ANALYSIS II

OLIVER C. SCHNÜRER

Zusammenfassung. Bei diesem Manuskript handelt es sich um Notizen zu einer Vorlesung Analysis II.

Inhaltsverzeichnis

Dies ist eine Fortsetzung des Skriptes zur Vorlesung Analysis I, siehe [3].

Wir orientieren uns an [2], was an [1] orientiert ist, und benutzen manchmal [5].

Wir empfehlen wiederum wie im ersten Semester, Resultate mit Banachräumen zunächst im Fall $\mathbb{R}$ oder $\mathbb{R}^n$ zu verstehen. Wünschenswert wäre im zweiten Semester

Date: 26. März 2020.

2000 *Mathematics Subject Classification.* 26-01.

Vielen Dank an Andrey Zakharov und die Hörer der Vorlesung für Korrekturen. In den Sommersemestern 2015 und 2020 in Konstanz benutzt.

nun insbesondere ab der Differentialrechnung in mehreren Variablen ein Verständnis zumindest im Falle von $\mathbb{R}^n$.

Mo 20.04.2020

4. Differentiation in einer Variablen

Wir beschränken uns hier auf den Fall von auf Teilmengen von $\mathbb{R}$ definierten Funktionen. Vieles lässt sich direkt auf den komplexen Fall übertragen, dies wird aber auch noch in der Vorlesung Funktionentheorie behandelt. Somit betrachten wir hier in der Regel auch Banachräume über dem Körper $\mathbb{R}$.

4.1. **Differenzierbare Funktionen.**

Definition 4.1. Sei $\Omega \subset \mathbb{R}$ offen, F eine Banachraum und $f\colon \Omega \to F$ eine Abbildung.

(i) Dann heißt f in $x_0 \in \Omega$ *differenzierbar,* falls

$$\lim_{\substack{\Omega \ni x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Der Quotient heißt *Differenzenquotient.* Sein Grenzwert heißt *Ableitung.* Wir bezeichnen ihn mit $f'(x_0)$ oder $\frac{df}{dx}(x_0)$. Es gilt $f'(x_0) \in F$.

(ii) Ist f in jedem Punkt $x_0 \in \Omega$ differenzierbar, so heißt f *in* Ω *differenzierbar.*

(iii) Ist f in Ω differenzierbar und ist die Abbildung $f'\colon \Omega \to F$ mit $x \mapsto f'(x)$ stetig, so heißt f in Ω *stetig differenzierbar.*

Definition 4.2 (Landausymbole)**.** Seien E, F Banachräume, $\Omega \subset E$ offen, $f\colon \Omega \to F$ und $g\colon \Omega \to \mathbb{R}_{>0}$ Abbildungen. Sei $x_0 \in \Omega$. Bei den nachfolgenden Definitionen ist stets kenntlich zu machen, dass x_0 der Bezugspunkt ist.

(i) Gilt

$$\lim_{\substack{x \to x_0 \\ \Omega \ni x \neq x_0}} \frac{f(x)}{g(x)} = 0,$$

so schreiben wir $f(x) \in o(g(x))$. Wir sagen dann, f sei ein klein-o von $g(x)$.

(ii) Gilt

$$\varlimsup_{\substack{x \to x_0 \\ \Omega \ni x \neq x_0}} \frac{\|f(x)\|}{g(x)} < \infty,$$

so schreiben wir $f(x) \in O(g(x))$. Wir sagen dann, f sei ein groß-O von $g(x)$.

Beispiele 4.3.

(i) Sei $f\colon \mathbb{R} \to \mathbb{R}$ die Abbildung $x \mapsto x^n$ für ein $n \in \mathbb{N}_{>0}$. Dann gelten für $x_0 = 0$

$$f(x) = o\left(|x|^\alpha\right) \quad \text{für alle} \quad 0 < \alpha < n \quad \text{und} \quad f(x) = O\left(|x|^n\right).$$

(ii) Für $f\colon \mathbb{R}^2 \to \mathbb{R}$ mit $f(x) = x^1 \cdot x^2$ gilt $f(x) = O\left(|x|^2\right)$ und daher folgt für $x_0 = 0$ und alle $0 < \alpha < 2$

$$f(x) = o\left(|x|^\alpha\right).$$

(iii) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x^2 \cos\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

ist in ganz $\mathbb{R}$ differenzierbar, aber f' ist im Punkt 0 nicht stetig.

Beweis. Übung für später oder unter Verwendung von Schulkenntnissen. $\square$

Proposition 4.4. *Sei F ein Banachraum und sei $\Omega \subset \mathbb{R}$ offen. Dann ist $f\colon \Omega \to F$ genau dann in $x_0 \in \Omega$ differenzierbar, wenn es ein $a \in F$ mit*

$$f(x) = f(x_0) + a \cdot (x - x_0) + o(|x - x_0|) \tag{4.1}$$

für alle $x \in \Omega$ gibt. In diesem Falle gilt $f'(x_0) = a$.

Beweis. „$\Rightarrow$“: Ist f im Punkt x_0 differenzierbar, so wählen wir $a = f'(x_0)$ und erhalten

$$f(x) - f(x_0) - a \cdot (x - x_0) = o(\|x - x_0\|)$$

nach Definition der Ableitung.

„$\Leftarrow$“: Nach (4.1) gilt für $x \neq x_0$ mit $x \in \Omega$

$$\left\| \frac{f(x) - f(x_0)}{x - x_0} - a \right\| = \frac{\|o(|x - x_0|)\|}{|x - x_0|} \to 0$$

für $x \to x_0$. Somit ist f im Punkt x_0 differenzierbar und es gilt $f'(x_0) = a$. □

Da die linearen stetigen Abbildungen $A\colon \mathbb{R} \to F$ genau von der Form $x \mapsto a \cdot x$ für einen Vektor $a = A(1) \in F$ sind, erhalten wir unter Verwendung der Konvention, dass wir Argumente von linearen Abbildungen in spitzen Klammern schreiben

Proposition 4.5. *Sei F ein Banachraum und $\Omega \subset \mathbb{R}$ offen. Dann ist $f\colon \Omega \to F$ genau dann in $x_0 \in \Omega$ differenzierbar, wenn es eine stetige lineare Abbildung $A \in L(\mathbb{R}, F)$ mit*

$$f(x) = f(x_0) + A\langle x - x_0 \rangle + o(\|x - x_0\|)$$

für alle $x \in \Omega$ gibt. Wir schreiben $A = Df(x_0)$.

Proposition 4.6. *Sei $\Omega \subset \mathbb{R}$ offen. Seien $f,\, g\colon \Omega \to \mathbb{R}$ in $x_0 \in \Omega$ differenzierbar.*

(i) Es gelten $(f + g)'(x_0) = f'(x_0) + g'(x_0)$ sowie $(\lambda f)'(x_0) = \lambda f'(x_0)$ für alle $\lambda \in \mathbb{R}$.

(ii) Es gilt die Produktregel

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).$$

(iii) Ist $g(x_0) \neq 0$, so gilt die Quotientenregel

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

Beweis von Proposition 4.6.

(i) Klar nach Definition und Grenzwertsätzen.

(ii) Es gilt

$$f(x)g(x) - f(x_0)g(x_0) = (f(x) - f(x_0))g(x) + (g(x) - g(x_0))f(x_0).$$

Wir dividieren nun durch $x - x_0$ und gehen zum Grenzwert $x \to x_0$ über. Die Behauptung folgt.

(iii) Die Behauptung folgt direkt aus der Produktregel und der noch zu zeigenden Identität

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{g^2(x_0)}.$$

Da $g(x_0) \neq 0$ ist, gilt auch $g(x) \neq 0$ in einer ganzen Umgebung von x_0, da g in x_0 stetig ist. Wir erhalten also

$$\frac{1}{g(x)} - \frac{1}{g(x_0)} = -\frac{g(x) - g(x_0)}{g(x)g(x_0)}.$$

Wir dividieren wiederum durch $x - x_0$ und gehen zum Grenzwert $x \to x_0$ über. Die Behauptung folgt. □

Korollar 4.7. *Sei $n \in \mathbb{N}_{>0}$. Dann ist $f_n\colon \mathbb{R} \to \mathbb{R}$ mit $x \mapsto x^n$ in ganz $\mathbb{R}$ differenzierbar und es gilt*
$$f_n'(x) = nx^{n-1}.$$

Beweis. Die Behauptung für $n = 1$ ist leicht einzusehen. Gelte daher die Aussage bereits für alle Funktionen f_m mit $m < n$. Mit Produktregel erhalten wir daraus
$$\begin{aligned} f_n'(x) &= (f_{n-1}f_1)'(x) = f_{n-1}'(x)f_1(x) + f_{n-1}(x)f_1'(x) \\ &= (n-1)x^{n-2} \cdot x + x^{n-1} \cdot 1 = nx^{n-1}. \end{aligned}$$ □

Theorem 4.8 (Kettenregel)**.** *Sei $\Omega, \tilde{\Omega} \subset \mathbb{R}$ offen. Seien $g\colon \Omega \to \mathbb{R}$ in x_0 und $f\colon \tilde{\Omega} \to \mathbb{R}$ in $g(x_0) \in \tilde{\Omega}$ differenzierbar, so ist $f \circ g$ in einer Umgebung von x_0 definiert, in x_0 differenzierbar und es gilt*
$$(f \circ g)'(x_0) = f'(g(x_0)) \cdot g'(x_0) = f'(g(x_0))\langle g'(x_0)\langle \cdot \rangle\rangle.$$

Beweis.

(i) g ist in x_0 differenzierbar, somit auch stetig und daher ist $f \circ g$ in einer Umgebung von x_0 definiert. Betrachte daher ab jetzt nur noch Werte in einer solchen Umgebung.

(ii) Aufgrund der Differenzierbarkeit von f im Punkt $g(x_0)$ gilt nach Proposition 4.5 mit $g(x)$ und $g(x_0)$ statt x und x_0
$$\begin{aligned} (f \circ g)(x) - (f \circ g)(x_0) &= f(g(x)) - f(g(x_0)) \\ &= f'(g(x_0))\langle g(x) - g(x_0)\rangle + o(|g(x) - g(x_0)|). \end{aligned}$$
Weiterhin gilt aufgrund der Differenzierbarkeit von g in x_0
$$g(x) - g(x_0) = g'(x_0)\langle x - x_0\rangle + o(|x - x_0|).$$
Zusammengenommen erhalten wir
$$\begin{aligned} (f \circ g)(x) - (f \circ g)(x_0) = f'(g(x_0))&\langle g'(x_0)\langle x - x_0\rangle\rangle \\ &+ f'(g(x_0))\langle o(|x - x_0|)\rangle + o(|g(x) - g(x_0)|). \end{aligned}$$

(iii) Es fehlt noch der Nachweis, dass die beiden Terme in der letzten Zeile in $o(|x - x_0|)$ liegen. Für den ersten Term ist dies klar. Da wir eine Funktion in $o(|x - x_0|)$ stets in der Form $\varepsilon(|x - x_0|) \cdot |x - x_0|$ mit $\varepsilon(|x - x_0|) \to 0$ für $x \to x_0$ schreiben können, genügt der Nachweis, dass
$$\lim_{x \to x_0} \varepsilon(|g(x) - g(x_0)|) \cdot \frac{|g(x) - g(x_0)|}{x - x_0} = 0$$
gilt. Dies folgt nun aus der Stetigkeit und Differenzierbarkeit von g in x_0 und Grenzwertsätzen. □

Mit Hilfe der Kettenregel kann man die Ableitung der Inversen bestimmen. Es gilt

Korollar 4.9. *Seien $\Omega, \tilde{\Omega} \subset \mathbb{R}$ offen. Sei $f\colon \Omega \to \tilde{\Omega}$ invertierbar und sei $g\colon \tilde{\Omega} \to \Omega$ die Inverse. Sei f in x_0 und g in $f(x_0)$ differenzierbar, so gilt*
$$g'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Beweis. Es gilt $\mathrm{id} = g \circ f$. Wir leiten dies ab und erhalten mit Hilfe der Kettenregel
$$1 = \mathrm{id}' = g'(f(x_0)) \cdot f'(x_0).$$
Somit ist insbesondere $f'(x_0) \neq 0$ und wir erhalten die behauptete Identität. □

Do 23.04.2020

Wir definieren iterativ höhere Ableitungen und zugehörige Funktionenräume.

Definition 4.10. Sei F ein Banachraum und sei $\Omega \subset \mathbb{R}$ offen.

(i) Fassen wir die Ableitung einer Funktion $f\colon \Omega \to F$ als Funktion $Df\colon \Omega \to F$ auf, so können wir induktiv die Abbildungen $D^0 f = f\colon \Omega \to F$, $D^1 f = Df\colon \Omega \to F$, $D^2 f = D(Df)\colon \Omega \to F$ sowie $D^{n+1} f = D\left(D^n f\right) : \Omega \to F$ definieren, soweit dies für die gegebene Funktion f möglich ist. Wir schreiben auch $D^n f \equiv f^{(n)} \equiv \frac{d^n f}{dx^n}$. Existiert $f^{(n)}$ (in einem Punkt $x_0 \in \Omega$), so heißt f (in x_0) *n-mal differenzierbar* und $f^{(n)}$ die *n-te Ableitung* von f.

