (0,0) = Df(\xi)((1,1) - (0,0))$$

Es gilt

$$Df(u,u) = \begin{pmatrix} 2u & 0 \\ 0 & 3u^2 \end{pmatrix} \quad \forall\, u \in \mathbb{R}$$

Damit ist

$$Df(u,u)(1,1) = \begin{pmatrix} 2u & 0 \\ 0 & 3u^2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2u \\ 3u^2 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \forall\, u \in \mathbb{R}$$

7.2.6 Konvexe Mengen

Bemerkung. Eine Teilmenge $U \subseteq X$ mit $[[x,y]] \subseteq U$ für alle $x, y \in U$ heißt **konvex**.

Mit diesem Begriff lassen sich nun folgende Korollare formulieren:

Korollar 7.2.6. Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ wie immer Banachräume, $\emptyset \neq U$ eine offene, konvexe Teilmenge von X, $f : U \to Y$ differenzierbar. Dann gilt

$$\|f(x) - f(y)\|_Y \leq \underbrace{\sup_{t \in U} \|f'(t)\|_{\mathcal{L}(X,Y)}}_{=:M} \|x - y\|_X \quad x, y \in U$$

Im Falle $M < \infty$ ist f lipschitz-stetig mit Konstante M.

Korollar 7.2.7. Sei $\emptyset \neq U \subseteq X$ offen, $f : U \to Y$ differenzierbar. Dann gilt

$$\boxed{f \text{ konstant} \iff f'(x) = 0 \quad \forall\, x \in U}$$

7.2.7 Mittelwertsatz in Integralform

Im Spezialfall $Y = \mathbb{R}$ gilt folgender Mittelwertsatz,

Satz 7.2.8. Sei $\emptyset \neq U \subseteq X$ offen, $f : U \to \mathbb{R}$ stetig differenzierbar. Dann ist für $x, y \in U$ mit $[[x,y]] \subseteq U$

$$f(y) - f(x) = \int_0^1 f'(x + t(y - x))(y - x)\,\mathrm{d}t$$

Beweis. Seien $x, y \in U$ mit $[[x,y]] \subseteq U$ fest. Definiere

$$\begin{aligned} \varphi :\ & [0,1] \to \mathbb{R} \\ & \varphi(t) = f(x + t(y - x)) \end{aligned}$$

φ ist wegen $x + t(y - x) \in U$ für alle $t \in U$ wohldefiniert und als Komposition stetig differenzierbarer Funktionen stetig differenzierbar. Nach dem Hauptsatz der Differenzial- und Integralrechnung (Satz 5.2.11) folgt die Behauptung:

$$\begin{aligned} \varphi(1) - \varphi(0) &= \int_0^1 \varphi'(t)\,\mathrm{d}t && \Leftrightarrow \\ f(y) - f(x) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f(x + t(y - x))\,\mathrm{d}t = \int_0^1 f'(x + t(y - x))(y - x)\,\mathrm{d}t \end{aligned}$$

□

7.2.8 Integralrechnung mit vektorwertigen Funktionen

Um dieses Resultat auf vektorwertige Funktionen zu übertragen, benötigt man den Begriff des Riemann-Integrals für Funktionen $f : [a, b] \to Y$, die ihre Werte in Banachräumen oder allgemeiner in normierten Vektorräumen annehmen. Dies wollen wir kurz tun:

Für jede Zerlegung $Z = \{a = t_0, t_1, \ldots, t_n = b\}$ von $[a, b]$ und jeden Zwischenvektor $\xi = (\xi_1, \ldots, \xi_n)$ erklärt man die **Riemann'sche Summe**

$$\sigma(Z, \xi, f) = \sum_{k=1}^{n} (t_k - t_{k-1}) \cdot f(\xi_k)$$

Eine **Riemann'sche Summenfolge** ist eine Folge von Riemann'schen Summen $\sigma(Z_n, \xi_n^k, f)$ mit $(Z_n)_{n\in\mathbb{N}}$ Zerlegungsnullfolge und $(\xi_k^n)_{n\in\mathbb{N}}$ beliebige zugehörige Zwischenvektorenfolge.

Strebt nun jede Riemann'sche Summenfolge gegen einen - und damit notwendigerweise ein- und denselben - Grenzwert in Y, so nennt man f **Riemann-integrierbar** auf $[a, b]$, und den Grenzwert der Riemann-Folge bezeichnet man mit

$$\int_a^b f(t)\,\mathrm{d}t \quad (\in Y)$$

Wie im Reellen definiert man

$$\int_a^b f(t)\,\mathrm{d}t = -\int_b^a f(t), \int_a^a f(t)\,\mathrm{d}t = 0$$

und wie im Reellen gelten auch folgende Sätze

1. Cauchy'sches Integrabilitätskriterium (nur wenn Y ein Banachraum ist):

   Satz. $f : [a, b] \to Y$ ist Riemann-integrierbar auf $[a, b] \iff$

   $$\forall \epsilon > 0 \quad \exists \delta > 0 \ : \ \left\|\sigma(Z, \xi, f) - \sigma\left(\tilde{Z}, \tilde{\xi}, f\right)\right\|_Y < \epsilon$$

   für alle Zerlegungen $Z, \tilde{Z}$, für die $|Z|, |Z| < \delta$ gilt.

2. Jede stetige Funktion $f : [a, b] \to Y$ ist Riemann-integrierbar.

3. Es gilt die Dreiecksungleichung

   $$\left\|\int_a^b f(t)\,\mathrm{d}t\right\|_Y \leq \int_a^b \|f(t)\|_Y$$

   und damit insbesondere

   $$\left\|\int_a^b f(t)\,\mathrm{d}t\right\|_Y \leq \sup_{t\in[a,b]} \|f(t)\|_Y (b - a)$$

4. Für alle $p_1, p_2, p_3 \in [a, b]$ ist

   $$\int_{p_1}^{p_2} f(t)\,\mathrm{d}t + \int_{p_2}^{p_3} f(t)\,\mathrm{d}t = \int_{p_1}^{p_3} f(t)\,\mathrm{d}t$$

5. Hauptsatz der Differential- und Integralrechnung für Banachraum-wertige Funktionen:

   Satz. Wenn $f: [a,b] \to Y$ stetig ist, $F: [a,b] \to Y$ definiert durch

   $$F(x) = \int_a^x f(t)\,\mathrm{d}t \quad \forall\, x \in [a,b]$$

   Dann ist F stetig differenzierbar und es gilt $F'(x) = f(x)$ für alle $x \in [a,b]$.

   Beweis. Es ist für $x, x+h \in [a,b]$

   $$F(x+h) - F(x) = \int_x^{x+h} f(t)\,\mathrm{d}t \qquad \Rightarrow$$

   $$F(x+h) = F(x) + f(x)h + \underbrace{\int_x^{x+h} (f(t) - f(x))\,\mathrm{d}t}_{=:\tilde{r}(x+h)}$$

   Da

   $$\|\tilde{r}(x+h)\|_Y \leq |h| \max_{t\in[x,x+h]} \|f(t) - f(x)\|_Y$$

   folgt mit

   $$r(x+h) := \begin{cases} \frac{\tilde{r}(x+h)}{h}, & h \neq 0, \\ 0, & h = 0 \end{cases}$$

   dass

   $$F(x+h) = F(x) + f(x)h + r(x+h)\,|h| \quad \forall\, x, x+h \in [a,b]$$

   und wegen der Stetigkeit von f: $\lim_{h\to 0} r(x+h) = 0$.
   Damit ist gezeigt, dass F in allen Punkten $x \in [a,b]$ differenzierbar ist mit $F' = f$. □

Beachten wir weiter, dass für eine stetige Funktion $A: [a,b] \ni t \mapsto A(t) \in \mathcal{L}(X,Y)$, wobei $(X, \|\cdot\|_x), (Y, \|\cdot\|_Y)$ Banachräume sind, gilt

$$\int_a^b A(t)h\,\mathrm{d}t = \left(\int_a^b A(t)\,\mathrm{d}t \right) h \quad \forall\, h \in X$$

Dies folgt aus der entsprechenden Beziehung der zugehörigen Riemann-Summen

$$\sum_{k=1}^{n} (t_k - t_{k-1})\, A(\xi_k)h = \left(\sum_{k=1}^{n} (t_k - t_{k-1})\, A(\xi_k) \right) h$$

7.2.9 Allgemeiner Mittelwertsatz in Integralform

Nun können wir den allgemeinen Mittelwertsatz in Integralform formulieren.

Satz 7.2.9. Sei $f : U \subseteq X \to Y$ stetig differenzierbar, $x, y \in U$ mit $[[x, y]] \subseteq U$. Dann gilt

$$f(y) - f(x) = \int_0^1 f'(x + t(y-x))\,\mathrm{d}t(y-x)$$

Beweis. Analog wie im reellwertigen Fall definieren wir die Hilfsfunktion

$$\begin{aligned} \varphi :\ &[a, b] \to X \\ &\varphi(t) = f(x + t(y-x)) \quad t \in [a, b] \end{aligned}$$

φ ist stetig differenzierbar und mit dem Hauptsatz der Differenzial- und Integralrechung für vektorwertige Funktionen folgt

$$\begin{aligned} \varphi(1) - \varphi(0) &= \int_0^1 \varphi'(t)\,\mathrm{d}t && \Leftrightarrow \\ f(y) - f(x) &= \int_0^1 f'(x + t(y-x))\,\mathrm{d}t(y-x) \end{aligned}$$

□

7.3 Höhere Ableitungen und Taylor'sche Formel

Seien weiterhin $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banachräume und $U \neq \emptyset$ offene Teilmenge des jeweils angegebenen Raumes.

7.3.1 Höhere Ableitungen

Sei $f : U \subseteq X \to Y$ differenzierbar, dann ist es möglich, dass die Ableitung $f' : U \to \mathcal{L}(X, Y)$ wieder in einem Punkt $x_0 \in U$ respektive auf ganz U differenzierbar ist.

Definition. Die Ableitung $(f')'(x_0) =: f''(x_0)$ heißt dann entsprechend **2. Ableitung** von f in x_0, respektive heißt die Abbildung

$$\begin{aligned} f'' :\ U &\to \mathcal{L}(X, \mathcal{L}(X, Y)) \\ x &\mapsto (f')'(x_0) \end{aligned}$$

(falls existent) die **2. Ableitung** von f auf U. f heißt in diesem Fall **2-mal differenzierbar** in x_0 respektive auf U (und 2-mal stetig differenzierbar auf U, wenn f'' stetig ist).

Nun ist klar, wie man induktiv die k-te Ableitung und k-fache (stetige) Differenzierbarkeit definieren kann. Die k-te Ableitung hat die Signatur

$$f^{(k)} :\ U \to \mathcal{L}(\underbrace{X, (X, \ldots, \mathcal{L}(X}_{k\text{-mal}}, Y)) \ldots)$$

Das Problem ist: Die Räume werden immer komplizierter. Die lineare Algebra schafft hier Abhilfe; es gilt nämlich

$$\mathcal{L}(X, \mathcal{L}(X, Y)) \cong \mathcal{L}(X \times X, Y)$$

wobei eine stetige lineare Abbildung $A \in \mathcal{L}(X, \mathcal{L}(X, Y))$ der stetigen **bilinearen Abbildung** $X \times X \ni (x, \tilde{x}) \mapsto (Ax)\tilde{x}$ entspricht und umgekehrt. Iteriert man dieses Verfahren, so erhält man

$$\mathcal{L}(X, (X, \ldots, \mathcal{L}(X, Y)) \ldots) \cong \underbrace{\mathcal{L}(X \times X \times \ldots \times X, Y)}_{\text{stetige multilineare Abbildungen}}$$

7.3.2 Mehrfache partielle Ableitungen

Bevor wir diese allgemeine Situation betrachten, wenden wir uns zunächst einmal dem einfachen Fall von reellwertigen Funktionen $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ zu und studieren mehrfache partielle Differenzierbarkeit:

Definition. Besitzt eine solche Funktion f partielle Ableitungen $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$, so können diese wieder partiell differenzierbar sein. In diesem Fall bezeichnet man die Funktionen

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) =: \frac{\partial^2}{\partial x_i \partial x_j} f =: f_{x_i x_j} \quad i, j \in \{1, \ldots, n\}$$

als **partielle Ableitungen 2. Ordnung** von f.