(ii) Ist $f^{(n)}\colon \Omega \to F$ stetig, so heißt f eine *n-mal stetig differenzierbare Funktion.* Existiert $f^{(n)}$ für alle $n \in \mathbb{N}$, so heißt f unendlich oft differenzierbar.

(iii) Wir definieren die folgenden Funktionenräume:

$$\begin{aligned} C^0(\Omega, F) &:= \{f\colon (f\colon \Omega \to F) \text{ ist stetig}\},\\ C^0\left(\overline{\Omega}, F\right) &:= \left\{f\colon f \text{ ist stetig und beschränkt auf } \overline{\Omega} \text{ fortsetzbar}\right\},\\ C^k(\Omega, F) &:= \{f\colon \Omega \to F\colon\ f \text{ ist in } \Omega\ k\text{-mal stetig differenzierbar}\},\\ C^k\left(\overline{\Omega}, F\right) &:= \Big\{f \in C^k(\Omega, F)\colon\ \text{für alle } 0 \le n \le k \text{ ist } f^{(n)}\\ &\qquad\qquad \text{stetig und beschränkt auf } \overline{\Omega} \text{ fortsetzbar}\Big\} \end{aligned}$$

sowie den Raum der glatten Funktionen

$$C^\infty(\Omega, F) := \bigcap_{k \in \mathbb{N}} C^k(\Omega, F).$$

Ist $F = \mathbb{R}$, so schreiben wir $C^0(\Omega) \equiv C^0(\Omega, \mathbb{R})$, $C^k(\Omega) \equiv C^k(\Omega, \mathbb{R})$, $\ldots$.

Definition 4.11.

(i) Auf dem $\mathbb{R}$-Vektorraum $C^0\left(\overline{\Omega}, F\right)$ definieren wir eine Norm für eine Funktion $f\colon \Omega \to F$ durch

$$\|f\|_{C^0} \equiv \|f\|_{L^\infty} := \sup_{x \in \Omega} \|f(x)\|_F.$$

(ii) Sei $k \in \mathbb{N}$. Auf dem $\mathbb{R}$-Vektorraum $C^k\left(\overline{\Omega}, F\right)$ definieren wir eine Norm durch

$$\|f\|_{C^k(\Omega,F)} := \sum_{i=0}^{k} \left\|f^{(i)}\right\|_{C^0(\Omega,F)}.$$

Theorem 4.12. *Sei $\Omega \subset \mathbb{R}$ offen und sei F ein Banachraum. Sei $k \in \mathbb{N}$. Dann ist $C^k\left(\overline{\Omega}, F\right)$ mit der C^k-Norm ein Banachraum.*

Beweis. Übung. Benutze das später unabhängig davon bewiesene Theorem 4.31 und dass $C^0\left(\overline{\Omega}, F\right)$ ein Banachraum ist. □

Proposition 4.13. *Sei $\Omega \subset \mathbb{R}$ offen, sei $n \in \mathbb{N}_{>0}$ und sei $f\colon \Omega \to \mathbb{R}^n$ mit $f = \left(f^i\right)_{1 \le i \le n}$ eine Abbildung. Dann ist f genau dann in x_0 differenzierbar, wenn die Komponenten f^i, $1 \le i \le n$, in x_0 differenzierbar sind und es gilt dann*

$$f'(x_0) = \left(\left(f^i\right)'(x_0)\right)_{1 \le i \le n}.$$

Beweis. Übung. □

Definition 4.14. Seien E ein metrischer Raum und $f\colon E \to \mathbb{R}$ eine Abbildung. Sei $x_0 \in E$.

(i) Dann besitzt f in x_0 ein *lokales Minimum,* falls ein $U \in \mathcal{U}(x_0)$ mit $f(x_0) \le f(x)$ für alle $x \in U$ existiert.

(ii) Dann besitzt f in x_0 ein *(globales) Minimum,* falls $f(x_0) \le f(x)$ für alle $x \in E$ gilt.

(iii) Gilt sogar $f(x_0) < f(x)$ für alle entsprechenden $x \neq x_0$, so heißt x_0 ein *striktes lokales/globales Minimum.*
(iv) *Lokale, globale* und *strikte Maxima* sind analog oder als entsprechende Minima von $-f$ definiert.
(v) f besitzt in x_0 ein *(lokales) Extremum,* falls f in x_0 ein (lokales) Minimum oder Maximum besitzt.
(vi) Besitzt f in x_0 ein Extremum, so heißt x_0 eine *Extremalstelle.*

Proposition 4.15. *Seien $a < b \in \mathbb{R}$, also a, $b \in \mathbb{R}$ und $a < b$, sei $f\colon (a,b) \to \mathbb{R}$ eine Funktion mit einem lokalen Extremum in $x_0 \in (a,b)$. Ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.*

Beweis. Wir betrachten ohne Einschränkung nur den Fall, dass f in x_0 ein lokales Minimum annimmt.

Dann gibt es $\varepsilon > 0$, so dass für alle $x \in B_\varepsilon(x_0)$ die Ungleichung $f(x_0) \leq f(x)$ gilt. Wir erhalten
$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0$$
für alle $x \in B_\varepsilon(x_0)$ mit $x > x_0$. Somit folgt $f'(x_0) \geq 0$.

Durch analoge Betrachtungen für $x < x_0$ erhalten wir $f'(x_0) \leq 0$. Somit ist $f'(x_0) = 0$. ☐

Definition 4.16. Sei $\Omega \subset \mathbb{R}$ offen. Sei $f\colon \Omega \to \mathbb{R}$ eine in $x_0 \in \Omega$ differenzierbare Funktion mit $f'(x_0) = 0$. Dann heißt x_0 *kritischer Punkt* von f.

Bemerkung 4.17.

(i) Proposition 4.15 zeigt, dass jedes lokale Extremum einer auf einer offenen Teilmenge von $\mathbb{R}$ definierten Funktion ein kritischer Punkt ist.
(ii) $\mathbb{R} \ni x \mapsto x^4 - x^2 \in \mathbb{R}$ besitzt drei lokale Extrema und dies sind auch genau die kritischen Punkte dieser Funktion.
(iii) $\mathbb{R} \ni x \mapsto x^3$ besitzt in $x = 0$ einen kritischen Punkt, aber keine lokalen Extrema.

Proposition 4.18 (Satz von Rolle)**.** *Seien $a < b \in \mathbb{R}$. Sei $f \in C^0([a,b])$ in (a,b) differenzierbar und gelte $f(a) = f(b)$. Dann gibt es ein $\xi \in (a,b)$ mit $f'(\xi) = 0$.*

Beweis. Definiere
$$m := \inf_{[a,b]} f \qquad \text{und} \qquad M := \sup_{[a,b]} f.$$
Da $[a,b]$ kompakt ist, werden Minimum und Maximum auch angenommen. Es gilt $M \geq f(a) = f(b) \geq m$.

(i) Gilt in beiden Fällen Gleichheit, so ist f konstant und somit auch $f' \equiv 0$.
(ii) Sonst gelte ohne Einschränkung $M > f(a)$. Dann gibt es ein $\xi \in (a,b)$ mit $f(\xi) = M$. Nach Proposition 4.15 folgt also $f'(\xi) = 0$. ☐

Theorem 4.19 (Mittelwertsatz (MWS))**.** *Seien $a < b \in \mathbb{R}$. Sei $f \in C^0([a,b])$ und in (a,b) differenzierbar. Dann gibt es ein $\xi \in (a,b)$ mit*
$$f(b) - f(a) = f'(\xi) \cdot (b - a).$$

Beweis. Definiere die Funktion $g\colon [a,b] \to \mathbb{R}$ durch
$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$
Dann erfüllt g die Voraussetzungen des Satzes von Rolle und $\xi \in (a,b)$ mit $g'(\xi) = 0$ erfüllt gerade die Behauptung. ☐

Für einen Mittelwertsatz für vektorwertige Funktionen benötigen wir noch zwei Resultate:

Proposition 4.20. *Sei $\Omega \subset \mathbb{R}$ offen, F ein Banachraum und sei $f\colon \Omega \to F$ in x_0 differenzierbar. Sei $\varphi\colon F \to \mathbb{R}$ ein stetiges lineares Funktional. Dann ist $g := \varphi \circ f$ in x_0 differenzierbar und es gilt $g'(x_0) = \varphi(f'(x_0)) \equiv \langle f'(x_0), \varphi\rangle$.*

Beweis. Wende φ auf die Gleichung (4.1) an und beachte, dass $\varphi(o(|x-x_0|)) \subset o(|x-x_0|)$ gilt. Details: Übung. □

Lemma 4.21. *Sei E ein normierter $\mathbb{K}$-Vektorraum. Sei $u \in E$. Dann gilt*

$$\|u\| = \sup_{\substack{\varphi\in E^*\\ \|\varphi\|=1}} |\langle u, \varphi\rangle|.$$

Beweis. Wir können aktuell nur eine Ungleichung zeigen. Die andere folgt aus dem Satz von Hahn-Banach, den wir hier verwenden und später unabhängig davon in der Funktionalanalysis beweisen werden.

„$\ge$“: Nach Definition der Operatornorm gilt für alle $\varphi \in E^*$ gerade

$$\|\varphi\| = \sup_{\|u\|=1} |\varphi(u)| = \sup_{\|u\|=1} |\langle u, \varphi\rangle|.$$

Somit folgt insbesondere $|\langle u, \varphi\rangle| \le \|\varphi\| \cdot \|u\|$.

„$\le$“: Der Satz von Hahn-Banach liefert, dass es ein stetiges lineares Funktional $\varphi_0 \in E^*$ mit $\|\varphi_0\| = 1$ und $\varphi_0(u) = \|u\|$ gibt. Somit folgt

$$\|u\| = |\langle u, \varphi_0\rangle| \le \sup_{\substack{\varphi\in E^*\\ \|\varphi\|=1}} |\langle u, \varphi\rangle|.$$

Im Spezialfall eines Hilbertraumes (und insbesondere des $\mathbb{R}^n$) nehmen wir ohne Einschränkung $\|u\| = 1$ an. Dann ist dieses Funktional direkt durch $\varphi_0 := \langle \cdot, u\rangle$, also $\varphi_0(x) := \langle x, u\rangle$ gegeben. In diesem Fall gilt $\|\varphi_0\| = \sup\limits_{\|x\|=1} |\langle x, u\rangle| \le \sup\limits_{\|x\|=1} \|x\| \cdot \|u\| = 1$ aufgrund der Cauchy-Schwarzschen Ungleichung. Schließlich gelten noch $\|\varphi_0\| = \sup\limits_{\|x\|=1} |\langle x, u\rangle| \ge |\langle u, u\rangle| = \|u\|^2 = 1$ und $\varphi_0(u) = \|u\|^2 = 1$. □

Theorem 4.22 (Vektorwertiger Mittelwertsatz)**.** *Seien $a < b \in \mathbb{R}$. Sei F ein Banachraum, sei $f\colon [a,b] \to F$ stetig und in (a,b) differenzierbar. Dann gilt*

$$\|f(b) - f(a)\| \le \sup_{x\in(a,b)} \|f'(x)\| \cdot (b-a) = \sup_{x\in(a,b)} \|f'(x)\| \cdot |b-a|.$$

Beweis. Sei $\varphi \in F^*$ mit $\|\varphi\| = 1$ beliebig. Wir definieren eine Funktion $g\colon [a,b] \to \mathbb{R}$ durch $g(x) := \langle f(x), \varphi\rangle$. Dann ist $g \in C^0([a,b])$ und in (a,b) differenzierbar mit $g'(x) = \langle f'(x), \varphi\rangle$. Wir wenden nun den Mittelwertsatz auf g an und erhalten für ein $\xi \in (a,b)$ mit der Abschätzung vom Anfang des Beweises von Lemma 4.21

$$\begin{aligned} \langle f(b) - f(a), \varphi\rangle &= g(b) - g(a) = g'(\xi)\cdot(b-a)\\ &= \langle f'(\xi), \varphi\rangle \cdot (b-a) \le \|f'(\xi)\| \cdot \underbrace{\|\varphi\|}_{=1} \cdot (b-a). \end{aligned}$$

Da dies für beliebige $\varphi \in F^*$ mit $\|\varphi\| = 1$ gilt, folgt die Behauptung nun aus Lemma 4.21. □

Im folgenden Korollar ist $[x_0, x]$ kein Intervall, sondern eine Verbindungsstrecke wie in der Definition von Konvexität. Somit müssen wir nicht $x_0 < x$ annehmen.