Damit ist klar, wie man induktiv partielle Ableitungen k-ter Ordnung definieren kann. Notation: $f_{x_{i_k} x_{i_{k-1}} \ldots x_{i_1}}$ oder ausführlich

$$\frac{\partial}{\partial x_{i_k}} \frac{\partial}{\partial x_{i_{k-1}}} \cdots \frac{\partial}{\partial x_{i_1}} f =: \frac{\partial^k}{\partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} f \qquad i_1, \ldots, i_k \in \{1, \ldots, n\}$$

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_i} \cdots \frac{\partial}{\partial x_i} f =: \frac{\partial^k}{\partial x_i^k} f \qquad i \in \{1, \ldots, n\}$$

7.3.3 Beispiel

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto x^2 \cdot \sin(y) \end{aligned}$$

besitzt partielle Ableitungen beliebig hoher Ordnung, insbesondere ist

$$\begin{aligned}
\frac{\partial}{\partial x} f(x, y) &= f_x(x, y) = 2x \sin(y) \quad \forall\, x, y \in \mathbb{R}^2 \\
\frac{\partial}{\partial y} f(x, y) &= f_y(x, y) = x^2 \cos(y) \quad \forall\, x, y \in \mathbb{R}^2 \\
\frac{\partial}{\partial y \partial x} f(x, y) &= f_{yx}(x, y) = \frac{\partial}{\partial y}(2x \sin(y)) = 2x \cos(y) \quad \forall\, x, y \in \mathbb{R}^2 \\
\frac{\partial}{\partial x \partial y} f(x, y) &= f_{xy}(x, y) = \frac{\partial}{\partial x}(x^2 \cos(y)) = 2x \cos(y) \quad \forall\, x, y \in \mathbb{R}^2
\end{aligned}$$

d.h. in diesem Fall gilt $f_{xy} = f_{yx}$, die gemischten Ableitungen stimmen überein.

Dies gilt im Allgemeinen so nicht, betrachte zum Beispiel

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases} \end{aligned}$$

f ist stetig differenzierbar auf $\mathbb{R}^2$, f_{xy} und f_{yx} existieren auf $\mathbb{R}^2$ und sind stetig in $\mathbb{R}^2 \setminus \{(0,0)\}$, aber es ist $f_{xy}(0,0) = 1$ und $f_{yx}(0,0) = -1$.

Es ist auch möglich, dass nur eine der beiden gemischten partiellen Ableitungen 2. Ordnung existiert.

7.3.4 Satz von Schwarz

Solche Situationen treten nicht auf, wenn die gemischte partielle Ableitung 2. Ordnung zusätzlich noch im betrachteten Punkt stetig ist. Das ist die Aussage des

Satz 7.3.1. Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. Für eine gewisse Wahl von $i, j \in \{1, \ldots, n\}$ mögen die partiellen Ableitungen $\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j}$ und $\frac{\partial^2 f}{\partial x_i \partial x_j}$ in einer Umgebung V eines Punktes x_0 existieren.

Wenn $\frac{\partial^2}{x_i \partial x_j} f$ in x_0 stetig ist, dann existiert auch die gemischte partielle Ableitung $\frac{\partial^2}{x_j \partial x_i} f$ in x_0 und es gilt

$$\frac{\partial^2}{x_i \partial x_j} f(x_0) = \frac{\partial^2}{x_j \partial x_i} f(x_0)$$

Beweis. Wir wollen zeigen, dass

$$\frac{\partial^2}{\partial x_j \partial x_i} f(x_0) = \lim_{t \to 0} \frac{\frac{\partial f}{\partial x_i}(x_0 + t \cdot e_j) - \frac{\partial f}{\partial x_i}(x_0)}{t}$$

existiert und $= \frac{\partial^2}{\partial x_i \partial x_j} f(x_0)$ ist. Es ist zumindest

$$\ldots = \lim_{t \to 0} \left[\frac{1}{t} \lim_{s \to 0} \left(\frac{f(x_0 + se_i + te_j) - f(x_0 + te_j)}{s} - \frac{f(x_0 + se_i) - f(x_0)}{s} \right) \right]$$

Definieren wir für genügend kleine $t \in \mathbb{R}$

$$\Phi(t) := f(x_0 + se_i + te_j) - f(x_0 + te_j)$$

dann ist Φ wohldefiniert und differenzierbar, da f partiell in einer Umgebung von x_0 nach der j-ten Koordinate differenzierbar ist. Nach dem Mittelwertsatz der reellen Analysis ergibt sich für ein $\xi \in \,]-1, 1[$

$$\begin{aligned} \Phi(t) - \Phi(0) &= \Phi'(\xi t)t \\ &= \left(\frac{\partial f}{\partial x_j}(x_0 + se_i + \xi te_j) - \frac{\partial f}{\partial x_j}(x_0 + \xi te_j) \right) t \end{aligned}$$

Damit ergibt sich

$$\frac{\partial^2}{\partial x_j \partial x_i} f(x_0) = \lim_{t\to 0} \lim_{s\to 0} \frac{\frac{\partial f}{\partial x_j}(x_0 + se_i + \xi te_j) - \frac{\partial f}{\partial x_j}(x_0 + \xi te_j)}{s}$$

Wir betrachten nun die Funktion

$$\psi(s) = \frac{\partial f}{\partial x_j}(x_0 + se_i + \xi te_j) - \frac{\partial f}{\partial x_j}(x_0 + \xi te_j)$$

Diese ist für genügend kleine s wohldefiniert und auch differenzierbar, da $\frac{\partial^2}{\partial x_i \partial x_j} f$ einer Umgebung von x_0 existiert. Somit folgt wieder aus dem Mittelwertsatz

$$\psi(s) - \psi(0) = \psi'(\eta s)s = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + \eta se_i + \xi te_j)$$

für ein $\eta \in \,]-1, 1[$ und es folgt

$$\frac{\partial^2}{\partial x_j \partial x_i} f(x_0) = \lim_{t\to 0} \lim_{s\to 0} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + \underbrace{\eta se_i}_{\to 0} + \underbrace{\xi te_j}_{\to 0}) = \frac{\partial^2}{\partial x_i \partial x_j} f(x_0)$$

da $\frac{\partial^2}{\partial x_i \partial x_j} f$ in x_0 stetig ist. □

Korollar 7.3.2. Sei $q \in \mathbb{N}$, $q \geq 2$. Wenn $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ q-mal stetig partiell differenzierbar ($\Leftrightarrow$ q-mal (total) differenzierbar auf U) ist, dann gilt für $k \leq q$.

$$\frac{\partial^k}{\partial x_{i_k} \partial x_{i_{k-1}} \dots \partial x_{i_1}} f = \frac{\partial^k}{\partial x_{i_{\pi(k)}} \partial x_{i_{\pi(k-1)}} \dots \partial x_{i_{\pi(1)}}} f$$

für jede Permutation $\pi : \{1, \dots, k\} \leftrightarrow \{1, \dots, k\}$.

7.3.5 Hesse-Matrix

Sei nun $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ differenzierbar, sodass

$$\begin{aligned} f' : U &\to \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^n \\ x &\mapsto f'(x) \end{aligned}$$

bezüglich der kanonischen Basis im $\mathbb{R}^m$ dargestellt durch

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial_n}(x) \right)$$

stetig differnzierbar ist, d.h. $f \in C^2(U) = C^2(U, \mathbb{R})$. dann ist

$$f'' = (f')' : U \to \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$$

und f'' kann mit einer $n \times n$-Matrix identifiziert werden. Bezüglich der kanonischen Basis im $\mathbb{R}^n$ hat diese die Form

$$\begin{aligned} J_{\nabla f}(x_0) &= \begin{pmatrix} \frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_1} f\right)(x_0) & \dots & \frac{\partial}{\partial x_n}\left(\frac{\partial}{\partial x_1} f\right)(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_n} f\right)(x_0) & \dots & \frac{\partial}{\partial x_n}\left(\frac{\partial}{\partial x_n} f\right)(x_0) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} f(x_0) & \dots & \frac{\partial^2}{\partial x_n x_1} f(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_1 \partial x_n} f(x_0) & \dots & \frac{\partial^2}{\partial x_n x_n} f(x_0) \end{pmatrix} \end{aligned}$$

Diese Matrix heißt **Hesse Matrix** von f in x_0, bezeichnet mit $\mathrm{Hess} f(x)$ oder $H_f(x)$. Nach dem Satz von Schwarz ist diese symmetrisch.

Wir wollen nun $f''(x_0)$ als symmetrische Bilinearform auffassen:

$$\begin{aligned} f''(x_0):\ \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (u, v) &\mapsto f''(x_0)(u, v) \end{aligned}$$

wobei

$$\begin{aligned} f''(x_0)(u,v) = (f''(x_0)u)\, v &= \left(\mathrm{Hess} f(x_0) \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}\right) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^n \frac{\partial^2}{\partial x_i \partial x_1} f(x_0) u_i \\ \vdots \\ \sum_{i=1}^n \frac{\partial^2}{\partial x_i \partial x_n} f(x_0) u_i \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &= \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(x_0) u_i v_j \\ &= u^T \mathrm{Hess} f(x_0) v = v^T \mathrm{Hess} f(x_0) u \\ &= D_u D_v f(x_0) = D_v D_u f(x_0) \end{aligned}$$

7.3.6 Multiindizes

Analog kann man für $f \in C^k(U)$ die k-te Ableitung als k-lineare Abbildung auffassen:

$$\begin{aligned} f^{(k)}(x_0) = D^k f(x_0):\ \mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n &\to \mathbb{R} \\ (v_1, \ldots, v_n) &\mapsto D_{v_1} \ldots D_{v_n} f(x_0) \end{aligned}$$

Nützlich ist nun für höhere Ableitungen folgende Notation:

Für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ (ein sogenannter **Multiindex**) setzt man

$$|\alpha| := \sum_{i=1}^{n} \alpha_i = \alpha_1 + \ldots + \alpha_n$$

$$\alpha! := \prod_{i=1}^{n} \alpha_i! = \alpha_1! \cdot \ldots \cdot \alpha_n!$$

und für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$

$$x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$$

$$D^\alpha f := D_1^{\alpha_1} \ldots D_n^{\alpha_n} f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} f$$

7.3.7 Differentialoperatoren, Laplace-Operator

Man kann nun bequem sogenannte lineare Differentialoperatoren definieren. Zunächst kann man mit Hilfe der Multiindizes Polynome in n Variablen darstellen. Ein Polynom p in n Variablen $\xi_1, \ldots, \xi_n$ vom Grad $m \in \mathbb{N}_0$ hat die Form

$$p(\xi_1, \ldots, \xi_n) = \sum_{|\alpha| \le m} a_\alpha \xi_1^{\alpha_1} \cdot \ldots \cdot \xi_n^{\alpha_n}$$

Ersetzt man nun in einem solchen Polynom die Variablen ξ_i durch partielle Ableitungsoperatoren $\frac{\partial}{\partial x_i}$, so erhält man einen linearen Differentialoperator

$$p(D) = \sum_{|\alpha| \le m} a_\alpha \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$$

genauer

$$\begin{aligned} p(D) :\ & C^m(U) \to C(U) \\ & f \mapsto \sum_{|\alpha| \le m} a_\alpha \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f \end{aligned}$$

Zum Beispiel erzeugt das Polynom

$$p(\xi_1, \ldots, \xi_n) = \xi_1^2 + \ldots + \xi_n^2$$

den Differentialoperator

$$\triangle = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

den sogenannten **Laplace-Operator**.

7.3.8 Eindimensionale Taylor-Formel mit Integralrestglied

Satz. Für $f \in C^m(I)$, wobei I ein Intervall, gilt

$$f(x_0 + h) = \sum_{k=0}^{m} \frac{f^{(k)}(x_0) h^k}{k!} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \left(f^{(m)}(x_0 + th) - f^{(m)}(x_0) \right) h^m \, \mathrm{d}t$$

Beweis. durch vollständige Induktion:
Für $m = 1$ gilt

$$\sum_{k=0}^{1} \frac{f^{(k)}(x_0)h^k}{k!} + \int_0^1 \frac{(1-t)^{1-1}}{(1-1)!} \left(f^{(1)}(x_0+th) - f^{(1)}(x_0)\right) h^1 \,\mathrm{d}t$$

$$= f(x_0) + f'(x_0)h + \int_0^1 \left(f'(x_0+th) - f'(x_0)\right) h \,\mathrm{d}t$$

$$= f(x_0) + f'(x_0)h - f'(x_0)h + \int_0^1 f'(x_0+th)h \,\mathrm{d}t$$

$$= f(x_0) + [f(x_0+th)]_0^1 = f(x_0) + f(x_0+h) - f(x_0) = f(x_0+h)$$

Induktionsschritt $(m-1) \to m$:

$$\overbrace{\sum_{k=0}^{m} \frac{f^{(k)}(x_0)h^k}{k!}}^{=:\Sigma} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \left(f^{(m)}(x_0+th) - f^{(m)}(x_0)\right) h^m \,\mathrm{d}t$$

$$= \Sigma - f^{(m)}(x_0)h^m \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \,\mathrm{d}t + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} f^{(m)}(x_0+th)h^m \,\mathrm{d}t$$

$$= \Sigma - f^{(m)}(x_0)h^m \left[-\frac{(1-t)^m}{m!}\right]_0^1 + \left[\frac{(1-t)^{m-1}}{(m-1)!} f^{(m-1)}(x_0+th)h^{m-1}\right]_0^1$$

$$- \int_0^1 -\frac{(1-t)^{m-2}}{(m-2)!} f^{(m-1)}(x_0+th)h^{m-1} \,\mathrm{d}t$$

$$= \sum_{k=0}^{m} \frac{f^{(k)}(x_0)h^k}{k!} - \frac{f^{(m)}(x_0)h^m}{m!} - \frac{f^{(m-1)}(x_0)h^{m-1}}{(m-1)!}$$

$$+ \int_0^1 \frac{(1-t)^{m-2}}{(m-2)!} f^{(m-1)}(x_0+th)h^{m-1} \,\mathrm{d}t$$

$$= \sum_{k=0}^{m-1} \frac{f^{(k)}(x_0)h^k}{k!} - \int_0^1 \frac{(1-t)^{m-2}}{(m-2)!} f^{(m-1)}(x_0)h^{m-1} \,\mathrm{d}t$$

$$+ \int_0^1 \frac{(1-t)^{m-2}}{(m-2)!} f^{(m-1)}(x_0+th)h^{m-1} \,\mathrm{d}t$$

$$= \sum_{k=0}^{m-1} \frac{f^{(k)}(x_0)h^k}{k!} + \int_0^1 \frac{(1-t)^{m-2}}{(m-2)!} \left(f^{(m-1)}(x_0+th) - f^{(m-1)}(x_0)\right) h^{m-1} \,\mathrm{d}t$$

$$\overset{\mathrm{IV}}{=} f(x_0+h)$$

□

7.3.9 Mehrdimensionaler Satz von Taylor

Satz 7.3.3. Seien $f \in C^{m+1}(U)$, U offene Teilmenge des $\mathbb{R}^n$, $x_0, x_0 + h \in U$ sodass $[[x_0, x_0 + h]] \subseteq U$. Dann existiert ein $\theta \in]0,1[$ mit

$$f(x_0 + h) = \sum_{|\alpha| \leq m} \frac{D^\alpha f(x_0) h^\alpha}{\alpha!} + \sum_{|\alpha| = m+1} \frac{D^\alpha f(x_0 + \theta h)}{\alpha!} h^\alpha$$

Beweis.