Korollar 4.23. *Sei $\Omega \subset \mathbb{R}$ ein offenes Intervall, sei F ein Banachraum und sei $f\colon \Omega \to F$ differenzierbar. Seien x, $z_0 \in \Omega$. Dann gilt*

$$\|f(x) - f(x_0) - f'(z_0)\langle x - x_0\rangle\| \le \sup_{y\in[x_0,x]} \|f'(y) - f'(z_0)\| \cdot |x - x_0|.$$

Beweis. Für festes $z_0 \in \Omega$ definieren wir $g(x) := f(x) - f'(z_0)\langle x\rangle = f(x) - f'(z_0)\cdot x$. Wegen $g'(x) = f'(x) - f'(x_0)$ folgt die Behauptung nun direkt aus dem (vektorwertigen) Mittelwertsatz. □

Korollar 4.24. *Sei $\Omega \subset \mathbb{R}$ ein offenes Intervall, F ein Banachraum und $f\colon \Omega \to F$ differenzierbar. Gilt $f' \equiv 0$, so ist f eine konstante Abbildung.*

Beweis. Dies folgt direkt aus dem vektorwertigen Mittelwertsatz. □

Mo 27.04.2020

Definition 4.25. Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f\colon I \to \mathbb{R}$ heißt

(i) *monoton wachsend,* falls für alle x_1, $x_2 \in I$ aus $x_1 \le x_2$ die Ungleichung $f(x_1) \le f(x_2)$ folgt.
(ii) *strikt monoton wachsend,* falls für alle x_1, $x_2 \in I$ aus $x_1 < x_2$ die Ungleichung $f(x_1) < f(x_2)$ folgt.

Proposition 4.26. *Sei $f\colon (a,b) \to \mathbb{R}$ differenzierbar.*

(i) Dann ist f genau dann monoton wachsend, wenn $f'(x) \ge 0$ für alle $x \in (a,b)$ gilt.
(ii) Gilt $f'(x) > 0$ für alle $x \in (a,b)$, so ist f strikt monoton wachsend.

Beweis.

(a) Ist f monoton wachsend, so ist der Differenzenquotient nichtnegativ und folglich auch $f'(x) \ge 0$ für alle x.
(b) Ist $f'(x) \ge 0$ oder $f'(x) > 0$ für alle x, so folgt die (strikte) Monotonie aus dem eindimensionalen Mittelwertsatz. □

Theorem 4.27. *Sei $I = (a,b) \subset \mathbb{R}$. Sei $f\colon I \to \mathbb{R}$ differenzierbar und gelte $f' \ne 0$ in I. Dann ist $J := f(I) \subset \mathbb{R}$ offen und f besitzt eine differenzierbare Inverse $f^{-1}\colon J \to I$ mit*

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)} \quad \text{für alle } x \in I.$$

Beweis.

(i) f ist injektiv, denn sonst gäbe es nach dem Satz von Rolle ein $x \in (a,b)$ mit $f'(x) = 0$.
(ii) J ist als stetiges Bild einer zusammenhängenden Menge wieder zusammenhängend und somit ein Intervall.
(iii) J ist offen, denn wenn einer der Endpunkte zu J gehört, so nimmt f in I ein Maximum oder Minimum an, was aber wegen $f'(x) \ne 0$ für $x \in (a,b)$ ausgeschlossen ist.
(iv) Dieselbe Argumentation liefert, dass f eine offene Abbildung ist. Somit ist $f^{-1}\colon J \to I$ stetig, $f\colon I \to J$ also ein Homöomorphismus.
(v) Da f ein Homöomorphismus ist, folgt

$$x \to x_0 \quad \Longleftrightarrow \quad f(x) \to f(x_0).$$

Wir setzen nun $\xi = f(x)$ und $\xi_0 = f(x_0)$. Dann gilt

$$\frac{f^{-1}(\xi) - f^{-1}(\xi_0)}{\xi - \xi_0} = \frac{x - x_0}{f(x) - f(x_0)} = \left(\frac{f(x) - f(x_0)}{x - x_0}\right)^{-1}.$$

Wir betrachten nun den Grenzwert $x \to x_0$ mit $x \ne x_0$ oder, aufgrund der obigen Äquivalenz und der Injektivität äquivalent dazu, $\xi \to \xi_0$ mit $\xi \ne \xi_0$ und erhalten die Behauptung. □

4.2. **Die Regeln von de l'Hospital.** Die Regeln von (de) l'Hospital oder die l'Hospitalschen Regeln erlauben es, Grenzwerte wie $\lim\limits_{x\to 0} \frac{\sin x}{x} = 1$ zu bestimmen.

Als wichtiges Hilfsmittel benötigen wir den folgenden Satz.

Theorem 4.28 (verallgemeinerter Mittelwertsatz)**.** *Sei $I = (a,b) \subset \mathbb{R}$ ein Intervall und seien die Funktionen f, $g \in C^0([a,b])$ in (a,b) differenzierbar. Sei weiterhin $g'(x) \neq 0$ für alle $x \in (a,b)$. Dann gibt es ein $\xi \in I$ mit*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Beweis von Theorem 4.28.

(i) Wegen $g'(x) \neq 0$ ist g nach dem Satz von Rolle injektiv. Aufgrund der Stetigkeit ist g daher auch monoton. Ohne Einschränkung dürfen wir annehmen, dass g monoton wachsend ist. Dann folgt $g([a,b]) = [\alpha,\beta]$ mit $\alpha = g(a)$ und $\beta = g(b)$. Weiterhin gilt $g' > 0$ in I.

(ii) Nach Theorem 4.27 besitzt g daher eine in $J := (\alpha,\beta)$ differenzierbare Inverse φ mit $\varphi'(g(x)) = \frac{1}{g'(x)}$ für alle $x \in I$. Wir definieren $h := f \circ \varphi$ und erhalten aus dem Mittelwertsatz

$$h(\beta) - h(\alpha) = h'(\gamma)(\beta - \alpha)$$

für ein $\gamma \in (\alpha,\beta)$. Es gelten $h(\beta) = f(b)$, $h(\alpha) = f(a)$, $\beta = g(b)$, $\alpha = g(a)$ sowie

$$h'(\gamma) = f'(\varphi(\gamma))\varphi'(\gamma) = \frac{f'(\varphi(\gamma))}{g'(\varphi(\gamma))}.$$

Setzen wir dies oben ein, so folgt das Theorem mit $\xi = \varphi(\gamma)$. □

Wir formulieren das folgende Resultat nur für rechtsseitige Umgebungen von a.

Proposition 4.29. *Seien $a \in \mathbb{R}$. Sei U eine offene rechtsseitige Umgebung von a, d. h. eine offene Umgebung von a in $[a,\infty)$. Seien f, $g \in C^0(U)$ in $U \setminus \{a\}$ differenzierbar. Sei weiterhin $f(a) = g(a) = 0$ und $g'(x) \neq 0$ für alle $x \in U \setminus \{a\}$. Dann existiert*

$$\lim_{\substack{U \ni x \to a\\ x \neq a}} \frac{f(x)}{g(x)}, \qquad \textit{falls} \qquad \lim_{\substack{U \ni x \to a\\ x \neq a}} \frac{f'(x)}{g'(x)}$$

existiert, und die beiden Limites stimmen überein.

Beweis. Gelte $[a,b] \subset U$. Sei $x \in (a,b]$ beliebig. Dann gibt es aufgrund des verallgemeinerten Mittelwertsatzes ein $\xi = \xi(x) \in (a,x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Wir lassen nun $x \to a$ und somit $\xi(x) \to a$ und erhalten die Behauptung. □

Bemerkung 4.30.

(i) Es gilt unter Benutzung von Schulwissen ($\sin' x = \cos x$)

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1.$$

4.3. **Differentiation von Funktionenfolgen.**

Theorem 4.31. *Seien E, F metrische Räume. Sei F vollständig. Sei $(f_n)_n$ eine Folge von stetigen Funktionen $f_n \colon E \to F$, die gleichmäßig gegen eine Funktion $f \colon E \to F$ konvergiert. Dann ist f stetig.*

Beweis. Sei $x_0 \in E$ beliebig. Sei weiterhin $\varepsilon > 0$ beliebig. Aufgrund der gleichmäßigen Konvergenz gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq n_0$ und alle $x \in E$

$$d(f(x), f_n(x)) < \frac{\varepsilon}{3} \quad \text{gilt.}$$

Da f_{n_0} in x_0 stetig ist, gibt es ein $\delta > 0$, so dass für alle $x \in B_\delta(x_0)$

$$d(f_{n_0}(x), f_{n_0}(x_0)) < \frac{\varepsilon}{3} \quad \text{gilt.}$$

Insgesamt erhalten wir also für sämtliche $x \in B_\delta(x_0)$

$$d(f(x), f(x_0)) \leq \underbrace{d(f(x), f_{n_0}(x))}_{<\varepsilon/3} + \underbrace{d(f_{n_0}(x), f_{n_0}(x_0))}_{<\varepsilon/3} + \underbrace{d(f_{n_0}(x_0), f(x_0))}_{<\varepsilon/3} < \varepsilon.$$

Somit ist f (in x_0) stetig. ☐

Gleichmäßige Konvergenz sichert auch, dass wir den Grenzwert in der Folge und die Ableitung, letztlich auch ein Grenzwert, vertauschen dürfen, dass also

$$\lim_{n\to\infty} f_n' = \left(\lim_{n\to\infty} f_n\right)'$$

gilt.

Theorem 4.32. *Sei $\Omega \subset \mathbb{R}$ offen. Sei F ein Banachraum und $f_n \colon \Omega \to F$ eine Folge stetig differenzierbarer Funktionen mit $f_n \to f$ und $f_n' \rightrightarrows g$ für eine Funktion $g \colon \Omega \to F$. Dann ist auch f stetig differenzierbar und es gilt $f' = g$.*

Beweis. Wir wollen nachweisen, dass f differenzierbar ist und dass $f' = g$ gilt. Dazu genügt natürlich der Nachweis von $f' = g$. Die Stetigkeit der Ableitung folgt aus Theorem 4.31, da $f_n' \rightrightarrows g$ gilt.

Sei $x_0 \in \Omega$ beliebig. Sei $\varepsilon > 0$. Wir wollen nachweisen, dass es ein $\delta > 0$ gibt, so dass für alle $x \in B_\delta(x_0)$ mit $x \neq x_0$

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

gilt. Es gilt

$$\begin{aligned}
&\left\| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right\| \\
\leq &\left\| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_n(x) - f_n(x_0)}{x - x_0} \right\| + \left\| \frac{f_n(x) - f_n(x_0)}{x - x_0} - g(x_0) \right\| \\
\equiv &I_1 + I_2.
\end{aligned}$$

Wir zeigen zunächst, dass der Term I_2 klein wird. Nach Korollar 4.23, einem verallgemeinerten Mittelwertsatz, gilt für $x \in B_\delta(x_0)$

$$\begin{aligned}
I_2 &\leq \sup_{\xi \in B_\delta(x_0)} \|f_n'(\xi) - g(x_0)\| \\
&\leq \sup_{\xi \in B_\delta(x_0)} \|f_n'(\xi) - g(\xi)\| + \sup_{\xi \in B_\delta(x_0)} \|g(\xi) - g(x_0)\| \\
&< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{2\varepsilon}{4} = \frac{\varepsilon}{2},
\end{aligned}$$

falls wir zunächst $\delta > 0$ so klein wählen, dass der letzte Term aufgrund der Stetigkeit von g kleiner als $\varepsilon/4$ ist und dann n so vergrößern, dass der erste Term aufgrund der gleichmäßigen Konvergenz $f_n' \rightrightarrows g$ kleiner als $\varepsilon/4$ ist. All dies können wir so machen (und tun es auch, ebenso im Folgenden), dass dies für alle größeren n ebenfalls wahr bleibt.

Wir fixieren nun $x \in B_\delta(x_0)$ und erhalten $I_1 \leq \varepsilon/4$ für große n. Zusammengenommen folgt gerade die Behauptung. ☐

Damit können wir die Ableitung einer Potenzreihe im Inneren ihres Konvergenzradiusses durch gliedweises Differenzieren ermitteln. Do 30.04.2020

Dies funktioniert genauso für Potenzreihen in $\mathbb{C}$, die in der Funktionentheorie noch eingehend untersucht werden.

Theorem 4.33. *Sei $((a_n x^n))_{n\in\mathbb{N}}$ eine Potenzreihe in $\mathbb{R}$ mit Konvergenzradius $r > 0$. Dann ist die Funktion*

$$f\colon B_r(0) \to \mathbb{R},$$
$$f(x) := \sum_{n=0}^{\infty} a_n x^n$$

in $B_r(0)$ stetig differenzierbar und es gilt für $x \in B_r(0)$

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

Insbesondere ist eine Potenzreihe innerhalb ihres Konvergenzradiusses also beliebig oft (stetig) differenzierbar.

Beweis.

(i) Die Potenzreihe $((na_n x^{n-1}))_{n\in\mathbb{N}}$ hat (mindestens (es gilt sogar Gleichheit)) wieder den Konvergenzradius r: Man sieht direkt durch Ausklammern eines Faktors x, dass die Reihe $((na_n x^n))_n$ genau für die Werte von x konvergiert, für die die Reihe $\big((na_n x^{n-1})\big)_n$ konvergiert.