1. Schritt: Definiere die Hilfsfunktion

$$\begin{aligned} g :\ [0,1] &\to \mathbb{R} \\ t &\mapsto f(x_0 + th) \end{aligned}$$

Diese ist nach der Kettenregel $(m+1)$-mal stetig differenzierbar und es gilt

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} g(t) = \sum_{i_1, \dots, i_k = 1}^{n} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x_0 + th) h_{i_1} \dots h_{i_k} \quad \forall\, k = 1, \dots, m+1$$

Beweis durch Induktion

Für $k = 1$ gilt:

$$g'(t) = \nabla f(x_0 + th) \cdot h = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x_0 + th) \cdot h_i$$

Induktionsschritt $k \to k+1$:

$$\begin{aligned} g^{(k+1)}(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(g^{(k)}(t) \right) \\ &\overset{\text{IV}}{=} \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i_1, \dots, i_k = 1}^{n} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x_0 + th) h_{i_1} \cdot \ldots \cdot h_{i_k} \right) \\ &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\sum_{i_1, \dots, i_k = 1}^{n} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x_0 + th) h_{i_1} \cdot \ldots \cdot h_{i_k} \right) h_i \\ &\overset{i = i_{k+1}}{=} \sum_{i_1, \dots, i_k, i_{k+1} = 1}^{n} \frac{\partial}{\partial x_{i_{k+1}}} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x_0 + th) h_{i_1} \cdot \ldots \cdot h_{i_k} \cdot h_{i_{k+1}} \end{aligned}$$

2. Schritt: Kommt unter den Indizes $i_1, \dots, i_k$ die 1 α_1-mal vor, die 2 α_2-mal vor, …, n α_n-mal vor, so ist

$$\frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f = \frac{\partial^k}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f$$

Mit Hilfe eines kombinatorischen Arguments überprüft man, dass es unter den k-Tupeln $(i_1, \dots, i_k)$, wobei $1 \leq i_1, \dots, i_k \leq n$, genau $\frac{k!}{\alpha_1 \cdot \ldots \cdot \alpha_n} = \frac{k!}{\alpha!}$ viele gibt, bei denen die Zahl $r \in \{1, \dots, n\}$ genau α_r-mal vorkommt. Also ist

$$g^{(k)}(t) = \sum_{|\alpha| = k} \frac{k!}{\alpha!} D^\alpha f(x_0 + th) h^\alpha$$

3. Schritt: Nach dem Satz von Taylor aus Analysis I existiert ein $\theta \in]0,1[$ mit

$$g(1) = \sum_{k=0}^{m} \frac{g^{(k)}(0)}{k!} + \frac{g^{(m+1)}(\theta)}{(m+1)!}$$

also

$$\begin{aligned} f(x_0+h) &= \sum_{k=0}^{m} \sum_{|\alpha|=k} \frac{D^\alpha f(x_0)}{\alpha!} h^\alpha + \sum_{|\alpha|=m+1} \frac{D^\alpha f(x_0+\theta h) h^\alpha}{\alpha!} \\ &= \sum_{|\alpha|\leq m} \frac{D^\alpha f(x_0)}{\alpha!} h^\alpha + \sum_{|\alpha|=m+1} \frac{D^\alpha f(x_0+\theta h) h^\alpha}{\alpha!} \end{aligned}$$

□

7.3.10 Mehrdimensionaler Satz von Taylor mit Integralrestglied

Ganz ähnlich beweist man nun

Satz 7.3.4. Seien $f \in C^m(U)$, U offene Teilmenge des $\mathbb{R}^n$, $x_0, x_0+h \in U$, sodass $[[x_0, x_0+h]] \subseteq U$. Dann gilt

$$f(x_0+h) = \sum_{|\alpha|\leq m} \frac{D^\alpha f(x_0) h^\alpha}{\alpha!} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \sum_{|\alpha|=m} \left(D^\alpha f(x_0+h) - D^\alpha f(x_0)\right) h^\alpha \,\mathrm{d}t$$

Beweis. Der 1. und 2. Schritt verlaufen wie im Beweis zuvor nur mit m an Stelle von $m+1$. Im 3. Schritt wendet man dann den Satz von Taylor mit Integraldarstellung auf g an:

$$g(1) = \sum_{k=0}^{m} \frac{g^{(k)}(0)}{k!} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \left(g^{(m)}(t) - g^{(m)}(0)\right) \mathrm{d}t$$

Die Formel für $g^{(k)}$ eingesetzt ergibt die Behauptung. □

Korollar 7.3.5. Seien $f \in C^m(U)$, U offene Teilmenge des $\mathbb{R}^n$, $x_0 \in U$, $\delta > 0$ mit $B_\delta(x_0) \subseteq U$. Dann gilt für alle $h \in \mathbb{R}^n$, $\|h\| < \delta$

$$f(x_0+h) = \sum_{|\alpha|\leq m} \frac{D^\alpha f(x_0)}{\alpha!} h^\alpha + r(h)$$

mit einem Rest $r(h)$ für den gilt

$$\lim_{h\to 0} \frac{r(h)}{|h|} = 0 = r(0)$$

Beweis. Nach dem Satz von Taylor, Satz 7.3.3, existiert zu jedem $h \in \mathbb{R}^n$, $\|h\| < \delta$, ein $\theta \in]0,1[$ mit

$$\begin{aligned} f(x_0+h) &= \sum_{|\alpha|\leq m-1} \frac{D^\alpha f(x_0) h^\alpha}{\alpha!} + \sum_{|\alpha|=m} \frac{D^\alpha f(x_0+\theta h) h^\alpha}{\alpha!} \\ &= \sum_{|\alpha|\leq m} \frac{D^\alpha f(x_0) h^\alpha}{\alpha!} + \underbrace{\sum_{|\alpha|=m} \frac{\left(D^\alpha f(x_0+\theta h) - D^\alpha f(x_0)\right) h^\alpha}{\alpha!}}_{=:r(h)} \end{aligned}$$

Für $|\alpha| = m$ ist

$$\frac{|h^\alpha|}{\|h\|^m} = \frac{|h_1^{\alpha_1} \cdot \ldots \cdot h_n^{\alpha_n}|}{\|h\|^{\alpha_1} \cdot \ldots \cdot \|h\|^{\alpha_n}}$$

Da alle Normen auf $\mathbb{R}^n$ äquivalent sind, kann man von $\|\cdot\| = \|\cdot\|_\infty$ ausgehen, also ist

$$\|h\|^{\alpha_i} = \|h\|_2^{\alpha_i} \geq |h_i|^{\alpha_i} \quad \forall\, i = 1, \ldots, n$$

Zusammen ergibt das: $\frac{|h|^\alpha}{\|h\|^m} \leq 1$. Es folgt

$$\begin{aligned}\frac{|r(h)|}{\|h\|^m} &\leq \sum_{|\alpha|=m} \frac{|D^\alpha f(x_0 + \theta h) - D^\alpha f(x_0)|}{\alpha!} \frac{|h^\alpha|}{\|h\|^m} \\ &\leq \sum_{|\alpha|=m} \frac{|D^\alpha f(x_0 + \theta h) - D^\alpha f(x_0)|}{\alpha!} \overset{h\to 0}{\longrightarrow} 0\end{aligned}$$

weil $f \in C^m(U)$. □

Spezialfälle:

1. $m = 0$: $f(x_0 + h) = f(x_0) + r(h)$ mit $r(h) \to 0 = r(0)$, daraus folgt gerade die Stetigkeit von f in x_0.
2. $m = 1$: $f(x_0 + h) = f(x_0) + \nabla f(x_0)h + r(h)$ mit $\frac{r(h)}{\|h\|} \to 0$; das zeichnet f als in x_0 differenzierbar aus.
3. $m = 2$: $f(x_0+h) = f(x_0)+\nabla f(x_0)h+\frac{1}{2}\mathrm{Hess}f(x_0)h\cdot h+r(h)$ mit $\frac{r(h)}{\|h\|^2} \to 0$.

   Es gibt zwei Typen von α-Tupeln mit $|\alpha| = 2$:

   - $\alpha = (2, 0, \ldots, 0), (0, 2, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, 2)$, d.h. $\alpha = 2e_i$ für ein $i = 1, \ldots, n$ und $\alpha! = 2$ und
   - $\alpha = (1, 1, 0, \ldots, 0), (1, 0, 1, 0, \ldots, 0), \ldots$, d.h. $\alpha = e_i + e_j$ für $1 \leq i < j \leq n$, also $\alpha! = 1$.

7.3.11 Allgemeine Taylor-Formel

Bemerkung. Für $f \in C^{m+1}(U)$, U offene Teilmenge des $\mathbb{R}^n$, $[[x_0, x_0 + h]] \subseteq U$, kann die Taylor-Formel auch äquivalent geschrieben werden als

$$f(x_0 + h) = \sum_{k=0}^{m} \frac{f^{(k)}(x_0) \overbrace{(h, \ldots, h)}^{k\text{-Tupel}}}{k!} + \frac{f^{(m+1)}(x_0 + \theta h) \overbrace{(h, \ldots, h)}^{m+1\text{-Tupel}}}{(m + 1)!}$$

oder alternativ für $f \in C^m(U)$ mit dem Integralrestglied

$$f(x_0 + h) = \sum + \int_0^1 \frac{(1 - t)^{m-1}}{(m - 1)!} \left(f^{(m)}(x_0 + th) - f^{(m)}(x_0)\right) \overbrace{(h, \ldots, h)}^{m\text{-Tupel}} \mathrm{d}t$$

denn es gilt

$$\begin{aligned}\frac{f^{(k)}(x_0)(h, \ldots, h)}{k!} &= \frac{\left(\left(\ldots\left(f^{(k)}(x_0)h\right) \ldots h\right) h\right)}{k!} = \frac{D_h \ldots D_h f(x_0)}{k!} \\ &\overset{!}{=} \sum_{|\alpha|=k} \frac{D^\alpha f(x_0)h^\alpha}{\alpha!}\end{aligned}$$

Beweis. der letzten Gleichung durch Induktion

$k = 1$:

$$\frac{f^{(1)}(x_0)h}{1!} = D_h f(x_0) = \nabla f(x_0)h = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x_0)h_i = \sum_{|\alpha|=1} D^\alpha f(x_0)h^\alpha$$

$k \to k+1$:

$$\begin{aligned}
\frac{f^{(k+1)}(x_0)(h,\dots,h)}{(k+1)!} &= D_h\left[D_h \dots D_h f(x_0)\right] \cdot \frac{1}{(k+1)\cdot k!} \\
&\overset{\text{IV}}{=} \frac{1}{k+1} D_h \left(\sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(x_0) h^\alpha \right) \\
&= \frac{1}{k+1} \nabla \left(\sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(\cdot) h^\alpha \right) \Bigg|_{x=x_0} \cdot h \\
&= \frac{1}{k+1} \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(x_0) h^\alpha \right) \cdot h_i
\end{aligned}$$

Wie beim Beweis der Taylor-Formel schon bemerkt, kann man die innere Summe unter Anpassung des Vorfaktors wie folgt auflösen

$$\begin{aligned}
\dots &= \frac{1}{k+1} \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\sum_{i_1,\dots,i_k=1}^{n} \frac{\frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x_0) h_{i_1} \cdot \dots \cdot h_{i_k}}{k!} \right) h_i \\
&= \sum_{i_1,\dots,i_k,i_{k+1}=1}^{n} \frac{\frac{\partial}{\partial x_{i_{k+1}}} \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f(x_0) h_{i_1} \cdot \dots \cdot h_{i_k} \cdot h_{i_{k+1}}}{(k+1)!} \\
&= \sum_{|\alpha|=k+1} \frac{D^\alpha f(x_0) h^\alpha}{\alpha!}
\end{aligned}$$

□

In der obigen Form macht die Taylor-Formel Sinn für allgemeine vektorwertige Funktionen in Banachräumen. Ohne Beweis folgender

Satz 7.3.6. Sei $U \subseteq X$, $f \in C^{m+1}(U,Y)$. Dann gilt für alle $x_0, x_0+h \in U$ mit $[[x_0, x_0+h]] \subseteq U$

$$\exists\, \theta \in [0,1] \ : \ f(x_0+h) = \sum_{k=0}^{m} \frac{f^{(k)}(x_0)(h,\dots,h)}{k!} + \frac{f^{(m+1)}(x_0+\theta h)(h,\dots,h)}{(m+1)!}$$

Bemerkung. Eine analoge Formel gilt mit der entsprechenden Integraldarstellung des Restglieds für $f \in C^m(U,Y)$.