Nach Lemma 4.34 gilt $\lim\limits_{n\to\infty} n^{1/n} = 1$. Damit folgt

$$\frac{1}{r} = \varlimsup_{n\to\infty} |a_n|^{1/n} = \varlimsup_{n\to\infty} (n|a_n|)^{1/n}$$

wie behauptet.

(ii) Sei $\rho < r$. Dann konvergiert die Potenzreihe $\big((na_n x^{n-1})\big)_n$ in $B_\rho(0)$ gleichmäßig (Analysis I) und ist nach Theorem **??** (wir dürfen Differentiation und $n \to \infty$ für die Partialsummen vertauschen) als gliedweise Ableitung der Potenzreihe $((a_n x^n))_n$ in $B_\rho(0)$ die Ableitung von f. Da ρ mit $0 < \rho < r$ beliebig war, hat f' die angegebene Gestalt.

(iii) Per Induktion folgt, dass f in $B_r(0)$ beliebig oft differenzierbar ist. Somit sind auch sämtliche Ableitungen von f in $B_r(0)$ stetig. $\square$

Lemma 4.34. *Es gilt*

$$\lim_{n\to\infty} n^{1/n} = 1.$$

Beweis. Sei $\varepsilon > 0$. Wäre $n^{1/n} \geq 1 + \varepsilon$ für unendlich viele $n \in \mathbb{N}$, so folgte für diese aufgrund der binomischen Formel

$$n \geq (1+\varepsilon)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k}\varepsilon^k \geq 1^n + \frac{n}{1}1^{n-1}\varepsilon + \frac{n(n-1)}{1\cdot 2}1^{n-2}\varepsilon^2.$$

Dies ist nicht möglich, da die linke Seite linear in n, aber die rechte Seite quadratisch in n wächst. $\square$

Beispiel 4.35. Für die Exponentialfunktion gilt mit Indexverschiebung

$$\exp'(x) = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!}\right)' = \sum_{n=1}^{\infty} n\frac{x^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x).$$

4.4. **Elementare Funktionen.** Wir beginnen mit einem Resultat über gewöhnliche Differentialgleichungen, das wir erst im dritten Semester beweisen werden.

Theorem 4.36. *Sei E ein reeller Banachraum, $A \in L(E)$ und $x_0 \in E$. Dann besitzt das Anfangswertproblem*

$$\begin{cases} \dot{x}(t) = Ax(t), & t \in \mathbb{R}, \\ x(0) = x_0 \end{cases}$$

eine eindeutig bestimmte Lösung $x \in C^\infty(\mathbb{R}, E)$.

Beweis. Drittes Semester. In diesem speziellen Fall kann man eine Lösung direkt hinschreiben, die Eindeutigkeit folgt aus dem Gronwallschen Lemma, das jedoch Integrale benutzt. Vergleiche [6]. □

Definition 4.37. Seien E, F metrische Räume und $f_n, f\colon E \to F$ Abbildungen.

(i) Dann heißt f_n *lokal gleichmäßig konvergent* mit Grenzwert f, wenn jeder Punkt $x_0 \in E$ eine Umgebung U besitzt, so dass $f_n \rightrightarrows f$ in U gilt.

(ii) Dann heißt f_n *kompakt konvergent* mit Grenzwert f, wenn für jede kompakte Teilmenge $K \subset E$ gleichmäßige Konvergenz $f_n|_K \rightrightarrows f|_K$ vorliegt, d. h. für jede kompakte Teilmenge $K \subset E$ und jedes $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ und alle $x \in K$ die Abschätzung $d(f_n(x), f(x)) < \varepsilon$ gilt.

Lemma 4.38. *Seien E, F metrische Räume und $f_n, f\colon E \to F$ Abbildungen.*

(i) Konvergiert f_n lokal gleichmäßig gegen f, so ist $(f_n)_n$ kompakt konvergent mit Grenzwert f.

(ii) Ist E lokal kompakt, d. h. besitzt jeder Punkt in E eine kompakte Umgebung, so folgt aus kompakter Konvergenz definitionsgemäß auch lokal gleichmäßige Konvergenz. (In dieser Situation spricht man eher von lokal gleichmäßiger Konvergenz als von kompakter Konvergenz.)

Beweis. Übung. □

Bemerkung 4.39. Unsere Aussagen über Reihen und deren gliedweise Ableitungen erlauben es, zu zeigen, dass

$$x(t) = e^{tA}x_0 \qquad \text{mit} \qquad e^{tA} = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}$$

gilt. Dabei konvergieren die Reihe und die ihre Ableitung darstellende Reihe lokal gleichmäßig absolut.

Beweis. Übung. □

Wir erwähnen zunächst bereits Bekanntes oder direkt daraus Ableitbares über die Potenzfunktionen und die Exponentialfunktion.

Bemerkung 4.40 (Potenzfunktionen)**.** Seien $n, m \in \mathbb{N}_{>0}$. Dann ist die Funktion $f_n\colon \mathbb{R}_+ \to \mathbb{R}_+$ mit $f(x) = x^n$ unendlich oft differenzierbar mit $f_n'(x) = nx^{n-1}$ und wegen $f_n'(x) > 0$ für $x > 0$ strikt monoton wachsend. Somit ist $f_n\colon \mathbb{R}_+ \to \mathbb{R}_+$ ein Homöomorphismus (kleine Übung). Die Umkehrfunktion $g_n(x) = x^{1/n}$ ist in $\mathbb{R}_+$ stetig und wegen $f_n'(x) > 0$ für $x > 0$ in $\mathbb{R}_{>0}$ (unendlich oft) differenzierbar.

Sei $a > 0$. Dann gilt $(a^n)^{\frac{1}{m}} = \left(a^{\frac{1}{m}}\right)^n$, wie man durch Potenzieren mit m sieht. Wir definieren

$$a^{\frac{n}{m}} = (a^n)^{\frac{1}{m}} \qquad \text{und} \qquad a^{-\frac{n}{m}} = \left(a^{-1}\right)^{\frac{n}{m}}.$$

Damit gelten die Potenzgesetze

$$a^{x+y} = a^x \cdot a^y \qquad \text{und} \qquad (a^x)^y = a^{x \cdot y}$$

für $x, y \in \mathbb{Q}$, vergleiche auch die zugehörige Übungsaufgabe.

Bemerkung 4.41 (Exponentialfunktion)**.** Wir hatten die Exponentialfunktion als

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

für $x \in \mathbb{R}$ definiert. Wir wissen bereits, dass

$$\exp(x + y) = \exp(x) \cdot \exp(y)$$

für alle $x, y \in \mathbb{R}$ gilt. Wegen $1 = \exp(0) = \exp x \cdot \exp(-x)$ gilt somit $\exp(x) \neq 0$ für alle $x \in \mathbb{R}$ und wegen $\exp(x) = \exp\left(\frac{x}{2}\right) \cdot \exp\left(\frac{x}{2}\right) > 0$ oder der letzten Identität ist $\exp(x) > 0$ für alle $x \in \mathbb{R}$. Schließlich gilt $\exp' = \exp$.

Wegen $\exp'(x) > 0$ für alle $x \in \mathbb{R}$ ist die Exponentialfunktion invertierbar. Sie ist ein Homöomorphismus $\exp\colon \mathbb{R} \to \mathbb{R}_{>0}$, da $\exp(x) \to \infty$ für $x \to \infty$ und somit wegen $1 = \exp(x) \cdot \exp(-x)$ auch $\exp(x) \to 0$ für $x \to -\infty$ gelten. Wir definieren den Logarithmus, $\log\colon \mathbb{R}_{>0} \to \mathbb{R}$, als Umkehrfunktion der Exponentialfunktion. Aufgrund der Regel für die Ableitung der Umkehrfunktion gilt $\log'(x) = \frac{1}{x}$. Daher ist der Logarithmus beliebig oft differenzierbar.

Lemma 4.42. *Seien $x \in \mathbb{R}$ und $y \in \mathbb{Q}$. Dann gilt*

$$(\exp x)^y = \exp(xy).$$

Insbesondere folgt mit $x = \log a$ für $a > 0$ die Identität $a^y = \exp(y \log a)$ für alle $y \in \mathbb{Q}$.

Beweis. Wir definieren $f(x) := (\exp x)^y$ und $g(x) := \exp(xy)$. Es gelten $f(0) = 1$ und $g(0) = 1$. Die Funktionen f und g lösen beide eine Differentialgleichung, nämlich $f'(x) = yf(x)$ bzw. $g'(x) = yg(x)$. Da die Lösung dieser Differentialgleichung bei gleichem Anfangswert eindeutig bestimmt ist, folgt $f = g$ und somit die Behauptung. ☐

Motiviert durch diese Beziehung für rationale Exponenten definieren wir

Definition 4.43. Seien $a > 0$ und $x \in \mathbb{R}$. Dann definieren wir

$$a^x = \exp(x \log a).$$

Bemerkung 4.44.

(i) Für $x \in \mathbb{Q}$ stimmt diese Definition mit der bereits bekannten Definition überein.
(ii) Für $e := \exp(1) = a$ und somit $\log(e) = 1$ erhalten wir daher $e^x = \exp(x)$.

Die Exponentialfunktion wächst schneller als jede Potenz für $x \to \infty$ und fällt schneller als jede Potenz für $x \to -\infty$.

Proposition 4.45. *Es gelten für beliebiges $n \in \mathbb{N}$*

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty \qquad \textit{und} \qquad \lim_{x\to-\infty} e^x |x|^n = 0.$$

Beweis. Für $x > 0$ erhalten wir aus der Exponentialreihe $e^x \geq \frac{x^{n+1}}{(n+1)!}$. Hieraus folgt die erste Behauptung. Für die zweite Behauptung benutzen wir $e^x |x|^n = \frac{|-x|^n}{e^{-x}}$ und die erste Behauptung. ☐

Proposition 4.46. *Der Logarithmus erfüllt für alle $x, y, \varepsilon \in \mathbb{R}_{>0}$ und alle $\lambda \in \mathbb{R}$*

$$\log(xy) = \log x + \log y, \quad \log x^\lambda = \lambda \log x \quad \textit{und} \quad \lim_{x\to\infty} \frac{\log x}{x^\varepsilon} = 0.$$

Somit wächst der Logarithmus langsamer als jede Potenz.

Beweis. Die ersten beiden Identitäten zeigt man durch Anwenden der Exponentialfunktion, die dritte folgt aus der Regel von l'Hospital. □

Mo 04.05.2020

Definition 4.47. Sei $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Sei $x \in C^\infty\left(\mathbb{R}, \mathbb{R}^2\right)$ die eindeutig bestimmte Lösung der Differentialgleichung

$$\begin{cases} \dot{x}(t) = A\,x(t), & t \in \mathbb{R}, \\ x(0) = (0,1)^T. \end{cases}$$

Dann definieren wir die Funktionen $\sin\colon \mathbb{R} \to \mathbb{R}$ und $\cos\colon \mathbb{R} \to \mathbb{R}$ durch $\sin(t) := x^1(t)$ sowie $\cos(t) := x^2(t)$, also $x(t) \equiv \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}$.

Wir schreiben $\sin^k t \equiv (\sin t)^k$ bzw. $\cos^k t \equiv (\cos t)^k$ für alle zulässigen Exponenten, also zumindest für alle $k \in \mathbb{N}$.

Bemerkung 4.48.

(i) Es gelten $\sin' t = \cos t$ sowie $\cos' t = -\sin t$ für alle $t \in \mathbb{R}$.
(ii) Wir haben früher bereits die Potenzreihen für die Sinus- und die Kosinusfunktion angegeben. Nun wollen wir jedoch die Definition als Lösung einer gewöhnlichen Differentialgleichung verwenden und daraus dann später die Potenzreihendarstellung herleiten.

Proposition 4.49. *Es gilt*

$$\sin^2 t + \cos^2 t = 1$$

für alle $t \in \mathbb{R}$.

Beweis. Wir definieren $\varphi(t) := \sin^2 t + \cos^2 t$. Es folgt

$$\dot{\varphi}(t) = 2 \sin t \cos t - 2 \cos t \sin t = 0$$

und wir erhalten die Behauptung, da $(\sin(0), \cos(0)) = (0,1)$ gilt. □

Sinus und Kosinus sind ungerade bzw. gerade Funktionen.

Proposition 4.50. *Es gelten* $\sin t = -\sin(-t)$ *und* $\cos t = \cos(-t)$.

Beweis. Für $t = 0$ gilt die Behauptung. Definiere $\varphi(t) := -\sin(-t)$ und $\psi(t) := \cos(-t)$ Nun gilt

$$\frac{d}{dt}\begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix} = \begin{pmatrix} \cos(-t) \\ \sin(-t) \end{pmatrix} = \begin{pmatrix} \psi(t) \\ -\varphi(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix}.$$

Damit lösen $(\sin(t), \cos(t))^T$ und $(\varphi(t), \psi(t))^T$ dasselbe Anfangswertproblem mit demselben Anfangswert für $t = 0$ und stimmen daher überein. □

Um sich die folgenden Additionstheoreme zu merken, vergleicht man eine Drehmatrix um den Winkel $\alpha + \beta$ mit dem Produkt von zwei Drehmatrizen um die Winkel α und β. Dies benutzt jedoch implizit, dass die Abbildung, die α auf die Drehmatrix um den Winkel α abbildet, ein Gruppenhomomorphismus ist.