7.4 Lokale Extrema reellwertiger Funktionen

Sei im folgenden Abschnitt $U \subseteq \mathbb{R}^n$ offen und stets $f : U \to \mathbb{R}$.

7.4.1 Lokale Extrema

Definition 7.4.1. f hat in einen Punkt $x_0 \in U$ ein **lokales Maximum** (respektive **lokales Minimum**), wenn es eine Umgebung V von x_0 gibt, sodass

$$f(x_0) \geq f(x) \quad \forall\, x \in V \qquad \text{(respektive } f(x_0) \leq f(x) \quad \forall\, x \in V)$$

Ist sogar

$$f(x_0) > f(x) \quad \forall\, x \in V \setminus \{x_0\} \qquad \text{(respektive } f(x_0) < f(x) \quad \forall\, x \in V \setminus \{x_0\}\,)$$

so hat f in x_0 ein **isoliertes (strenges/striktes) lokales Maximum** (respektive **Minimum**).

f hat in x_0 ein **(isoliertes) lokales Extremum**, wenn f in x_0 ein (isoliertes) lokales Maximum oder ein (isoliertes) lokales Minimum besitzt.

7.4.2 Notwendiges Kriterium

Satz und Definition 7.4.1.1. Sei f partiell differenzierbar. Wenn f in $x_0 \in U$ ein lokales Extremum hat, dann ist notwendig $\nabla f(x_0) = 0$, d.h. $\frac{\partial}{\partial x_i} f(x_0) = 0$ für jedes $i = 1, \ldots, n$.

Ein Punkt $x_0 \in U$ mit $\nabla f(x_0) = 0$ heißt **kritischer Punkt** von f.

Beweis. O.B.d.A. besitze f in x_0 ein lokales Maximum, d.h. es existiert ein $\epsilon > 0$, sodass

$$f(x_0) \geq f(x) \quad \forall\, x \in B_\epsilon(x_0)$$

Betrachte nun, für $i = 1, \ldots, n$, die Funktionen

$$]-\epsilon, \epsilon[\; \ni t \mapsto g_i(t) = f(x_0 + te_i)$$

Nach Voraussetzung besitzt g_i in $t = 0$ ein lokales Maximum und ist differenzierbar, da f partiell differenzierbar ist. Aus der Analysis I ist bekannt, dass dann folgt $g_i'(0) = 0$. D.h. aber gerade

$$0 = g_i'(0) = \lim_{t \to 0} \frac{g_i(t) - g_i(0)}{t} = \lim_{t \to 0} \frac{f(x_0 + te_i) - f(x_0)}{t} = \frac{\partial f}{\partial x_i}(x_0)$$

für jedes $i = 1, \ldots, n$, also $\nabla f(x_0) = 0$. $\square$

7.4.3 Erinnerung: Hinreichendes Kriterium im Eindimensionalen

Für 2-mal differenzierbare Funktionen f einer reellen Variable ist ein hinreichendes Kriterium für das Vorliegen eines Extremums in x_0: $f''(x_0) \neq 0$. Genauer: ist $f''(x_0) > 0$, so liegt ein lokales Minimum vor, ist $f''(x_0) < 0$, so liegt ein lokales Maximum vor.

Für $f \in C^2(U)$, wobei nun $U \subseteq \mathbb{R}^n$, erhält man ein hinreichendes Kriterium für das Vorliegen eines Extremums mit Hilfe der Hesse-Matrix.

7.4.4 Definitheit symmetrischer Matrizen

Definition 7.4.2. Sei A eine reelle symmetrische $n \times n$-Matrix. Dann heißt A

- **positiv definit**, falls $\langle A\xi, \xi \rangle > 0 \quad \forall\, \xi \in \mathbb{R}^n \setminus \{0\}$,

- **positiv semidefinit**, falls $\langle A\xi, \xi\rangle \geq 0 \quad \forall \xi \in \mathbb{R}^n$,
- **negativ definit**, falls $-A$ positiv definit, d.h $\langle A\xi, \xi\rangle < 0 \quad \forall \xi \in \mathbb{R}^n \setminus \{0\}$,
- **negativ semidefinit**, falls $-A$ positiv semidefinit und
- **indefinit**, falls A weder positiv definit noch negativ definit ist, d.h. wenn $\xi, \eta \in \mathbb{R}^n$ existieren mit $\langle A\xi, \xi\rangle > 0$ und $\langle A\eta, \eta\rangle < 0$.

Aus der linearen Algebra ist bekann, dass eine reelle symmetrische $n \times n$-Matrix n reelle Eigenwert $\lambda_0, \dots, \lambda_n$ (gezählt nach Vielfachheit) und eine Orthonormalbasis $v_1, \dots, v_n \in \mathbb{R}^n$ bestehend aus Eigenvektoren besitzt:

$$Av_i = \lambda_i v_i \quad \forall\, i = 1, \dots, n$$

Somit gilt für $v = \sum\limits_{i=1}^{n} \mu_i v_i$ mit $\mu_i \in \mathbb{R}$:

$$\begin{aligned}
\langle Av, v\rangle = Av \cdot v &= \left(A \sum_{i=1}^{n} \mu_i v_i\right) \cdot \sum_{i=1}^{n} \mu_i v_i \\
&= \left(\sum_{i=1}^{n} \mu_i A v_i\right) \cdot \sum_{i=1}^{n} \mu_i v_i \\
&= \left(\sum_{i=1}^{n} \mu_i \lambda_i v_i\right) \cdot \sum_{i=1}^{n} \mu_i v_i \\
&\overset{\text{ONB}}{=} \sum_{i=1}^{n} \mu_i^2 \lambda_i
\end{aligned}$$

Also gilt

Satz.

- A ist positiv (respektive negativ) definit $\iff$ alle Eigenwerte von A sind positiv (respektive negativ).
- A ist positiv (respektive negativ) semidefinit $\iff$ alle Eigenwerte von A sind ≥ 0 (bzw. ≤ 0).
- A ist indefinit $\iff$ A besitzt sowohl positive als auch negative Eigenwerte.

Weiter ist aus der linearen Algebra das sogenannte Hauptminorenkriterium bekannt

Satz. Für jedes $k = 1, \dots, n$ heißt

$$A_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

Hauptminor von A. Es gilt

- A ist positiv definit $\iff$ $\det A_k > 0$ für jedes $k = 1, \dots, n$.
- A ist negativ definit $\iff$ $(-1)^k \det A_k > 0$ für jedes $k = 1, \dots, n$.

7.4.5 Hinreichendes Kriterium

Satz 7.4.3. Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar, $x_0 \in U$ sei ein kritischer Punkt. Dann gilt

i) Ist $\mathrm{Hess} f(x_0)$ positiv definit, dann besitzt f in x_0 ein isoliertes lokales Minimum.

ii) Ist $\mathrm{Hess} f(x_0)$ negativ definit, dann besitzt f in x_0 ein isoliertes lokales Maximum.

iii) Ist $\mathrm{Hess} f(x_0)$ indefinit, so besitzt f in x_0 kein Extremum.

Beweis.

i) Sei $A := \mathrm{Hess} f(x_0)$. Nach Voraussetzung ist $A\xi \cdot \xi > 0$ für alle $\xi \in \mathbb{R}^n \setminus \{0\}$. Insbesondere folgt daraus: Die stetige Funktion

$$\{\xi \in \mathbb{R}^n \mid \|\xi\|_2 = 1\} \ni \xi \mapsto A\xi \cdot \xi$$

nimmt auf ihrem Definitionsbereich, der kompakten Einheitssphäre, ein positives Minimum an. D.h.

$$\min_{\xi \in \overline{B_1(0)}} A\xi \cdot \xi =: \alpha > 0$$

Nach der Taylor-Formel für $m = 2$ für f gilt

$$f(x_0 + h) = f(x_0) + \nabla f(x_0) \cdot h + \frac{1}{2} \mathrm{Hess} f(x_0) h \cdot h + r(h)$$

für $h \in \mathbb{R}^n$, $\|h\|_2$ klein genug, wobei $\frac{|r(h)|}{\|h\|_2^2} \to 0$ für $h \to 0$. Somit existiert ein $\delta > 0$ mit

$$\frac{|r(h)|}{\|h\|_2^2} < \frac{\alpha}{4} \quad \forall\, h \in \mathbb{R}^n,\ \|h\|_2 < \delta$$

d.h. $|r(h)| < \frac{\alpha}{4} \|h\|_2^2$. Es folgt

$$\begin{aligned} f(x_0 + h) &= f(x_0) + \frac{1}{2} A h \cdot h + r(h) \\ &= f(x_0) + \frac{1}{2} \|h\|_2^2 A \frac{h}{\|h\|_2} \cdot \frac{h}{\|h\|_2} + r(h) \\ &> f(x_0) + \frac{1}{2} \|h\|_2^2 \alpha + r(h) \\ &\geq f(x_0) + \frac{\alpha}{4} \|h\|_2^2 \end{aligned}$$

ii) folgt aus i) durch den Übergang zu $-f$.

iii) Wenn $H_f(x_0)$ indefinit ist, dann existieren $\xi, \eta \in \mathbb{R}^n$ mit

$$\alpha := H_f(x_0)\xi \cdot \xi > 0,\ \beta := H_f(x_0)\eta \cdot \eta < 0$$

Sei $t \in \mathbb{R}$, $|t|$ genügend klein, sodass $x_0 + t\xi \in U$, dann gilt nach Taylor-Formel

$$f(x_0 + t\xi) = f(x_0) + t \cdot \underbrace{\nabla f(x_0)}_{=0} \cdot \xi + \frac{t^2}{2} H_f(x_0)\xi \cdot \xi + r(t\xi)$$

mit $\frac{r(t\xi)}{\|t\xi\|_2^2} \to 0$, wenn $t \to 0$, d.h. $\frac{r(t\xi)}{t^2} \to 0$, also existiert ein $\delta > 0$

$$|r(t\xi)| < \frac{\alpha}{2}t^2 \quad \forall\, t \in \mathbb{R},\ |t| < \delta$$

womit $r(t\xi) > -\frac{\alpha}{2}t^2 = \frac{t^2}{2}H_f(x_0)\xi \cdot \xi$, also

$$f(x_0 + h) = f(x_0) + \frac{t^2}{2}H_f(x_0)\xi \cdot \xi + r(t\xi) > f(x_0) \quad \forall\, t \in \mathbb{R},\ |t| < \delta$$

gilt, d.h. f kann in x_0 kein lokales Maximum besitzen. Eine analoge Argumentation für η liefert: f kann in x_0 kein lokales Minimum besitzen.

$\square$

7.5 Implizite Funktionen

7.5.1 Motivation

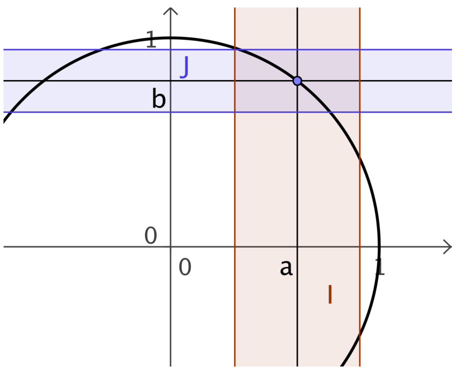
Gegeben sei $F : U \subseteq X \times Y \to Z$, $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y), (Z, \|\cdot\|_Z)$ Banachräume, $(x_0, y_0) \in U$ mit $F(x_0, y_0) = 0$.

Unter welchen Voraussetzungen an F hat die Gleichung $F(x, y) = 0$ für jedes x in einer Umgebung von x_0 genau eine Lösung $y \in Y$? Existiert eine solche Umgebung V, dann gibt es eine eindeutig bestimmte Funktion $g : V \to Y$ mit $F(x, g(x)) = 0$ für alle $x \in V$. Man sagt dann: $F(x, y) = 0$ **lässt sich nach** y **auflösen** und die Funktion g wird durch $F(x, g(x) = 0 \quad \forall\, x \in V$ **implizit** definiert. g heißt dann auch **Auflösung** von $F(x, y) = 0$ nach y.