Proposition 4.51 (Additionstheoreme)**.** *Seien* x, $y \in \mathbb{R}$. *Dann gelten*

$$\sin(x+y) = \sin x \cdot \cos y + \cos x \cdot \sin y$$

und

$$\cos(x+y) = \cos x \cdot \cos y - \sin x \cdot \sin y.$$

Beweis. Wir definieren $\Phi(x) := (\sin(x+y), \cos(x+y))^T$ und

$$\Psi(x) := \begin{pmatrix} \sin x \cdot \cos y + \cos x \cdot \sin y \\ \cos x \cdot \cos y - \sin x \cdot \sin y \end{pmatrix}.$$

Wir behaupten nun, dass $\Phi(x) = \Psi(x)$ für alle $x \in \mathbb{R}$ gilt. Für $x = 0$ gilt

$$\Phi(0) = (\sin(y), \cos(y))^T = \Psi(0).$$

Weiterhin gelten

$$\frac{d}{dx}\Phi(x) = \begin{pmatrix} \cos(x+y) \\ -\sin(x+y) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Phi(x)$$

sowie

$$\frac{d}{dx}\Psi(x) = \begin{pmatrix} \cos x \cdot \cos y - \sin x \cdot \sin y \\ -\sin x \cdot \cos y - \cos x \cdot \sin y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Psi(x).$$

Da Φ und Ψ also dieselbe lineare gewöhnliche Differentialgleichung erfüllen und $\Phi(0) = \Psi(0)$ gilt, folgt $\Phi = \Psi$. □

Aus $\sin^2 x + \cos^2 x = 1$ folgt $\sin x, \cos x \in [-1, 1]$ für alle $x \in \mathbb{R}$. Wir wollen nun zeigen, dass $\operatorname{im} \sin = \operatorname{im} \cos = [-1, 1]$ gilt. Dazu zeigen wir zunächst

Lemma 4.52. *Es gibt ein $x > 0$ mit $\cos x = 0$.*

Beweis. Es gilt $\cos 0 = 1$. Falls es kein $x > 0$ mit $\cos x = 0$ gibt, so folgt aufgrund des Zwischenwertsatzes $\cos x > 0$ für alle $x > 0$. Somit ist $[0, \infty) \ni x \mapsto \sin x$ wegen $\sin' x = \cos x > 0$ strikt monoton wachsend und wir erhalten

$$0 = \sin 0 < \sin 1 \leq \sin x \quad \text{für alle } x \geq 1.$$

Damit liefert aber der Mittelwertsatz für den Kosinus auf dem Intervall $[1, x]$ für ein $\xi \in [1, x]$

$$\cos x - \cos 1 = -\sin \xi \cdot (x - 1) \leq -\sin 1 \cdot (x - 1).$$

Für $x \to \infty$ konvergiert die rechte Seite gegen minus Unendlich. Dies widerspricht $\cos^2 x \leq 1$. □

Definition 4.53 (Die Kreiszahl π)**.** Wir haben gezeigt, dass

$$M := \{x \in \mathbb{R}_+ : \cos x = 0\}$$

nichtleer ist. Aufgrund der Stetigkeit des Kosinusses ist M abgeschlossen. Somit gilt für $\pi := 2 \cdot \inf M$ die Identität $\cos \frac{\pi}{2} = 0$.

Proposition 4.54. *Auf $\left[0, \frac{\pi}{2}\right]$ ist der Sinus strikt monoton wachsend und es gilt*

$$0 = \sin 0 < \sin x < \sin \frac{\pi}{2} = 1 \quad \textit{für alle } 0 < x < \frac{\pi}{2}.$$

Beweis. Nach Definition von π gilt $\sin' x = \cos x > 0$ für $0 < x < \frac{\pi}{2}$. Wegen $\cos \frac{\pi}{2} = 0$, der Monotonie und $\sin^2 x + \cos^2 x = 1$ insbesondere für $x = \frac{\pi}{2}$ erhalten wir $\sin \frac{\pi}{2} = 1$. □

Korollar 4.55. *Aufgrund der Additionstheoreme erhalten wir daraus*

$$\begin{aligned}
\sin \pi &= \sin \frac{\pi}{2} \cdot \cos \frac{\pi}{2} + \cos \frac{\pi}{2} \cdot \sin \frac{\pi}{2} = 1 \cdot 0 + 0 \cdot 1 = 0, \\
\cos \pi &= \cos \frac{\pi}{2} \cdot \cos \frac{\pi}{2} - \sin \frac{\pi}{2} \cdot \sin \frac{\pi}{2} = 0 \cdot 0 - 1 \cdot 1 = -1, \\
\sin 2\pi &= \sin \pi \cdot \cos \pi + \cos \pi \cdot \sin \pi = 2 \cdot 1 \cdot 0 = 0, \\
\cos 2\pi &= \cos \pi \cdot \cos \pi - \sin \pi \cdot \sin \pi = (-1) \cdot (-1) - 0 \cdot 0 = 1.
\end{aligned}$$

Beweis. Bereits in die Formulierung eingebaut. □

Allgemeiner gilt

Korollar 4.56. *Für alle $x \in \mathbb{R}$ gelten*

$$\sin\left(x + \frac{\pi}{2}\right) = \cos x, \qquad \sin(x+\pi) = -\sin x, \qquad \sin(x+2\pi) = \sin x,$$
$$\cos\left(x + \frac{\pi}{2}\right) = -\sin x, \qquad \cos(x+\pi) = -\cos x, \qquad \cos(x+2\pi) = \cos x.$$

Beweis. Benutze die Additionstheoreme. $\square$

Korollar 4.57. *Es gilt* $\sin(\mathbb{R}) = [-1,1] = \cos(\mathbb{R})$.

Beweis. Es gelten $\sin\frac{\pi}{2} = 1$, $\sin\frac{3\pi}{2} = \cos\pi = -1$ und $\cos 0 = 1$. Somit werden aufgrund des Zwischenwertsatzes sämtliche Funktionswerte im Intervall $[-1,1]$ vom Sinus und Kosinus angenommen. Wegen $\sin^2 x + \cos^2 x = 1$ für alle $x \in \mathbb{R}$ ist dies bereits das gesamte Bild. $\square$

Nach Korollar 4.56 sind der Sinus und der Kosinus mit Periode 2π periodisch.

Definition 4.58. Sei E ein Vektorraum und F eine Menge. Sei $0 \neq \xi \in E$. Eine Abbildung $f\colon E \to F$ heißt mit Periode ξ periodisch, falls

$$f(x+\xi) = f(x) \quad \text{für alle } x \in E$$

gilt. Hieraus folgt natürlich auch $f(x + k\xi) = f(x)$ für alle $x \in E$ und alle $k \in \mathbb{Z}$.

Proposition 4.59. *Die Abbildung*

$$\begin{aligned} \varphi\colon [0,2\pi) &\to \mathbb{R}^2, \\ t &\mapsto (\cos t, \sin t) \end{aligned}$$

ist injektiv und es gilt $\operatorname{im}\varphi = \mathbb{S}^1 \subset \mathbb{R}^2$.

Beweis. Im Intervall $\left[0, \frac{\pi}{2}\right]$ ist der Sinus strikt monoton wachsend von 0 bis 1. Somit ist φ dort injektiv. Wegen $\cos' x = -\sin x < 0$ für $0 < x < \frac{\pi}{2}$ ist der Kosinus in diesem Intervall von 1 nach 0 fallend, ist also insbesondere nichtnegativ. $\sin x$ durchläuft in diesem Intervall alle Werte in $[0,1]$. Da es zu jedem $a = \sin x$ genau ein $b \geq 0$ mit $a^2 + b^2 = 1$, das den Einheitskreis beschreibt, gibt und $b = \cos x$ die Gleichung $a^2 + b^2 = 1$ erfüllt, folgt

$$\mathbb{S}^1 \cap \Gamma_+ \equiv \mathbb{S}^1 \cap \left\{x = \left(x^i\right) \in \mathbb{R}^2 \colon x^i > 0 \quad \text{für alle } i = 1,2\right\} \subset \operatorname{im}\varphi.$$

Γ_+ heißt positiver Quadrant und wird im $\mathbb{R}^n$ entsprechend als die Menge definiert, in der alle Koordinaten bezüglich der Standardbasis positiv sind.

Analoge Überlegungen für die weiteren Quadranten und die Intervalle $\left[\frac{\pi}{2}, \pi\right]$, $\left[\pi, \frac{3\pi}{2}\right]$ sowie $\left[\frac{3\pi}{2}, 2\pi\right]$ liefern nun die Behauptung über das Bild. Die Injektivität folgt aus der Injektivität auf jedem dieser Intervalle und Vorzeichenbetrachtungen. $\square$

Proposition 4.60 (Eulersche Formel)**.** *Für $t \in \mathbb{R}$ gilt*

$$e^{it} = \cos t + i \sin t.$$

Für $t = \pi$ vereint $e^{i\pi} + 1 = 0$ viele wichtigen Konstanten der Analysis in einer einzigen Formel.

Beweis. Wir können auch komplexe Potenzreihen definieren und innerhalb ihres Konvergenzradiusses ableiten. Hier bedeutet der Ausdruck e^{it} daher, dass wir it in die Potenzreihe der Exponentialfunktion eingesetzt haben.

Nach Kettenregel folgt für $z(t) := e^{it}$

$$\dot z(t) = ie^{it} = iz(t).$$

Für $\zeta(t) := \cos t + i \sin t$ gilt ebenfalls $\dot{\zeta}(t) = -\sin t + i \cos t = i\zeta(t)$. Da die Anfangswerte $z(0) = 1 = \zeta(0)$ übereinstimmen, erhalten wir $z = \zeta$ und daher die Behauptung. $\square$

Do 07.05.2020

Hieraus ergibt sich die angekündigte Potenzreihenentwicklung für den Sinus und den Kosinus.

Proposition 4.61. *Es gelten für alle $x \in \mathbb{R}$*

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \ldots$$

und

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \ldots.$$

Insbesondere haben beide Potenzreihen den Konvergenzradius $+\infty$.

Beweis. Wir betrachten die Exponentialfunktion an der Stelle ix, $x \in \mathbb{R}$, und erhalten

$$e^{ix} = \sum_{n=0}^{\infty} \frac{i^n x^n}{n!}.$$

Da diese Reihe absolut konvergiert, können wir sie umordnen, nach Real- und Imaginärteil sortieren und erhalten

$$e^{ix} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

Die Eulersche Formel liefert nun die Behauptung. $\square$

Proposition 4.62. *Es gilt für $z \in \mathbb{C}$*

(i) $e^z = 1$ genau dann, wenn $z = 2\pi ik$ mit $k \in \mathbb{Z}$ und
(ii) $e^z = -1$ genau dann, wenn $z = i\pi + 2\pi ki$ mit $k \in \mathbb{Z}$.

Beweis. Wir schreiben $z = x + iy$ mit $x, y \in \mathbb{R}$. Dann folgt $e^z = e^x \cdot e^{iy}$ und $|e^z| = e^x \cdot \left|e^{iy}\right| = e^x$. Somit kann $e^z = \pm 1$ nur für $x = 0$ erfüllt sein. Nun gilt $e^{iy} = \cos y + i \sin y$. Da der Sinus und der Kosinus 2π-periodisch sind und $[0, 2\pi) \ni y \mapsto (\cos y, \sin y) \in \mathbb{S}^1$ bijektiv ist, kann es bis auf Vielfache von $2\pi i$ jeweils nur genau ein solches z geben. Da die angegebenen Zahlen z die Identitäten $e^z = \pm 1$ erfüllen, folgt die Behauptung. $\square$

Allgemeiner folgt direkt aus der Eulerschen Formel

Korollar 4.63. *Die Exponentialfunktion $\mathbb{C} \ni z \mapsto e^z \in \mathbb{C}$ ist $2\pi i$-periodisch.*

Definition 4.64.

(i) Für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ definieren wir den *Tangens,* $\tan x$, durch $\tan x = \frac{\sin x}{\cos x}$.

Bemerkung 4.65.

(i) Aus der Eulerschen Formel erhalten wir

$$\cos x = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \qquad \text{und} \qquad \sin x = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right).$$

Somit definieren wir die komplexen trigonometrischen Funktionen durch

$$\cos z = \frac{1}{2}\left(e^{iz} + e^{-iz}\right) \qquad \text{und} \qquad \sin z = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right).$$

Die Potenzreihendarstellungen stimmen mit den reellen Potenzreihendarstellungen überein. Additionstheoreme und Ableitungsregeln gelten genauso wie im Reellen, jedoch ist $|\sin z| \le 1$ nun im Allgemeinen falsch.

Beweis. Übung. □

Bemerkung 4.66 (Polarkoordinaten in $\mathbb{R}^2$)**.** Sei $x \in \mathbb{R}^2 \setminus \{0\}$. Dann gilt $x = |x| \cdot \frac{x}{|x|} \equiv r \cdot \xi$ für $r \ge 0$ und $\xi \in \mathbb{S}^1$. Somit lässt sich $\xi \in \mathbb{S}^1$ in eindeutiger Weise als $(\cos t, \sin t)$ mit $t \in (-\pi, \pi]$, oder, wenn wir $\mathbb{R}^2$ mit $\mathbb{C}$ identifizieren, als e^{it} darstellen. (r, t) heißen dann Polarkoordinaten des Vektors $x = re^{it}$.

t heißt auch Argument der komplexen Zahl $z = re^{it}$, $t = \arg z$. Die Argumentfunktion ist in $\mathbb{R}_{\le 0} \subset \mathbb{C}$ unstetig.

Lemma 4.67.

(i) Die Exponentialfunktion bildet den Streifen

$$S = \{z = x + iy \colon -\pi < y \le \pi\}$$

bijektiv auf $\mathbb{C} \setminus \{0\}$ ab.

(ii) Sei $\log\colon \mathbb{C} \setminus \{0\} \to S$ die Umkehrfunktion von $\exp\colon S \to \mathbb{C} \setminus \{0\}$, so gilt $\log' z = \frac{1}{z}$ für alle Punkte im Inneren des Streifens S.

Beweis.

(a) Wir schreiben Punkte in S in der Form $x+iy$ und sehen zunächst, dass $e^z = e^w$ nur bei gleichem Realteil von z und w möglich ist. Die Imaginärteile müssen ebenfalls übereinstimmen, da $(-\pi, \pi] \ni y \mapsto e^{iy} \in \mathbb{S}^1$ bijektiv ist.
(b) Die Surjektivität folgt aus der Darstellung in Polarkoordinaten.
(c) Die Ableitung ergibt sich wie im reellen Fall aus der Formel für die Ableitung der Umkehrfunktion. □

4.5. **Taylorsche Formeln.**

Bemerkung 4.68. Für Polynome

$$f(x) = \sum_{k=0}^{n} a_k (x - x_0)^k$$

mit $x \in \mathbb{R}$ und a_k aus einem Banachraum gilt $a_k = \frac{1}{k!} f^{(k)}(x_0)$ und somit

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0) \cdot (x - x_0)^k.$$

Taylorreihen verallgemeinern solch eine Darstellung für beliebige Funktionen.

Theorem 4.69. *Sei E ein reeller Banachraum, $I = (a, b)$ ein Intervall mit $0, 1 \in I$ und sei $f \in C^n(I, E)$ mit $n \ge 1$. Dann gilt*

$$f(1) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0) + R_n(f, 0)(1),$$

wobei sich das Restglied $R_n(f, 0)(1)$ in der Form

$$\|R_n(f, 0)(1)\| \le \frac{1}{(n-1)!} \sup_{0<t<1} \left\| f^{(n)}(t) - f^{(n)}(0) \right\|$$

abschätzen lässt.

Beweis. Wir müssen nur die Normabschätzung für das Restglied beweisen, da es solch ein Restglied nämlich stets gibt: Definiere es als die Differenz der anderen beiden Terme.

Wir definieren für $t \in I$ die Funktion

$$\varphi(t) = f(1) - \sum_{k=0}^{n-1} \frac{1}{k!} f^{(k)}(t) \cdot (1-t)^k - \frac{f^{(n)}(0)}{n!}(1-t)^n.$$

Man beachte die unterschiedlichen Argumente von f. Es gelten $\varphi(0) = R_n(f,0)(1)$ nach Definition des Restes sowie $\varphi(1) = 0$, da alle Terme mit $(1-t)^k$, $k \geq 1$, dort verschwinden und die restlichen beiden sich gerade aufheben. Für die Ableitung erhalten wir

$$\begin{aligned}
\varphi'(t) &= -\sum_{k=0}^{n-1} \frac{1}{k!} f^{(k+1)}(t) \cdot (1-t)^k + \sum_{k=1}^{n-1} \frac{k}{k!} f^{(k)}(t)(1-t)^{k-1} + \frac{f^{(n)}(0)}{(n-1)!}(1-t)^{n-1} \\
&= -\sum_{k=0}^{n-1} \frac{1}{k!} f^{(k+1)}(t) \cdot (1-t)^k + \sum_{k=0}^{n-2} \frac{1}{k!} f^{(k+1)}(t)(1-t)^{k} + \frac{f^{(n)}(0)}{(n-1)!}(1-t)^{n-1}
\end{aligned}$$

nach einer Indexverschiebung in der zweiten Summe

$$= \frac{1}{(n-1)!}\left(f^{(n)}(0) - f^{(n)}(t)\right) \cdot (1-t)^{n-1}.$$

Die Behauptung folgt nun aus dem Mittelwertsatz

$$\begin{aligned}
\|R_n(f,0)(1)\| = \|\varphi(1) - \varphi(0)\| &\leq \sup_{0<t<1} \|\varphi'(t)\| \\
&\leq \frac{1}{(n-1)!} \sup_{0<t<1} \left\| f^{(n)}(t) - f^{(n)}(0) \right\|. \qquad \square
\end{aligned}$$

Das folgende Theorem gilt auch noch für auf konvexen Teilmengen von $\mathbb{C}$ definierten Funktionen.

Theorem 4.70 (Taylorsche Formel)**.** *Sei $\Omega \subset \mathbb{R}$ ein offenes Intervall, sei E ein reeller Banachraum und sei $f \in C^n(\Omega, E)$ mit $n \geq 1$. Seien $x, x_0 \in \Omega$, dann lässt sich f in der Form*

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0) \cdot (x-x_0)^k + R_n(f,x_0)(x)$$

darstellen und es gilt die Restgliedabschätzung

$$\|R_n(f,x_0)(x)\| \leq \frac{1}{(n-1)!} \sup_{\xi \in (x_0,x)} \left\| f^{(n)}(\xi) - f^{(n)}(x_0) \right\| \cdot |x-x_0|^n.$$

Dabei bezeichne (x, x_0) die offene Verbindungsstrecke zwischen x und x_0, also das Intervall (x, x_0) oder das Intervall (x_0, x).

Beweis. Wir definieren

$$\varphi(t) := f(tx + (1-t)x_0)$$

für $0 \leq t \leq 1$. φ ist auch noch etwas über $[0,1]$ hinaus von der Klasse C^n. Daher können wir Theorem 4.69 anwenden und erhalten mit

$$\varphi^{(k)}(t) = f^{(k)}(tx + (1-t)x_0) \cdot (x-x_0)^k$$

sowie $R_n(\varphi, 0)(1) = R_n(f, x_0)(x)$ die Behauptung. $\square$

Mo 11.05.2020

Definition 4.71 (Taylorsche Polynome)**.** Sei E ein Banachraum über $\mathbb{K}$.

(i) Sei $\Omega \subset \mathbb{K}$ offen und $f \in C^n(\Omega, E)$, $n \geq 1$, so heißt $T_n(f, x_0) \in P(\mathbb{K}, E)$ mit

$$T_n(f, x_0)(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0) \cdot (x - x_0)^k$$

das *Taylorpolynom* n-ten Grades von f im Punkt x_0.

(ii) In Ω definieren wir das n-te *Restglied* $R_n(f, x_0)$ durch

$$R_n(f, x_0) := f - T_n(f, x_0).$$

(iii) Ist $f \in C^\infty(\Omega, E)$, so heißt die Potenzreihe

$$\left(\left(\frac{1}{n!} f^{(n)}(x_0) \cdot (x - x_0)^n\right)\right)_n$$

Taylorreihe von f im Punkt $x_0 \in \Omega$. Ihre n-te Partialsumme ist gerade $T_n(f, x_0)$. Für alle x, für die die Taylorreihe konvergiert, definieren wir

$$T(f, x_0)(x) := \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0) \cdot (x - x_0)^k.$$

Aus der Taylorschen Formel erhalten wir

Proposition 4.72. *Sei $\Omega \subset \mathbb{R}$ ein offenes Intervall, sei E ein reeller Banachraum, sei $f \in C^n(\Omega, E)$ mit $n \geq 1$ und sei $x_0 \in \Omega$. Dann gilt*

$$f(x) = T_n(f, x_0)(x) + o\left(|x - x_0|^n\right)$$

und $T_n(f, x_0)$ ist das einzige Polynom mit dieser Eigenschaft.

Beweis. Aus der Taylorschen Formel und der Stetigkeit der n-ten Ableitung von f erhalten wir, dass das Restglied von der Ordnung $o\left(|x - x_0|^n\right)$ ist.

Gäbe es ein weiteres Polynom, so liefert Differenzbildung ein Polynom vom Grade $\leq n$ in $o\left(|x - x_0|^n\right)$. Dies kann nur das Nullpolynom sein und wir erhalten die behauptete Eindeutigkeit. $\square$

Bemerkung 4.73.

(i) Sei $f \in C^\infty(\Omega, E)$. Dann gilt $f = T(f, x_0)$ in $B_\varepsilon(x_0)$ nach Definition genau dann, wenn die n-ten Restglieder $R_n(f, x_0)$ in $B_\varepsilon(x_0)$ für $n \to \infty$ punktweise gegen Null konvergieren, wenn also für alle $x \in B_\varepsilon(x_0)$

$$\lim_{n \to \infty} R_n(f, x_0)(x) = 0$$

gilt.

(ii) Gibt es zu $x_0 \in \Omega$ ein solches $\varepsilon > 0$ mit $\lim\limits_{n \to \infty} T_n(f, x_0)(x) = f(x)$ bzw., äquivalent dazu, $\lim\limits_{n \to \infty} R_n(f, x_0)(x) = 0$ für alle $x \in B_\varepsilon(x_0)$, so heißt f in x_0 *(reell) analytisch.* Ist $A \subset \Omega$ und f in jedem $x \in A$ analytisch, so schreiben wir dafür $f \in C^\omega(A, E)$.

(iii) In $B_{1/2}(0) \subset \mathbb{C}$ ist die Funktion $f(z) = \log(1 + z)$ durch ihre Taylorreihe mit Entwicklungspunkt im Ursprung darstellbar. Es gilt also

$$f(z) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} z^k.$$

Beweis. Es gilt $f \in C^\infty(B_1(0))$. Per Induktion erhalten wir

$$f^{(n)}(z) = (-1)^{n-1} \frac{(n-1)!}{(1+z)^n}$$

für alle $n \geq 1$.

Wir müssen also noch nachweisen, dass

$$\lim_{n\to\infty} R_n(f,0)(z) = 0$$

für alle $z \in B_{1/2}(0)$ gilt. Aus der Taylorschen Formel und unserer Darstellung für die Ableitung erhalten wir nacheinander

$$\begin{aligned} |R_n(f,0)(z)| &\le \sup_{|\zeta|<\frac{1}{2}} \left| \frac{1}{(1+\zeta)^n} - 1 \right| \cdot |z|^n, \\ \left|\frac{1}{1+\zeta}\right| &\le \frac{1}{1-|\zeta|} \le 2 \quad \text{für } |\zeta| \le \frac{1}{2}, \\ |R_n(f,0)(z)| &\le (1+2^n) \cdot |z|^n \end{aligned}$$

und wegen $|z| < \frac{1}{2}$ die Behauptung. □

(iv) Das folgende Beispiel zeigt, dass $T(f,x_0)$ zwar konvergieren kann, aber trotzdem nicht $f = T(f,x_0)$ zu gelten braucht.

Betrachte dazu die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x < 0, \\ 0, & x \ge 0. \end{cases}$$

Per Induktion überzeugt man sich direkt, dass $f \in C^\infty$ und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}$ gelten. Somit folgt $T(f,x_0) \equiv 0$, jedoch gilt in keiner Umgebung des Ursprunges $f \equiv 0$.

Wir leiten noch eine weitere Restgliedabschätzung her.