7.5.2 Beispiel

Betrachte

$$\begin{aligned} f :\ &\mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ &(x, y) \mapsto x^2 + y^2 - 1 \end{aligned}$$



Ist $(a, b) \in \mathbb{R}^2$ mit $b > 0$ (respektive $b < 0$) und $f(a, b) = 0$, dann existieren offene Intervalle I, J mit $a \in I, b \in J$, sodass für jedes $x \in I$ genau ein $y \in J$ existiert mit $f(x, y) = 0$.

Die Zuordnung $I \ni x \mapsto y \in J$ definiert eindeutig eine Abbildung $g : I \to J$ mit der Eigenschaft

$$f(x, g(x)) = 0 \quad \forall\, x \in I$$

Bemerkung.

1. In diesem Beispiel kann g explizit angegeben werden: Ist $b > 0$, so ist $g(x) = \sqrt{1-x^2}$, ist $b < 0$, so ist $g(x) = -\sqrt{1-x^2}$ für alle $x \in I$.
2. Im Allgemeinen wird es nicht möglich sein, die Auflösung g explizit anzugeben.
3. In keiner der Umgebungen von $(1, 0)$ oder $(-1, 0)$ lässt sich die Gleichung $f(x, y) = 0$ nach y auflösen. Denn
$$\frac{\partial}{\partial y} f(x,y) = 2y \begin{cases} \neq 0, & y \neq 0 \\ = 0, & y = 0 \end{cases}$$
4. Damit y eindeutig bestimmt ist, muss im Beispiel J geeignet gewählt werden, ansonsten erhält man zu $x \in I$ eventuell 2 Lösungen: $\pm\sqrt{1-x^2}$.

7.5.3 Satz über implizite Funktionen

Im Folgenden versehen wir den Produktraum $X \times Y$ mit der Norm

$$\|(x,y)\| := \|x\|_X + \|y\|_Y \quad \forall\,(x,y) \in X \times Y$$

$(X \times Y), \|\cdot\|)$ ist dann selbst ein Banachraum. Weiterhin vereinbaren wir

Wenn $U \subseteq X \times Y$ offen, $F : U \to Z$ in $(x_0, y_0) \in U$ differenzierbar ist, dann definieren wir

$$D_X F(x_0,y_0) \in \mathcal{L}(X,Z),\; D_X F(x_0,y_0)h := F'(x_0,y_0)(h,0) \quad \forall\, h \in X$$
$$D_Y F(x_0,y_0) \in \mathcal{L}(Y,Z),\; D_Y F(x_0,y_0)h := F'(x_0,y_0)(0,h) \quad \forall\, h \in Y$$

Satz 7.5.1. Sei $f : U \subseteq X \times Y \to Z$ stetig differenzierbar, $(x_0, y_0) \in U$ mit $f(x_0, y_0) = 0$. Falls $D_Y f(x_0, y_0)$ invertierbar ist, dann kann man $F(x, y) \neq 0$ lokal eindeutig in einer Umgebung von (x_0, y_0) nach y auflösen. Genauer gilt

i) Es existieren $\epsilon, \delta > 0$ mit $B_\delta(x_0) \times B_\epsilon(y_0) \subseteq U$ und so, dass für alle $x \in B_\delta(x_0)$ die Gleichung $F(x, y) = 0$ genau eine Lösung $y \in B_\epsilon(y_0)$ besitzt.

ii) Die durch i) eindeutig bestimmte Funktion $g : B_\delta(x_0) \to B_\epsilon(y_0)$ mit
$$F(x, g(x)) = 0 \quad \forall\, x \in B_\delta(x_0)$$
ist stetig differenzierbar.

iii) Für alle $x \in B_\delta(x_0)$ ist $D_Y F(x, g(x))$ invertierbar und es gilt
$$g'(x) = -\left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))$$

Vor dem großen Satz und seinem Beweis das folgende

Lemma 7.5.2. Wenn $C \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^q)$ mit $\|C\| < 1$, dann ist die Abbildung $I - C$ invertierbar.

Beweis. Es ist $\|C^n\| \leq \|C\|^n \quad \forall\, n \in \mathbb{N}$. Daher ist

$$S_n := \sum_{k=0}^{n} C^k \quad \forall\, n \in \mathbb{N},$$

die sogenannte **Neumann'sche Reihe**, eine Cauchy-Folge in $\mathcal{L}(\mathbb{R}^q, \mathbb{R}^q)$, denn da die geometrische Reihe $\sum \|C\|^k$ wegen $\|C\| < 1$ konvergiert, ist

$$\|S_n - S_m\| = \left\| \sum_{k=m+1}^{n} C^k \right\| \leq \sum_{k=m+1}^{n} \|C\|^k$$

für alle $n < m$. Da $\mathcal{L}(\mathbb{R}^q, \mathbb{R}^)$ vollständig ist, konvergiert $(S_n)_{n\in\mathbb{N}}$ gegen ein $S \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^q)$. Es ist aber

$$S_n(I - C) = \sum_{k=0}^{n} C^k - \sum_{k=1}^{n+1} C^k = I - C^{n+1} = (I - C)S_n$$

also für $n \to \infty$

$$S(I - C) = I = (I - C)S$$

wobei die Konvergenz des mittleren Terms gilt, da $\left\|C^{n+1}\right\| \leq \|C\|^{n+1} \to 0$. Insgesamt ist also gezeigt, dass $I - C$ invertierbar ist mit $(I - C)^{-1} = S$. □

Beweis. des Satzes nur für endlich-dimensionale Vektorräume, d.h. o.B.d.A. $X = \mathbb{R}^p, Y = \mathbb{R}^q$.

i) Setze $D := D_Y F(x_0, y_0)$; nach Voraussetzung und weil X, Y endlich-dimensional sind, ist $D^{-1} \in \mathcal{L}(Z, Y)$. Bemerke, dass für $(x, y) \in U$ die Äquivalenz $F(x, y) = 0 \Leftrightarrow y - D^{-1}F(x, y) = y$ gilt. Letztere Gleichung ist eine Fixpunktgleichung.

   Da F und $D_Y F$ nach Voraussetzung auf U stetig sind, existiert eine δ-Umgebung $B_\delta(x_0)$ von x_0 und eine ϵ-Umgebung $B_\epsilon(y_0)$ von y_0 mit $B_\delta(x_0) \times B_\epsilon(y_0) \subseteq U$, sodass

$$\left\|I - D^{-1} D_Y F(x, y)\right\| \leq \frac{1}{2} \qquad \forall\, (x, y) \in B_\delta(x_0) \times B_\epsilon(y_0) \tag{1}$$

$$\left\|D^{-1} F(x, y_0)\right\| \leq \frac{\epsilon}{4} \qquad \forall\, x \in B_\delta(x_0) \tag{2}$$

   Sei $x \in B_\delta(x_0)$ fest. Betrachte die Abbildung

$$\begin{aligned} \Phi :\ B_\epsilon(y_0) &\to \mathbb{R}^q \\ y &\mapsto y - D^{-1}F(x, y) \end{aligned}$$

   Unter Zuhilfnahme der Voraussetzungen und der Kettenregel erkennt man, dass Φ differenzierbar ist mit der Ableitung

$$\Phi'(y) = I - \left[(D^{-1})'(F(x, y))\right] D_Y F(x, y) = I - D^{-1} D_Y F(x, y)$$

   (denn es ist $(D^{-1})'(z) = D^{-1}$ für alle $z \in Z$, siehe 7.1.6)

und nach Schrankensatz und (1) gilt für alle $y, z \in B_\epsilon(y_0)$

$$\begin{aligned}\|\Phi(y)-\Phi(z)\|_2 &\le \sup_{t\in]0,1[} \|\Phi'(y+t(z-y))\|_{\mathcal{L}(Y,Y)} \|y-z\|_2 \\ &= \sup_{t\in]0,1[} \left\|I - D^{-1}D_Y F(x, y+t(z-y))\right\|_{\mathcal{L}(Y,Y)} \|y-z\|_2 \\ &\overset{(1)}{\le} \frac{1}{2}\|y-z\|_2 \quad \forall\, z,y \in B_\epsilon(y_0)\end{aligned}$$

d.h. Φ ist eine Kontraktion. Zudem ist

$$\begin{aligned}\|\Phi(y)-y_0\|_2 &\le \|\Phi(y)-\Phi(y_0)\| + \|\Phi(y_0)-y_0\| \\ &= \|\Phi(y)-\Phi(y_0)\| + \left\|y_0 - D^{-1}F(x,y_0) - y_0\right\| \\ &\le \frac{1}{2}\underbrace{\|y-y_0\|_2}_{<\epsilon} + \underbrace{\left\|D^{-1}F(x,y_0)\right\|}_{\le\frac{\epsilon}{4} \text{ nach } (2)} < \epsilon\end{aligned}$$

Somit bildet Φ $B_\epsilon(y_0)$ wieder in $B_\epsilon(y_0)$ ab.

Mit dem Banach'schen Fixpunktsatz folgt dann die Existenz und Eindeutigkeit eines Fixpunkts $\hat{y} \in B_\epsilon(y_0)$ von Φ. Damit ist dann gezeigt, dass für alle $x \in B_\delta(x_0)$ die Gleichung $F(x,y) = 0$ genau eine Lösung $y \in B_\epsilon(y_0)$ besitzt.

ii) Nach i) ist nun in eindeutiger Weise eine Funktion $g : B_\delta(x_0) \to B_\epsilon(y_0)$ mit $F(x, g(x)) = 0 \quad \forall\, x \in B_\delta$ definiert. Wir zeigen zunächst, dass g stetig ist. In der Tat erweist sich die obige Funktion g als Fixpunkt der Abbildung

$$\begin{aligned}T :\ & M \to M \\ & u \mapsto u - D^{-1}F(\cdot, u(\cdot))\end{aligned}$$

wobei

$$M := \left\{u : B_\delta(x_0) \to \mathbb{R}^q \mid u \text{ stetig}, u(x_0) = y_0, u(B_\delta(x_0)) \subseteq B_{\frac{\epsilon}{2}}(y_0)\right\}$$

Es gilt: M ist eine abgeschlossene, nicht-leere Teilmenge des Banachraums $(C_b(B_\delta(x_0), \mathbb{R}^q), \|\cdot\|_\infty))$ (d.h. des Vektorraums der stetigen, beschränkten Abbildungen $B_\delta(x_0) \to \mathbb{R}^q$ ausgestattet mit der Supremumsnorm).

T ist wohldefiniert: D^{-1} und F sind nach Voraussetzung stetig ist; außerdem ist für $u \in M$ und für alle $x \in B_\delta(x_0)$

$$\begin{aligned}\|(Tu)(x) - y_0\|_2 &= \|\Phi(u(x)) - y_0\|_2 \\ &\le \|\Phi(u(x)) - \Phi(y_0)\|_2 + \|\Phi(y_0) - y_0\|_2 \\ &\le \frac{1}{2}\|u(x) - y_0\|_2 + \left\|y_0 - D^{-1}F(x,y_0) - y_0\right\|_2 \\ &\le \frac{1}{2}\cdot\frac{\epsilon}{2} + \frac{\epsilon}{4} = \frac{\epsilon}{2}\end{aligned}$$

Also ist klar, dass $Tu \in C(B_\delta(x_0), \mathbb{R}^q)$ für alle $u \in M$. Weiter ist

$$(Tu)(x_0) = u(x_0) - D^{-1}F(x_0, u(x_0)) = y_0 - D^{-1}\underbrace{F(x_0,y_0)}_{=0} = y_0$$

Zusammen hat man also $Tu \in M$.

T ist kontrahierend mit Konstante $\frac{1}{2}$, da für alle $u, v \in M, x \in B_\delta(x_0)$ gilt

$$\begin{aligned}\|(Tu)(x) - (Tv)(x)\|_2 = \|\Phi(u(x)) - \Phi(v(x))\|_2 \le \frac{1}{2}\|u(x) - v(x)\|_2 \quad \Rightarrow \\ \|Tu - Tv\|_\infty \le \frac{1}{2}\|u - v\|_\infty \quad \forall\, u, v \in M\end{aligned}$$

Somit erfüllt T alle notwendigen Voraussetzungen, um den Banach'schen Fixpunktastz anzuwenden. T besitzt einen eindeutigen Fixpunkt $\tilde{g} \in M$. Auf Grund der Eindeutigkeit der Lösung $y \in B_\epsilon(y_0)$ der Gleichung

$$F(x, y) = 0$$

für festes $x \in B_\delta(x_0)$ folgt unmittelbar $\tilde{g} = g$, d.h. g ist stetig.

Um die fehlenden Punkte zu zeigen, beweisen wir zunächst, dass $D_Y F(x, y)$ in einer (eventuelle kleinen Umgebung) $B_{\delta_1}(x_0) \times B_{\epsilon_1}(y_0)$, wobei $0 < d_1 \le \delta$ und $0 < \epsilon_1 \le \epsilon$, invertierbar ist.