Theorem 4.74. *Sei E ein reeller Banachraum, sei $I \subset \mathbb{R}$ ein offenes Intervall mit $0, 1 \in I$, sei $f\colon I \to E$ eine $(n+1)$-mal differenzierbare Funktion mit $n \ge 0$. Dann gilt*

$$f(1) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0) + R_n(f,0)(1)$$

mit

$$\|R_n(f,0)(1)\| \le \frac{1}{(n+1)!} \sup_{0<t<1} \left\| f^{(n+1)}(t) \right\|.$$

Beweis. Wir definieren

$$\varphi(t) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(t) \cdot (1-t)^k \quad \text{und} \quad g(t) = (1-t)^{n+1}.$$

φ und g sind in I differenzierbar und es gilt

$$\begin{aligned} \varphi(0) &= T_n(f,0)(1), \quad \varphi(1) = f(1), \quad \varphi(1) - \varphi(0) = R_n(f,0)(1), \\ g(0) &= 1, \quad g(1) = 0, \\ \varphi'(t) &= \sum_{k=0}^{n} \frac{1}{k!} f^{(k+1)}(t) \cdot (1-t)^k - \sum_{k=1}^{n} \frac{k}{k!} f^{(k)}(t) \cdot (1-t)^{k-1}, \\ &= \frac{1}{n!} f^{(n+1)}(t) \cdot (1-t)^n, \\ g'(t) &= -(n+1)(1-t)^n \end{aligned}$$

Hierauf wenden wir nun den verallgemeinerten Mittelwertsatz, Theorem **??** (oder Theorem 4.28), auf dem Intervall $(0,1)$ an und erhalten

$$\|\varphi(1) - \varphi(0)\| \le \sup_{0<t<1} \left\| \frac{\varphi'(t)}{g'(t)} \right\| \cdot |g(1) - g(0)| \le \frac{1}{(n+1)!} \sup_{0<t<1} \left\| f^{(n+1)}(t) \right\|$$

wie behauptet. □

Theorem 4.75 (Lagrangesche Restgliedabschätzung)**.** *Sei $\Omega \subset \mathbb{K}$ offen und konvex. Sei E ein Banachraum über $\mathbb{K}$ und $f\colon \Omega \to E$ sei $(n+1)$-mal differenzierbar mit $n \geq 0$. Dann gilt*

$$f(x) = T_n(f, x_0)(x) + R_n(f, x_0)(x)$$

mit

$$\|R_n(f, x_0)(x)\| \leq \frac{1}{(n+1)!} \sup_{\xi \in (x_0, x)} \left\| f^{(n+1)}(\xi) \right\| \cdot |x - x_0|^{n+1}.$$

Beweis. Folge dem Beweis von Theorem 4.70 und benutze dabei Theorem 4.74 statt Theorem 4.69. □

5. Integration in einer Variablen

Wir werden hier einen ersten Integralbegriff kennenlernen, das Riemannintegral. Im dritten Semester werden wir das Lebesgueintegral behandeln, das es beispielsweise erlaubt, weitere Grenzwertsätze zu zeigen und mehr Funktionen zu integrieren.

Wenn man Integrale berechnet, verwendet man meist das Riemannsche Integral, erhält aber für das Lebesguesche Integral bessere allgemeine Sätze.

5.1. **Das Riemannsche Integral.** Sei $[a, b] \subset \mathbb{R}$ kompakt und sei $f\colon [a, b] \to \mathbb{R}$ reell, nicht-negativ und beschränkt. Dann wollen wir mit dem Integral über die Funktion f den Flächeninhalt der Menge

$$Q(f) := \left\{(x, y) \in \mathbb{R}^2 \colon a \leq x \leq b \text{ und } 0 \leq y \leq f(x)\right\}$$

in einer Weise beschreiben, dass für konstante Funktionen mit $f(x) \equiv h$ für alle x gerade die Rechtecksfläche $(b - a) \cdot h$ herauskommt.

Um Q in kleine Stücke aufteilen zu können, definieren wir

Definition 5.1 (Partition)**.** Sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall.

(i) Dann definiert $P = \{x_0, x_1, \ldots, x_n\}$, $n \in \mathbb{N}$, mit $a = x_0 < x_1 < \ldots < x_n = b$ eine *Partition* von I, die aus den Teilintervallen $I_i = [x_i, x_{i+1}]$ besteht.

Wir nennen P mit den angeordneten Punkten x_i und die Intervalle I_i eine Partition von I. Anders als bei einer Partition aus Analysis I lassen wir Überschneidungen an den endlich vielen Endpunkten der Intervalle zu. Die Punkte aus P heißen *Teil-* oder *Endpunkte*.

(ii) Sei $P = \{x_0, x_1, \ldots, x_n\}$ eine Partition. Dann heißt

$$\eta(P) := \sup_{0 \leq i \leq n-1} |x_{i+1} - x_i|$$

die *Feinheit* der Partition P.

(iii) Eine Partition P' heißt *feiner* als P, $P \prec P'$, falls jeder Endpunkt der Partition P auch Endpunkt von P' ist.

(iv) Seien P_1 und P_2 zwei Partitionen von I. Dann heißt die Partition $P_1 \vee P_2$, deren Menge der Endpunkte gerade die Vereinigung der Menge der Endpunkte von P_1 und P_2 ist, die gemeinsame *Verfeinerung* von P_1 und P_2.

Lemma 5.2. *Seien P_1 und P_2 Partitionen von $I = [a, b]$. Dann gilt*

$$P_1 \prec P_2 \quad \Longrightarrow \quad \eta(P_2) \leq \eta(P_1).$$

Do 14.05.2020

Beweis. Klar. □

Mit Partitionen können wir einen naheliegenden Flächeninhaltsbegriff motivieren.

Definiere $f_i := f|_{I_i}$, $\underline{Q}_i := I_i \times [0, \inf f_i]$ und $\overline{Q}_i := I_i \times [0, \sup f_i]$. Wegen

$$\bigcup_{i=0}^{n-1} \underline{Q}_i \subset Q \subset \bigcup_{i=0}^{n-1} \overline{Q}_i$$

ist der Flächeninhalt aller $\underline{Q}_i$'s addiert zu klein und der Flächeninhalt aller $\overline{Q}_i$'s addiert zu groß für einen sinnvollen Flächeninhaltsbegriff von Q: $|Q|$. Dass diese Rechtecke auf endlich vielen Linien überlappen trägt nichts zum Flächeninhalt bei. Wir fordern also

$$\sum_{i=0}^{n-1} |\underline{Q}_i| \equiv \sum_{i=0}^{n-1} \inf f_i \cdot |I_i| \le |Q| \le \sum_{i=0}^{n-1} |\overline{Q}_i| \equiv \sum_{i=0}^{n-1} \sup f_i \cdot |I_i|.$$

Der Ausdruck auf der linken Seite heißt *Untersumme,* der auf der rechten Seite *Obersumme.*

Wenn nun für eine beliebige Folge von Partitionen $(P_k)_k$ mit Feinheit $\eta(P_k) \to 0$ für $k \to \infty$ die Grenzwerte der linken und der rechten Seite existieren und übereinstimmen, so nennen wir dies den *Flächeninhalt* von Q.

Dies ist sicher nur möglich, falls es für jedes $\varepsilon > 0$ eine Partition P mit

$$\sum_{i=0}^{n-1} (\sup f_i - \inf f_i) \cdot |I_i| < \varepsilon$$

gibt.

Ersetzen wir $\sup f_i - \inf f_i$ durch die Oszillation von f_i, so können wir dieses Vorgehen auch auf banachraumwertige Funktionen verallgemeinern.

Definition 5.3. Sei $I = [a, b] \subset \mathbb{R}$ kompakt, sei E ein Banachraum und sei $f : I \to E$ beschränkt. Sei $P = \{x_0, x_1, \ldots, x_n\}$ eine Partition von I und seien I_i sowie f_i wie oben.

(i) Wir definieren die *Oszillation* von f als

$$\operatorname{osc} f := \sup_{x,\, y \in I} \|f(x) - f(y)\|.$$

(Für reellwertige Funktionen gilt $\operatorname{osc} f = \max\limits_{x \in I} f(x) - \min\limits_{y \in I} f(y)$.)

(ii) Multiplizieren wir die Oszillationen auf den Teilintervallen mit den Intervalllängen, so erhalten wir die Definition

$$\sigma(f, P) := \sum_{i=0}^{n-1} \operatorname{osc} f_i \cdot |I_i|.$$

(iii) Seien $\xi_i \in I_i$ beliebig und $Z := \{\xi_i : 0 \le i \le n-1\}$. Die Punkte ξ_i, $0 \le i \le n-1$, heißen *Zwischenpunkte* der Partition P. Dann definieren wir die *Riemannsche Summe* als

$$S(f, P, Z) \equiv S(P, Z) := \sum_{i=0}^{n-1} f(\xi_i) \cdot |I_i|.$$

Lemma 5.4. *Sei $f : I \to E$ beschränkt, seien P und P' zwei Partitionen von I mit $P \prec P'$ und habe P' genau λ Teilpunkte mehr als P. Dann gelten*

(i) $\sigma(f, P') \le \sigma(f, P)$ *und*

(ii) $|\sigma(f, P') - \sigma(f, P)| \le 2\lambda \|f\|_{L^\infty} \cdot \eta(P)$.

Beweis.

(i) Seien I_i, $0 \le i \le n-1$ die Teilintervalle der Partition P und I'_j, $0 \le j \le m-1$ die Teilintervalle von P'. Dann setzen sich die Intervalle I_i aus Intervallen I'_j zusammen und wir erhalten mit $f_i = f|_{I_i}$ und $f'_j = f|_{I'_j}$

$$\sigma(f, P') = \sum_{j=0}^{m-1} \operatorname{osc} f'_j \cdot |I'_j| = \sum_{i=0}^{n-1} \sum_{\{j :\, I'_j \subset I_i\}} \operatorname{osc} f'_j \cdot |I'_j|$$

$$\le \sum_{i=0} \operatorname{osc} f_i \sum_{\{j:\, I'_j \subset I_i\}} |I'_j| = \sum_{i=0}^{n-1} \operatorname{osc} f_i \cdot |I_i| = \sigma(f,P).$$

(ii) Wir führen den Beweis nur im Falle $\lambda = 1$. Der allgemeine Fall folgt daraus dann direkt per Induktion.

Sei dazu ohne Einschränkung I_1 ein Teilintervall, das sich aus den beiden Teilintervallen I'_1 und I'_2 (bis auf einen Punkt) disjunkt zusammensetzt. Alle anderen Teilintervalle stimmen überein. Dann folgt

$$\begin{aligned} 0 &\le \sigma(f,P) - \sigma(f,P') = \operatorname{osc} f_1 \cdot |I_1| - (\operatorname{osc} f'_1 \cdot |I'_1| + \operatorname{osc} f'_2 \cdot |I'_2|) \\ &\le \operatorname{osc} f_1 \cdot |I_1| \le 2 \cdot \|f\|_{L^\infty} \cdot |I_1| \le 2 \cdot \|f\|_{L^\infty} \cdot \eta(P). \end{aligned}$$

□

Lemma 5.5. *Seien $I = [a,b]$ und E wie bisher. Sei $f\colon I \to E$ beschränkt. Seien P und P' zwei Partitionen des Intervalles I mit Zwischenpunkten Z bzw. Z'.*

(i) Gilt $P \prec P'$, so folgt

$$\|S(P,Z) - S(P',Z')\| \le \sigma(f,P).$$

(ii) Sind P und P' beliebige Partitionen von I, so gilt

$$\|S(P,Z) - S(P',Z')\| \le \sigma(f,P) + \sigma(f,P').$$

Beweis.

(i) Seien I_i mit $0 \le i \le n-1$ die Teilintervalle von P und I'_j mit $0 \le j \le m-1$ die Teilintervalle von P'. Da sich die Intervalllängen aller Intervalle I'_j mit $I'_j \subset I_i$ zu $|I_i|$ aufaddieren, erhalten wir

$$\begin{aligned} &\left\| f(\xi_i) \cdot |I_i| - \sum_{\{j:\, I'_j \subset I_i\}} f\left(\xi'_j\right) \cdot \left|I'_j\right| \right\| \\ \le &\underbrace{\left\| f(\xi_i) \cdot |I_i| - \sum_{\{j:\, I'_j \subset I_i\}} f(\xi_i) \cdot \left|I'_j\right| \right\|}_{=0} + \sum_{\{j:\, I'_j \subset I_i\}} \underbrace{\left\| f(\xi_i) - f\left(\xi'_j\right) \right\|}_{\le \operatorname{osc} f_i} \cdot \left|I'_j\right| \\ \le &\operatorname{osc} f_i \cdot |I_i|. \end{aligned}$$

Diese Abschätzungen für die einzelnen Intervalle I_i addieren wir nun und erhalten

$$\begin{aligned} \|S(P,Z) - S(P',Z')\| &= \left\| \sum_{i=0}^{n-1} f(\xi_i)\, |I_i| - \sum_{i=0}^{n-1} \sum_{\{j:\, I'_j \subset I_i\}} f\left(\xi'_j\right) \cdot \left|I'_j\right| \right\| \\ &\le \sum_{i=0}^{n-1} \left\| f(\xi_i) \cdot |I_i| - \sum_{\{j:\, I'_j \subset I_i\}} f\left(\xi'_j\right) \cdot \left|I'_j\right| \right\| \\ &\le \sum_{i=0}^{n-1} \operatorname{osc} f_i \cdot |I_i| = \sigma(f,P). \end{aligned}$$

(ii) Sei $P^* = P \vee P'$ oder eine beliebige andere gemeinsame Verfeinerung von P und P' mit beliebigen Zwischenpunkten Z^*. Dann folgen $P \prec P^*$ und $P' \prec P^*$ und direkt nach der Dreiecksungleichung erhalten wir

$$\begin{aligned} &\|S(P,Z) - S(P',Z')\| \\ \le &\|S(P,Z) - S(P^*,Z^*)\| + \|S(P^*,Z^*) - S(P',Z')\| \\ \le &\sigma(f,P) + \sigma(f,P'). \end{aligned}$$

□

Lemma 5.6 (Riemannsches Integrabilitätskriterium)**.** *Sei $f\colon I = [a,b] \to E$ beschränkt. Dann sind die folgenden beiden Aussagen äquivalent:*

(i) $\underset{\varepsilon>0}{\forall}\ \underset{P}{\exists}\quad \sigma(f,P) < \varepsilon$ *und*

(ii) $\underset{\varepsilon>0}{\forall}\ \underset{\delta>0}{\exists}\ \underset{P}{\forall}\quad \eta(P) < \delta \quad\Longrightarrow\quad \sigma(f,P) < \varepsilon.$

Beweis.