Da F auf $B_\delta(x_0) \times B_\epsilon(y_0)$ stetig differenzierbar ist, existieren zunächst $0 < \delta_2 \le \delta$ und $0 < \epsilon_2 \le \epsilon$ mit

$$\|D_Y F(x, y) - D_Y F(x_0, y_0)\| < \frac{1}{\left\|(D_Y F(x_0, y_0))^{-1}\right\|}$$

für alle $(x, y) \in B_{\delta_2}(x_0) \times B_{\epsilon_2}(y_0)$. Mit den Bezeichnungen

$$A := D_Y F(x, y),\ B := D_Y F(x_0, y_0)$$

für festes (x, y) gilt also

$$\left\|B^{-1}A - I\right\| = \left\|B^{-1}(A - B)\right\| \le \left\|B^{-1}\right\| \cdot \|A - B\| < 1$$

Nach Lemma 7.5.2 folgt nun, dass $B^{-1}A = I - (I - B^{-1}A)$ in $C(\mathbb{R}^q, \mathbb{R}^q)$ invertierbar ist. Damit ist auch $A = D_Y F(x, y)$ invertierbar. Da (x, y) beliebig war, ist also $D_Y F(x, y)$ invertierbar für alle $(x, y) \in B_{\delta_2}(x_0) \times B_{\epsilon_2}(y_0)$.

Wähle nun $\epsilon_1 := \epsilon_2$. Da die Auflösung g aus i) bzw. ii) stetig ist, existiert nun aber $0 < \delta_1 \le \delta_2$ mit $g(B_{\delta_1}(x_0)) \subseteq B_{\epsilon_1}(y_0)$. Dann gilt die Aussage i) mit ϵ_1 und δ_1 an Stelle von ϵ und δ; außerdem gilt: $D_Y F(x, y)$ ist invertierbar.

Im weiteren Verlauf arbeiten wir mit ϵ_1, δ_1, die wir aber der Einfachheit halber wieder mit ϵ, δ bezeichnen.

Es bleibt zu zeigen, dass g stetig differenzierbar ist, und dass für die Ableitung gilt

$$g'(x) = -\left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x)) \quad \forall\, x \in B_\delta(x_0)$$

Nach dem Satz von Taylor ist für $(x, y), (x + h, y + k) \in B_\delta(x_0) \times B_\epsilon(y_0)$

$$\begin{aligned}F(x + h, y + k) &= F(x, y) + F'(x, y)h + r(h, k) \\ &= F(x, y) + D_X F(x, y)h + D_Y F(x, y)k + r(h, k)\end{aligned}$$

wobei $\frac{r(h,k)}{\|(h,k)\|} = \frac{r(h,k)}{\|h\|_2+\|k\|_2} \to 0$ für $(h,k) \to (0,0)$. Somit folgt

$$\begin{aligned} F(x+h, g(x+h)) &= F(x, g(x)) + D_X F(x, g(x))h \\ &\quad + D_Y F(x, g(x)) \left(g(x+h) - g(x)\right) + r(h, g(x+h) - g(x)) \end{aligned}$$

für $x, x+h \in B_\delta(x_0)$, d.h.

$$0 = D_X F(x, g(x)) + D_Y F(x, g(x))(g(x+h) - g(x)) + r(h, g(x+h) - g(x))$$

mit

$$\frac{r(h, g(x+h) - g(x))}{\|h\|_2 + \|g(x+h) - g(x)\|_2} \xrightarrow{h \to 0} 0 \quad (*)$$

Aber $D_Y F(x, g(x))$ ist invertierbar für alle $x \in B_\delta(x_0)$, womit folgt

$$g(x+h) - g(x) - \left[- \left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))h\right] \tag{3}$$

$$= \underbrace{(D_Y F(x, g(x)))^{-1} r(h, g(x+h) - g(x))}_{=:\tilde{r}(h)} \tag{4}$$

Wenn es gelingt zu zeigen, dass $\frac{\tilde{r}(h)}{\|h\|_2} \xrightarrow{h \to 0} 0$ (**), dann folgt, dass g differenzierbar ist und

$$g'(x) = - \left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x)) \quad \forall\, x \in B_\delta(x_0)$$

Die stetige Differenzierbarkeit von g folgt unmittelbar aus der Stetigkeit der Ableitung, die als Komposition stetiger Abbildungen stetig ist.

Tatsächlich gelingt der Nachweis von (**) nicht direkt. Aber für $\tilde{\epsilon}$ existiert $\tilde{\delta} > 0$, sodass

$$\frac{1}{2\left\|\left(D_Y F(x, g(x))\right)^{-1}\right\|} > \tilde{\epsilon} > 0$$

Wegen (*) und wegen der Stetigkeit von g kann man $\tilde{\delta}$ so klein wählen, dass gilt:

$$\begin{aligned} &\|r(h, g(x+h) - g(x))\| \\ &\le \tilde{\epsilon} \left(\|h\|_2 + \|g(x+h) - g(x)\|_2\right) \\ &\le \tilde{\epsilon} \left(\|h\|_2 + \left\|g(x+h) - g(x) - \left(- \left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))h\right)\right\|_2 \right. \\ &\quad \left. + \left\|\left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))h\right\|_2\right) \quad \forall\, \|h\| < \tilde{\delta} \end{aligned}$$

woraus folgt

$$\begin{aligned} &\|r(h, g(x+h) - g(x))\| \\ &\le \tilde{\epsilon} \left(1 + \left\|\left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))\right\|_2\right) \|h\|_2 \\ &\quad + \frac{\left\|g(x+h) - g(x) - \left(- \left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))h\right)\right\|_2}{2\left\|\left(D_Y F(x, g(x))\right)^{-1}\right\|} \end{aligned}$$

Daraus folgt mit (4)

$$\left\|g(x+h) - g(x) - \left[-\left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))h\right]\right\|$$
$$\leq 2\left\|\left(D_Y F(x, g(x))\right)^{-1}\right\| \tilde{\epsilon}\left(1 + \left\|\left(D_Y F(x, g(x))\right)^{-1} D_X F(x, g(x))\right\|_2\right) \|h\|_2$$

für alle $\|h\| < \tilde{\delta}$. Damit ist (**) gezeigt.

$\square$

7.5.4 Umkehrsatz

Unmittelbar aus dem Satz über implizite Funktionen folgt der

Satz 7.5.3. Sei $f : U \subseteq X \to Y$ stetig differenzierbar, $x_0 \in U$ so, dass $f'(x_0)$ invertierbar ist. Dann existiert eine offene Menge $W \subseteq U$, $x_0 \in W$, V Umgebung von $y_0 = f(x_0)$ in Y, sodass $f|_W : W \to V$ bijektiv und $g := f|_W{}^{-1} : V \to W$ stetig differenzierbar ist mit

$$g'(y) = f'(g(y_0))^{-1} = \left(f'(x_0)\right)^{-1}$$

Beweis. Die Gleichung $f(x) - y = 0$ muss nach x aufgelöst werden. Betrachte deshalb

$$\begin{aligned} F : U \times Y &\to Y \\ (x, y) &\mapsto f(x) - y \end{aligned}$$

F ist stetig differenzierbar, da f stetig differenzierbar ist; es gilt $F(x_0, y_0) = 0$ und $D_X F(x_0, y_0) = f'(x_0)$ ist invertierbar. Nach dem Satz 7.5.1 existieren $\epsilon, \delta > 0$ und $g : B_\epsilon(y_0) \to B_\delta(x_0)$ stetig differenzierbar mit $g(y_0) = x_0$ und $F(g(y), y) = 0$ für alle $y \in B_\delta(y_0)$, d.h. aber gerade $f(g(y)) = y$ für alle diese y. Mit $V := B_\epsilon(y_0), W := B_\delta(x_0) \cap f^{-1}(B_\delta(y_0))$ folgt:

$f|_W : W \to V$ ist bijektiv und $g = \left(f|_W\right)^{-1}$. nach Satz 7.5.1 ist g auch stetig differenzierbar und es gilt

$$I = (f \circ g)'(y_0) = f'(g(y_0))g'(y_0) = f'(x_0)g'(y_0)$$

also $g'(y_0) = \left(f'(x_0)\right)^{-1}$. $\square$

7.6 Extrema mit Nebenbedingungen

7.6.1 Problemstellung

Gesucht sind Extrema bzw. Extremstellen einer Funktion $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ unter einer Nebenbedingung $g(x) = 0$, wobei $g : U \to \mathbb{R}^p$ stetig differenzierbar ist; d.h. gesucht sind die Extrema bzw. Extremstellen von $f|_M$, wobei $M = \{x \in U \mid g(x) = 0\}$.

Man sagt: f besitzt in $x_0 \in M$ ein lokales Maximum (respektive lokales Minimum) unter der **Nebenbedingung** $g(x) = 0$, wenn $f(x) \leq f(x_0) \quad \forall x \in M \cap V$ für eine gewisse Umgebung V von x_0, $V \subseteq U$ (respektive $f(x) \geq f(x_0) \quad \forall x \in M \cap V$).

Es sei im Folgenden stets $1 \leq p < n$ (d.h. es gibt weniger Nebenbedingungen als Variablen). Manchmal ist es möglich, die p Nebenbedingungen explizit nach p Variablen aufzulösen.

7.6.2 Beispiel

Gesucht sind die Extrema von

$$\begin{aligned} f:\ \mathbb{R}^3 &\to \mathbb{R} \\ (x,y,z) &\mapsto x^2+y^2+z^2 \end{aligned}$$

unter der Nebenbedingung $g(x,y,z)=0$, wobei

$$\begin{aligned} g:\ \mathbb{R}^3 &\to \mathbb{R} \\ (x,y,z) &\mapsto 4x+3y-z \end{aligned}$$

Dann sind äquivalent: $g(x,y,z)=0 \Leftrightarrow z=4x+3y$, und das Bestimmen der Extrema von $f|_M$ mit

$$M=\big\{(x,y,z)\in\mathbb{R}^3 \mid g(x,y,z)=0\big\}=\big\{(x,y,z)\in\mathbb{R}^3 \mid z=4x+3y\big\}$$

ist äquivalent zum Bestimmen der Extrema der Funktion

$$\varphi:\ \mathbb{R}^2 \ni (x,y) \to f(x,y,4x+4y)=x^2+y^2+(4x+3y)^2 \in \mathbb{R}$$

Die ist ein wohlbekanntes Problem, das mit den Ergebnissen der vorhergehenden Kapitel gelöst werden kann.

7.6.3 Lagrange'sche Multiplikatoren

Achtung. Aber: Im Allgemeinen ist eine solche explizite Auflösung der Nebenbedingungen nicht möglich!

In diesem Fall ist manchmal folgender Satz hilfreich

Satz 7.6.1. Seien $f:U\subseteq\mathbb{R}^n\to\mathbb{R}$, $g:U\to\mathbb{R}^p$ stetig differenzierbar, $1\le p<n$, $x_0\in U$. Die Jacobi-Matrix von g habe Höchstrang in x_0 (d.h. $\mathrm{rg}(Dg(x_0))=p$) und es sei $g(x_0)=0$. Hat dann $f|_M$, wobei $M=\{x\in U \mid g(x)=0\}$ in x_0 ein Extremum, so existieren $\lambda_1,\dots,\lambda_p\in\mathbb{R}$ (sogenannte **Lagrange'sche Multiplikatoren**), sodass gilt

$$Df(x_0)+\sum_{j=1}^{p}\lambda_j g_j'(x_0)=0$$

d.h. es gilt

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x_0)+\lambda\frac{\partial g_1}{\partial x_1}(x_0)+\ldots+\lambda_p\frac{\partial g_p}{\partial x_n}(x_0)&=0 \\ &\vdots \\ \frac{\partial f}{\partial x_n}(x_0)+\lambda\frac{\partial g_1}{\partial x_1}(x_0)+\ldots+\lambda_p\frac{\partial g_p}{\partial x_n}(x_0)&=0 n \text{ Gleichungen} \end{aligned}$$

7.6.4 Nutzen der Satzes über Lagrange'sche Multiplikatoren

Bevor wir den Satz beweisen, wollen wir zunächst überlegen, wie der Satz für das Auflösen von Extrema unter Nebenbedingungen genutzt werden kann. Das allgemeine Vorgehen ist das folgende

Betrachte das Gleichungssystem (*) aus dem Satz. Dann gilt: jeder Punkt $x \in M$, der „Anfang" einer Lösung $(x^1, \ldots, x^n, \lambda_1, \ldots, \lambda_p)$ von (*) ist und für den gilt, dass $\operatorname{rg}(g'(x)) = 0$, kommt als mögliche Extremstelle in Frage. Weitere mögliche Kandidaten sind Punkte $x \in M$ mit $\operatorname{rg}(g'(x)) < p$. Im Einzelnen muss dann nachgeprüft werden, ob ein Extremum vorliegt.

7.6.5 Beispiel

Gesucht sind die Extrema von

$$\begin{aligned} f :\ \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto x \cdot y \end{aligned}$$

auf der Einheitssphäre $\left\{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\right\}$, d.h. unter der Nebenbedingung $g(x,y) = 0$, wobei

$$\begin{aligned} g :\ \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto x^2+y^2-1 \end{aligned}$$

f und g sind stetig differenzierbar, $g'(x,y) = (2x, 2y)$ hat Höchstrang $(= 1)$ in jedem Punkt $(x,y) \neq (0,0)$, d.h. insbesondere in jedem Punkt $(x,y) \in M$. Nach Satz 7.6.1 kommen deshalb als Kandidaten für (lokale) Extremstellen nur solche (x,y) in Frage, die „Anfang" einer Lösung (x,y,λ) des Gleichungssystems

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) + \lambda \frac{\partial g}{\partial x}(x,y) &= 0 \qquad &\Leftrightarrow \qquad y + \lambda 2x &= 0 \qquad &\text{(I)} \\ \frac{\partial f}{\partial y}(x,y) + \lambda \frac{\partial g}{\partial y}(x,y) &= 0 \qquad &\Leftrightarrow \qquad x + \lambda 2y &= 0 \qquad &\text{(II)} \\ g(x,y) &= 0 \qquad &\Leftrightarrow \qquad x^2+y^2 &= 1 \qquad &\text{(III)} \end{aligned}$$

sind. Löse nun dieses Gleichungssystem:

$$\begin{aligned} \text{(I)} &\Leftrightarrow y = -\lambda 2x \\ \text{(II)} &\overset{\text{(I)}}{\Rightarrow} x + \lambda 2 \cdot (-\lambda 2x) = 0 \Leftrightarrow (1-4\lambda^2)x = 0 \end{aligned}$$

also $x = 0$ oder $\lambda = \pm\frac{1}{2}$. Ersteres zurück in (I) eingesetzt ergibt $y = 0$, was einen Widerspruch zu (III), also $x^2+y^2 = 1$, ergibt. Setzt man die anderen Lösungen zurück in (I) ein, so erhält man

$$\begin{aligned} y &= \pm x \\ \text{(III)} \Rightarrow (x,y) &= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \end{aligned}$$

Dies sind die einzigen möglichen Kandidaten. Da

$$\begin{aligned} f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) &= \frac{1}{2} \\ f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) &= -\frac{1}{2} \end{aligned}$$

befinden sich an $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ (globale) Maxima und an $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ (globale) Minima.

7.6.6 Beweis des Satzes

Beweis. Nach eventuelle Umnummerierung der Variablen $x_1, \ldots, x_n$ darf man annehmen, dass

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x_0) & \ldots & \frac{\partial g_1}{\partial x_p}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1}(x_0) & \ldots & \frac{\partial g_p}{\partial x_p}(x_0) \end{pmatrix}$$

invertierbar ist. Schreibt man $x^* = (x_1, \ldots, x_p), y = (x_{p+1}, \ldots, x_n)$, dann ist $x_0 = (x_0^*, y_0)$ und man hat

$$\begin{aligned} g:\ \mathbb{R}^p \times \mathbb{R}^{n-p} &\to \mathbb{R}^p \\ (x^*, y) &\mapsto g(x^*, y) \end{aligned}$$

Dann ist die obige Matrix $= D_{X^*} g(x^*, y)$. Nach Satz 7.5.1 lässt sich dann die Gleichung $g(x) = g(x^*, y) = 0$ in x_0 lokal und stetig differenzierbar nach $x^* = (x_1, \ldots, x_p)$ auflösen:

$$\begin{aligned} x_1 &= h_1(x_{p+1}, \ldots, x_n) \\ &\vdots \\ x_p &= h_p(x_{p+1}, \ldots, x_n) \end{aligned}$$

wobei $h = (h_1, \ldots, h_p)$ die lokale Auflösungsfunktion ist. Also gilt $h(y_0) = x_0^*$ und in einer Umgebung von y_0

$$g(h(y), y) = 0$$

Nach Voraussetzung besitzt die Funktion $f|_M$ in x_0 ein Extremum. Dies impliziert, dass die in einer Umgebung von y_0 reellwertige, stetig differenzierbare Funktion $y \mapsto f(h(y), y)$ ein lokales Extremum in y_0 besitzt. Nach notwendigem Kriterium für das Vorliegen eins solchen folgt

$$0 = \nabla f\,\underbrace{(h(y_0), y_0)}_{=x_0} = \nabla f(h_1(x_{p+1}^0, \ldots, x_n^0), \ldots, h_p(x_{p+1}^0, \ldots, x_n^0), x_{p+1}^0, \ldots, x_n^0)$$

Nach Kettenregel folgt also

$$\begin{aligned} 0 &= \underbrace{(\nabla f)(x_0)}_{=\left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)\right)} \cdot \begin{pmatrix} \frac{\partial h_1}{\partial x_{p+1}}(y_0) & \ldots & \frac{\partial h_1}{\partial x_n}(y_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_p}{\partial x_{p+1}}(y_0) & \ldots & \frac{\partial h_p}{\partial x_n}(y_0) \\ 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{pmatrix} && (5) \\ &= D_{X^*} f(x_0) \cdot Dh(y_0) + D_Y f(x_0) && (6) \end{aligned}$$

Da $g(h(y), y) = 0$ in einer Umgebung von y_0, ist auch $Dg(h(y), y) = 0$ für alle y in dieser Umgebung, d.h. aber nach Kettenregel

$$\begin{aligned} D_{X^*}g(h(y_0), y_0) \cdot Dh(y_0) + D_Y g(h(y_0), y_0) \cdot I_{n-p} = 0 \qquad &\Leftrightarrow \\ \underbrace{D_{X^*}g(x_0)}_{\text{invertierbar}} Dh(y_0) + D_Y g(h(y_0, y_0) = 0 \Leftrightarrow& \\ Dh(y_0) = -\left(D_{X^*}g(x_0)\right)^{-1} D_Y g(h(y_0, y_0)& \end{aligned}$$

In (6) eingesetzt ergibt das

$$0 = -\underbrace{D_{X^*}f(x_0)}_{\in\mathbb{R}^{1\times p}} \underbrace{\left(D_{X^*}g(x_0)\right)^{-1}}_{\in\mathbb{R}^{p\times p}} D_Y g(h(y_0, y_0) + D_Y f(x_0)$$

Definiere $(\lambda_1, \ldots, \lambda_p) := -D_{X^*}f(x_0)\left(D_{X^*}g(x_0)\right)^{-1}$, so erhält man die $n-p$-Gleichungen

$$0 = \sum_{j=1}^{p} \lambda_j \cdot D_Y g_j(x_0) + D_Y f(x_0)$$

Es ist nach der Definition der $\lambda_1, \ldots, \lambda_n$

$$\begin{aligned} (\lambda_1, \ldots, \lambda_p) = -D_{X^*}f(x_0)\left(D_{X^*}g(x_0)\right)^{-1} \qquad &\Leftrightarrow \\ 0 = \sum_{j=1}^{p} \lambda_j \cdot D_{X^*}g_j(x_0) + D_{X^*}f(x_0)& \end{aligned}$$

Insgesamt also

$$0 = \sum_{j=1}^{p} \lambda_j \cdot Dg_j(x_0) + Df(x_0)$$

was zu zeigen war. $\square$

8 Differentialgleichungen

8.1 Gewöhnliche Differentialgleichungen erster Ordnung

8.1.1 Grundlegende Definitionen

Definition 8.1.1. Sei $(X, \|\cdot\|_X)$ ein Banachraum, $D \subseteq \mathbb{R} \times X$ offen, $g : D \to X$. Die Gleichung

$$y' = g(t, y)$$

heißt **gewöhnliche Differentialgleichung erster Ordnung (in expliziter Form)**. Ausgeschrieben bedeutet sie

$$y'(t) = g(t, y(t)) \quad \forall\, t$$

Eine auf einem Intervall $I \subseteq \mathbb{R}$ mit $\mathring{I} \neq \emptyset$ definierte Funktion $y : I \to X$ heißt **Lösung** der Differentialgleichung, wenn

$$(t, y(t)) \in D \quad \forall\, t \in I \quad \text{ und } \quad y'(t) = g(t, y(t)) \quad \forall\, t \in I$$

Hängt $g(t, y)$ nicht explizit von t ab (d.h. ist konstant in t), so kann g als Funktion von $D \subseteq X \to X$ aufgefasst werden und

$$y' = g(t)$$

heißt **autonome Differentialgleichung**. Entsprechend wird die allgemeine Form auch als **nicht-autonome Differentialgleichung** bezeichnet.

8.1.2 Beispiel: Populationsmodelle

1. Exponentielles Wachstumsmodell (Malthus, 1798)

   $y(t) :=$ Anzahl der Individuen in einer Population zum Zeitpunkt $t \geq 0$
   $y(0) = y_0$ Anzahl der Individuen zum Zeitpunkt $t = 0$ sei bekannt

   Gesucht ist: $y(t)$ für $t > 0$. Annahme

   $$\underbrace{y(t + \Delta t) - y(t)}_{\text{Zu-/Abnahme der Population}} \quad \overbrace{\approx}^{\text{proportional}} \quad \underbrace{\alpha}_{\text{Geburtenrate}} \; y(t)\Delta t - \overbrace{\beta}^{\text{Sterberate}} \; y(t)\Delta t$$

   Unter der Annahme, dass y in $t > 0$ differenzierbar ist, erhält man (als Approximation) mit $\Delta t \to 0$:

   $$y'(t) = \alpha y(t) - \beta y(t)$$

   d.h. $y' = (\alpha - \beta)y$. Dies ist eine Differentialgleichung mit der Anfangsbedingung $y(0) = y_0$.
   Die Differentialgleichung $y' = \gamma y$ besitzt die unendlich vielen Lösungen

   $$y(t) = c \cdot \mathrm{e}^{\gamma t} \quad \forall\, t \in \mathbb{R}$$

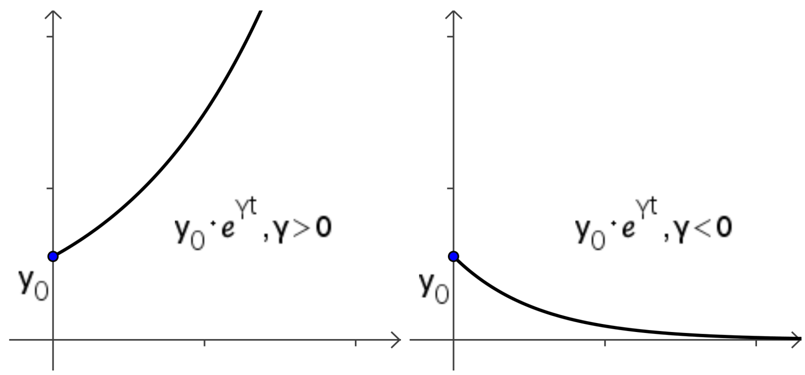
   wobei $c \in \mathbb{R}$ beliebig ist.

Das sogenannte **Anfangswertproblem** für diese Differentialgleichung, d.h.

$$\text{(AWP)} \begin{cases} y' = \gamma y \\ y(0) = y \end{cases}$$

besitzt die (wie wir sehen werden) eindeutige Lösung $y(t) = y_0 e^{\gamma t}$ für $t \in \mathbb{R}$.

Grafisch:



Man sieht ein, dass das Modell zu vereinfacht ist (zumindest für große Zeitintervalle). Es sind komplexere Modelle vorgeschlagen worden, etwa

2. Logistisches Modell (Verhulst, 1838)

$$y'(t) = \underbrace{(\alpha - \beta)}_{>0} y(t) \qquad \underbrace{-\gamma y^2(t)}_{\text{selbstregulierender Effekt}}$$

Mit $\gamma = (\alpha - \beta)\,\frac{1}{K}$ für geeignetes $K > 0$ hat man

$$y'(t) = \underbrace{(\alpha - \beta)\, y(t)}_{\text{Term aus dem exponentiellen Modell}} \underbrace{\left(1 - \frac{y(t)}{K}\right)}_{\text{Korrekturterm}}$$

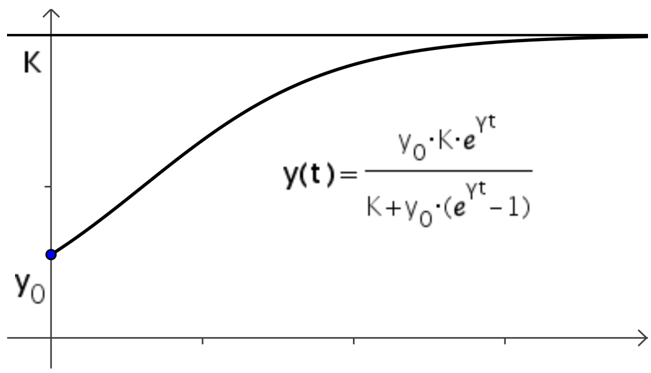
Für $y(t) \approx K$ ist der Korrekturterm ≈ 0, für kleine $y(t)$ ist er ≈ 1.

Das Anfangswertproblem

$$\begin{cases} y'(t) = (\alpha - \beta)\, y(t)(1 - \frac{y(t)}{K}) \\ y(0) = y_0 \end{cases}$$

für die logistische Gleichung, welches Ressourcenknappheit berücksichtigt, besitzt folgende eindeutige Lösung

$$y(t) = \frac{y_0 K \mathrm{e}^{(\alpha-\beta)t}}{K + y_0\left(\mathrm{e}^{(\alpha-\beta)t} - 1\right)} \qquad t > 0$$



3. Räuber-Beute-Modell

   Modellannahmen: Gegeben sind zwei Populationen, eine Räuber- und eine Beutepopulation (z.B. Füchse und Hasen, Hechte und Karpfen oder ähnliches), die sich gegenseitig beeinflussen.

   $$\begin{aligned} y(t) &:= \text{ Anzahl der Räuber zur Zeit } t \\ x(t) &:= \text{ Anzahl der Beute zur Zeit } t \end{aligned}$$

   Weitere Annahmen:

   - Beute: Es herrscht keine Nahrungsmittelknappheit oder sonstiger Ressourcenmangel. Sind also keine Räuber vorhanden, wir das Wachstum der Beutepopulation annähernd gut durch $x'(t) = ax(t)$ mit $a > 0$ beschrieben (Malthus-Modell).

     Sind Räuber vorhanden, verringert sich das Wachstum der Beute durch einen Term, der proportional zur Beute und zur Räuberpopulation ist: $-bx(t)y(t)$, mit $b > 0$ („Räuber fressen Beutetiere“). Also hat man
     $$x'(t) = ax(t) - bx(t)y(t)$$

   - Räuber: Räuber ernähren sich aussschließlich von Beute; sind keine Beutetiere vorhanden, verhungern die Räuber und sterben aus. Als Wachstumsmodell modelliert man dies durch $y'(t) = -cy(t)$ mit $c > 0$. Ist Beute vorhanden, kann die Räuberpopulation wachsen und sich reproduzieren. Das Malthusmodelle muss durch einen positiven Term der Form $+dx(t)y(t)$ $(d > 0)$ ergänzt werden:
     $$y'(t) = -cy(t) + dx(t)y(t)$$

   Als gesamtes Modell (Räuber-Beute-/Lotka-Volterra-Modell, 1925) ergibt sich

   $$\begin{aligned} x'(t) &= ax(t) - bx(t)y(t) \\ y'(t) &= -cy(t) + dx(t)y(t) \end{aligned}$$

   oder kurz

   $$\begin{aligned} x' &= ax - bxy = x(a - by) \\ y' &= -cy + dxy = y(dx - c) \end{aligned}$$

gegebenfalls ergänzt durch die Anfangsbedingung

$$x(0) = x_0 > 0,\ y(0) = y_0 > 0$$

Die Differentialgleichung aus dem Räuber-Beute-Modell besitzt keine Lösung, die explizit analytisch, in geschlossener Form angegeben werden kann. Aber man kann zeigen: das Anfangswertproblem

$$\begin{aligned} x' &= x(a - by) \\ y' &= y(dx - c) \\ x(0) &= x_0 \\ y(0) &= y_0 \end{aligned}$$

besitzt, für jedes Anfangswertepaar $(x_0, y_0) \in \mathbb{R}^2$ eine eindeutige Lösung $(x(t), y(t))$ auf $\mathbb{R}$. Diese kann mit numerischen Verfahren beliebig genau berechnet werden. Typischer Verlauf der Lösung:

8.1.3 Fragen an Differentialgleichungsmodelle

Gegeben ist ein Modell mit Differentialgleichungen/Anfangswertproblemen. Typische Fragen sind

- Existiert eine Lösung?
- Was ist das maximale Existenzintervall einer Lösung?
- Ist die Lösung eindeutig?
- Ist die Lösung explizit berechenbar? Ansonsten müssen numerische Verfahren zur Berechnung der Lösung entwickelt werden.
- Welche qualitativen Eigenschaften hat die Lösung (Monotonie, Periodizität, Langzeitverhalten usw.)?

Sie sind Gegenstand der Vorlesung „Differentialgleichungen 1“. Hier nur ganz kurz einige Überlegungen:

- Wir betrachten $X = \mathbb{R}$. Die Differentialgleichung $y'(t) = g(t)$ mit g als Dirichlet-Funktion. g besitzt keine Stammfunktion, weshalb die Differentialgleichung in keinem noch so kleinen Intervall eine Lösung besitzt. Allgemein sollte die rechte Seite der Differentialgleichung für die Existenz einer Lösung stetig sein.
- Klar: Die Lösung einer Differentialgleichung, wenn denn eine solche existiert, ist im Allgemeinen nicht eindeutig. Betrachte zum Beispiel $y'(t) = g(t)$ mit $g : I \to \mathbb{R}$ stetig. Die Gleichung besitzt die unendlich vielen Lösungen

  $$y_c(t) = \int\limits_{t_0}^{t} g(s)\,\mathrm{d}s + c, \quad c \in \mathbb{R}, t \in I$$

  Das zugehörige Anfangswertproblem

  $$\begin{aligned} y' &= g \\ y(0) &= y_0 \end{aligned}$$

besitzt dagegen die eindeutige Lösung

$$y(t) = y_0 + \int\limits_{t_0}^{t} g(s)\,\mathrm{d}s, \quad t \in I$$

Aber: Für allgemeine Differentialgleichungen mit stetiger rechter Seite ist es möglich, dass das Anfangswertproblem mehrere Lösungen besitzt. Betrachte zum Beispiel $y' = \sqrt[3]{y^2}$, wobei die rechte Seite,

$$\begin{aligned} y :\ \mathbb{R} &\to \mathbb{R} \\ y &\mapsto \sqrt[3]{y^2} \end{aligned}$$

offensichtlich stetig ist.

- Man kann zeigen, dass in endlich-dimensionalen Banachräumen ein Anfangswertproblem für eine Differentialgleichung mit stetiger rechter Seite stets eine Lösung besitzt (Satz von Peano). Aber im vorliegenden Fall besitzt das Anfangswertproblem

$$\begin{aligned} y' &= \sqrt[3]{y^2} \\ y(0) &= y_0 \end{aligned}$$

die unendlich vielen verschiedenen Lösungen

$$y_{\alpha,\beta}(t) = \begin{cases} \frac{1}{27}\left(t-\alpha\right)^3, & t \le \alpha \\ 0, & \alpha \le t \le \beta \\ \frac{1}{27}\left(t-\beta\right)^3, & \beta \le t \end{cases}$$

für $\alpha < 0 < \beta$.

8.1.4 Satz von Picard-Lindelöf, lokale Version

Im vorliegenden Beispiel ist die rechte Seite $g(t, y)$ zwar stetig (in (t, y)), aber bezüglich y nicht lipschitz-stetig in 0. Das ist das Problem, wie folgender Satz zeigt

Satz 8.1.2. Sei $(X, \|\cdot\|_X)$ ein Banachraum; seien $a, b \in \mathbb{R}$ mit $a < b$, $r > 0$, $u_0 \in X$, $g :\ [a, b] \times \overline{B_r(x_0)} \to X$ stetig und lipschitz-stetig bezüglich y, d.h. es existiert ein $L > 0$, sodass

$$\|g(t, y) - g(t, z)\|_X \le L\,\|y - z\|_X \quad \forall\, y, z \in \overline{B_r(x_0)}, t \in [a, b]$$

Dann besitzt das Anfangswertproblem

$$\begin{aligned} y' &= g(t, y) \\ y(t_0) &= u_0 \end{aligned}$$

für beliebiges $t_0 \in [a, b]$ eine eindeutige Lösung auf einem Intervall $[t_0 - \alpha, t_0 + \alpha]$, wobei $\alpha > 0$.

Beweis. Vorüberlegungen: Eine Funktion u ist Lösung des Anfangswertproblems auf dem Intervall $[t_0 - \alpha, t_0 + \alpha]$ für $\alpha > 0$ genau dann, wenn

$$\begin{aligned} u:\ [t_0 - \alpha, t_0 + \alpha] \to X & \text{ ist differenzierbar} \\ u(t_0) &= u_0 \\ u'(t) &= g(t, u(t)) \quad \forall\, t \in [t_0 - \alpha, t_0 + \alpha] \end{aligned}$$

Da $g(t, u(t))$ in t stetig ist, ist $u'(t)$ also auch stetig. Nach dem Hauptsatz der Differential- und Integralrechnung folgt aus der letzten Zeile (es gilt auch die Rückrichtung)

$$u(t) - u(t_0) = \int\limits_{t_0}^{t} g(s, u(s))\,\mathrm{d}s \quad \forall\, t \in [t_0 - \alpha, t_0 + \alpha] \qquad \Leftrightarrow$$

$$u(t) = u(t_0) + \int\limits_{t_0}^{t} g(s, u(s))\,\mathrm{d}s, \quad u \in C([t_0 - \alpha, t_0 + \alpha], X)$$

Letzteres kann als Fixpunktproblem formuliert werden und mit dem Banach'schen Fixpunktsatz (in eindeutiger Weise) gelöst werden.

Weitere Vorüberlegung: g ist beschränkt auf $[a, b] \times \overline{B_r(u_0)}$. Falls X endlich dimensional ist, ist dies klar, da $\overline{B_r(u_0)}$ kompakt ist und somit auch $[a, b] \times \overline{B_r(u_0)}$ kompakt ist, womit g als stetige Funktion insbesondere beschränkt sein muss. Im allgemeinen Fall hat man für alle $t \in [a, b], y \in \overline{B_r(u_0)}$

$$\begin{aligned} \|g(t, y)\|_X &\leq \|g(t, y) - g(t, u_0)\|_X + \|g(t, u_0)\|_X \\ &\leq L \underbrace{\|y - y_0\|_X}_{\leq r} + \|g(t, u_0)\|_X \end{aligned}$$

Also ist

$$\|g(t, y)\|_X \leq L \cdot r + \max_{t \in [a,b]} \Bigg\| \underbrace{g(t, u_0)}_{\text{stetig in } t} \Bigg\| =: M$$

Definiere nun $\alpha := \min\left\{\frac{1}{2L}, \frac{r}{M}\right\}$,

$$E := \{u \in C([a, b], X) \mid \|u(t) - u(t_0)\|_X \leq r \quad \forall\, t \in [t_0 - \alpha, t_0 + \alpha]\},$$

$$\begin{aligned} \Phi:\ E &\to C([a, b], X) \\ u &\mapsto \Phi(u) \end{aligned}$$

wobei

$$\Phi(u)(t) := u_0 + \int\limits_{t_0}^{t} g(s, u(s))\,\mathrm{d}s \quad \forall\, t \in [t_0 - \alpha, t_0 + \alpha]$$

E ist eine nicht-leere (zumindest konstante Funktion sind in E enthalten), abgeschlossene Teilmenge des Banachraums $(C([t_0 - \alpha, t_0 + \alpha], X), \|\cdot\|_\infty)$, wobei $\|\cdot\|_\infty$ die Supremumsnorm

$$\|u\|_\infty := \sup_{t \in [t_0 - \alpha, t_0 + \alpha]} \|u(t)\|_X$$

bezeichnet. Weiter gilt für alle $u \in E, t \in [t_0 - \alpha, t_0 + \alpha]$

$$
\begin{aligned}
\left\| \Phi(u)(t) - u_0 \right\|_X &= \left\| \int\limits_{t_0}^{t} g(s, u(t)) \,\mathrm{d}s \right\|_X \\
&\leq \int\limits_{\min\{t_0,t\}}^{\max\{t_0,t\}} \left\| g(s, u(s)) \right\|_X \mathrm{d}s \\
&\leq M \left| t_0 - t \right| \leq M\alpha \leq r
\end{aligned}
$$

Also ist $\Phi(u) \in E$ für alle $u \in E$, d.h. es ist eine Selbstabbildung. Außerdem ist für alle $u, v \in E, t \in [t_0 - \alpha, t_0 + \alpha]$

$$
\begin{aligned}
\left\| \Phi(u)(t) - \Phi(v)(t) \right\|_X &= \left\| \int\limits_{t_0}^{t} \left(g(s, u(s)) - g(s, v(s)) \right) \mathrm{d}s \right\| \\
&\leq \int\limits_{\min\{t_0,t\}}^{\max\{t_0,t\}} \left\| g(s, u(s)) - g(s, v(s)) \right\|_X \mathrm{d}s \\
&\leq \alpha \cdot L \cdot \left\| u(s) - v(s) \right\|_X \leq \alpha L \cdot \left\| u - v \right\|_\infty \leq \frac{1}{2} \left\| u - v \right\|_\infty
\end{aligned}
$$

Da dies für alle $t \in [t_0 - \alpha, t_0 + \alpha]$ gilt, hat man

$$
\left\| \Phi(u) - \Phi(v) \right\|_\infty \leq \frac{1}{2} \left\| u - v \right\|_\infty
$$

Φ ist also kontrahierend. Nach dem Banach'schen Fixpunktsatz folgt, dass es genau ein $u \in E$ gibt mit $\Phi(u) = u$, also genau ein $u \in C([t_0 - \alpha, t_0 + \alpha], X)$ mit $\|u(t) - u_0\| \leq r$ für alle $t \in [t_0 - \alpha, t_0 + \alpha]$, sodass

$$
u(t) = u_0 + \int\limits_{t_0}^{t} g(s, u(s)) \,\mathrm{d}s
$$

$\square$

9 Bildquellen

Alle weiteren Bilder erstellt mit **GeoGebra** und dem **GIMP**.