- „(i)$\Longrightarrow$(ii)“: Sei $\varepsilon > 0$ beliebig und P eine Partition mit $\sigma(f,P) < \varepsilon$. Wir nehmen an, dass P gerade λ Teilpunkte besitzt und dürfen weiterhin ohne Einschränkung annehmen, dass $\|f\|_{L^\infty} \neq 0$ gilt. Wir wählen
$$\delta := \frac{\varepsilon}{2\lambda\|f\|_{L^\infty}}.$$
Sei nun P' eine beliebige Partition von I mit $\eta(P') < \delta$. Sei $P^* := P \vee P'$. Dann besitzt P^* maximal λ zusätzliche Teilpunkte gegenüber P'. Es folgt
$$\begin{aligned}\sigma(f,P') &\le \sigma(f,P^*) + |\sigma(f,P') - \sigma(f,P^*)|\\ &\le \sigma(f,P) + 2\lambda\|f\|_{L^\infty}\cdot\eta(P') \qquad \text{(nach Lemma 5.4)}\\ &\le \varepsilon + 2\lambda\|f\|_{L^\infty}\cdot\delta = \varepsilon+\varepsilon = 2\varepsilon.\end{aligned}$$
- „(ii)$\Longrightarrow$(i)“: Ist klar. □

Damit definieren wir uns nun die Menge der Riemann integrierbaren Funktionen.

Definition 5.7 (Riemannsches Integral)**.** Sei $I = [a,b] \subset \mathbb{R}$ kompakt und sei E ein Banachraum.

(i) Eine beschränkte Funktion $f\colon I \to E$ heißt *Riemann integrabel* oder *integrabel,* falls f eine der Bedingungen aus Lemma 5.6 erfüllt.
Wir sagen dann, dass f das *Riemannsche Integrabilitätskriterium* erfüllt.

(ii) Sei f Riemann integrabel und $(P_k)_k$ eine Folge von Partitionen von I mit Zwischenpunkten $(Z_k)_k$ und $\eta(P_k) \to 0$ für $k \to \infty$. Dann definieren wir das *(bestimmte) Riemannsche Integral* von f über I durch
$$\int\limits_a^b f \equiv \int\limits_a^b f(x)\,dx \equiv \int\limits_I f := \lim_{k\to\infty} S(P_k, Z_k).$$
(Der Zusatz „bestimmt“ deutet dabei an, dass wir das Integrationsintervall I festgelegt haben.)

Theorem 5.8. *Das Riemannsche Integral ist wohldefiniert.*

Beweis.

(i) Sei $\varepsilon > 0$. Wähle $\delta > 0$, so dass für jede Partition P mit $\eta(P) < \delta$ die Ungleichung $\sigma(f,P) < \varepsilon$ folgt. Seien P und P' zwei Partitionen mit $\eta(P) < \delta$ und $\eta(P') < \delta$ sowie Zwischenpunkten Z und Z'. Nach Lemma 5.5 folgt dann

$$\|S(P,Z) - S(P',Z')\| \le \sigma(f,P) + \sigma(f,P') < 2\varepsilon. \tag{5.1}$$

(ii) Sei $(P_k)_k$ eine beliebige Folge von Partitionen von I mit $\eta(P_k) \to 0$ für $k \to \infty$ und Zwischenpunkten Z_k. Dann erhalten wir aufgrund der obigen Überlegungen für fast alle k, l die Abschätzung $\|S(P_k,Z_k) - S(P_l,Z_l)\| < 2\varepsilon$. Somit ist $(S(P_k,Z_k))_k \subset E$ eine Cauchyfolge und $\lim\limits_{k\to\infty} S(P_k,Z_k)$ existiert.

(iii) Ein bekanntes Argument liefert, dass dieser Grenzwert nicht von der speziellen Wahl der Folge $(P_k)_k$ und ihren Zwischenpunkten Z_k, $k \in \mathbb{N}$, abhängt: Sei eine weitere Folge von Partitionen $(P'_k)_k$ und Zwischenpunkten $(Z'_k)_k$ gegeben. Dann können wir die obige Argumentation auch auf die Folge

$$P_0,\, P'_0,\, P_1,\, P'_1,\, P_2,\, P'_2,\, \ldots$$

und ihre zugehörigen Zwischenpunkte anwenden und erhalten eine Cauchyfolge mit Gliedern $S(P_k, Z_k)$ und $S(P'_k, Z'_k)$. Dies ist aber nur möglich, wenn die Grenzwerte von $S(P_k, Z_k)$ und $S(P'_k, Z'_k)$ übereinstimmen. Somit folgt die Behauptung. □

Wir halten gesondert fest, dass wir, durch Betrachtung einer Folge $(P'_k, Z'_k)_k$, im Grenzwert $k \to \infty$ aus Gleichung (5.1) das folgende Korollar erhalten.

Korollar 5.9. *Für beliebiges $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle Partitionen P mit $\eta(P) < \delta$ und Zwischenpunkten Z die Ungleichung*

$$\left\| \int_a^b f - S(P, Z) \right\| < \varepsilon$$

Mo 18.05.2020

gilt.

5.2. **Integrationsregeln.**

Proposition 5.10. *Sei $f\colon I = [a, b] \to E$ Riemann integrabel und sei $[\alpha, \beta] \subset [a, b]$. Dann ist $f|_{[\alpha,\beta]}\colon [\alpha, \beta] \to E$ ebenfalls Riemann integrabel.*

Beweis. Sei $\varepsilon > 0$ und sei P eine Partition von I mit $\sigma(f, P) < \varepsilon$. Verfeinere P so zu P', dass auch α und β Teilpunkte der Partition P' sind. Wir erhalten $\sigma(f, P') \le \sigma(f, P)$. Sei schließlich P'' die Partition von $[\alpha, \beta]$, die genau die Teilpunkte aus P' als Teilpunkte besitzt, die zu $[\alpha, \beta]$ gehören. Dann folgt $\sigma(f|_{[\alpha,\beta]}, P'') \le \sigma(f, P')$ und wir erhalten die Behauptung. □

Proposition 5.11. *Sei $c \in [a, b]$. Sei f (oder genauer dessen Einschränkung auf die entsprechenden Intervalle) über $[a, c]$ und über $[c, b]$ integrierbar. Dann ist f auch über $[a, b]$ integrierbar und es gilt*

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

Beweis.

(i) Seien P und P' Partitionen von $[a, c]$ bzw. $[c, b]$. Dann liefert die Vereinigung der Teilpunkte von P und P' eine Partition P'' von $[a, b]$ und es gilt
$$\sigma(f, P'') = \sigma(f, P) + \sigma(f, P').$$
Hieraus folgt, dass f auch auf $[a, b]$ das Riemannsche Integrabilitätskriterium erfüllt.

(ii) Seien Z und Z' beliebige Zwischenpunkte der Partitionen P bzw. P' und $Z'' = Z \cup Z'$. Wir erhalten
$$S(P'', Z'') = S(P, Z) + S(P', Z').$$
Betrachten wir immer feinere Partitionen und gehen zum Grenzwert über, so erhalten wir daraus gerade die obige Behauptung für die bestimmten Integrale.

□

Definition 5.12. Sei f über $[a, b]$ integrierbar, so definieren wir

$$\int_a^a f = 0 \quad \text{und} \quad \int_b^a f = -\int_a^b f.$$

Mit Proposition 5.11 erhalten wir daraus

Proposition 5.13. *Sei $f\colon I = [a,b] \to E$ integrierbar. Seien $a_i \in I$, $1 \le i \le n$ beliebig. Dann folgt*

$$\int\limits_{a_1}^{a_n} f = \sum_{i=1}^{n-1} \int\limits_{a_i}^{a_{i+1}} f.$$

Beweis. Dies folgt direkt per Induktion nach n. Details: Übung. □

Theorem 5.14 (Dreiecksungleichung)**.** *Sei $f\colon [a,b] \to E$ (Riemann) integrierbar. Dann ist auch $\|f\|\colon [a,b] \to \mathbb{R}$ (Riemann) integrierbar und es gilt*

$$\left\| \int\limits_a^b f \right\| \le \int\limits_a^b \|f\|.$$

Beweis.

(i) Nach Dreiecksungleichung gilt für alle x, $y \in [a,b]$
$$\big|\|f(x)\| - \|f(y)\|\big| \le \|f(x) - f(y)\|$$
Somit erhalten wir $\operatorname{osc}\|f_i\| \le \operatorname{osc} f_i$ für alle Intervalle $I_i \subset [a,b]$. Daher gilt auch $\sigma(\|f\|, P) \le \sigma(f, P)$ für beliebige Partitionen von $[a,b]$ und erhalten die Integrabilität.

(ii) Sei $\varepsilon > 0$. Da f und $\|f\|$ integrabel sind, gibt es ein $\delta > 0$, so dass für alle Partitionen P von $[a,b]$ mit $\eta(P) < \delta$ und beliebigen Zwischenpunkten die Abschätzungen
$$\left\| \int\limits_a^b f - S(f,P,Z) \right\| < \varepsilon \quad \text{und} \quad \left| \int\limits_a^b \|f\| - S(\|f\|,P,Z) \right| < \varepsilon$$
gelten.

Wir verwenden nun die üblichen Bezeichnungen. Aus
$$S(\|f\|,P,Z) = |S(\|f\|,P,Z)|$$
folgt im Limes $\left| \int\limits_a^b \|f\| \right| = \int\limits_a^b \|f\|$. Somit erhalten wir
$$\begin{aligned}
\left\| \int\limits_a^b f \right\| &\le \left\| \int\limits_a^b f - S(f,P,Z) \right\| + \|S(f,P,Z)\| \\
&\le \varepsilon + \left\| \sum_{i=0}^{n-1} f(\xi_i) \cdot |I_i| \right\| \\
&\le \varepsilon + \sum_{i=0}^{n-1} \|f(\xi_i)\| \cdot |I_i| \\
&= \varepsilon + S(\|f\|,P,Z) = \varepsilon + |S(\|f\|,P,Z)| \\
&\le \varepsilon + \left| S(\|f\|,P,Z) - \int\limits_a^b \|f\| \right| + \left| \int\limits_a^b \|f\| \right| \\
&\le 2\varepsilon + \int\limits_a^b \|f\|.
\end{aligned}$$

Mit $\varepsilon \searrow 0$ folgt die Behauptung. $\square$

Proposition 5.15 (Monotonie des Integrals)**.** *Seien $f,\, g\colon [a,b] \to \mathbb{R}$ Riemann integrabel. Gilt $f \le g$, so folgt*

$$\int\limits_a^b f \le \int\limits_a^b g.$$

Insbesondere gilt

$$-(b-a)\cdot \|f\|_{L^\infty} \le \int\limits_a^b f \le (b-a)\cdot \|f\|_{L^\infty}.$$

Beweis.

(i) Die Ungleichung $S(f,P,Z) \le S(g,P,Z)$ bleibt auch im Grenzwert erhalten.
(ii) Benutze $-\|f\|_{L^\infty} \le f \le \|f\|_{L^\infty}$ und $\int\limits_a^b c = (b-a)\cdot c$ für alle $c \in \mathbb{R}$. $\square$

Der erste Teil des folgenden Korollars wird auch als Mittelwertsatz bezeichnet.

Korollar 5.16 (Mittelwertsatz)**.** *Sei $I = [a,b]$ und seien $f,\, g\colon I \to \mathbb{R}$ Riemann integrabel. Gelten $0 \le g$ und $m \le f \le M$, so folgen*

$$m\int\limits_a^b g \le \int\limits_a^b fg \le M\int\limits_a^b g \qquad \text{und} \qquad m(b-a) \le \int\limits_a^b f \le M(b-a).$$

Proposition 5.17. *Sei $f\colon I = [a,b] \to \mathbb{R}^n$ mit $f = \left(f^i\right)_{1\le i\le n}$ gegeben. Dann ist f genau dann Riemann integrabel, wenn dies für die Komponentenfunktionen zutrifft und es gilt*

$$\int\limits_a^b f = \left(\int\limits_a^b f^i\right)_{1\le i\le n}.$$

Dieses Resultat gilt auch für Funktionen mit Werten in einem Produktraum $\prod\limits_{i=1}^n E_i$ von Banachräumen E_i, $1 \le i \le n$.

Beweis.

(i) Beschränktheit: Es ist klar, dass f genau dann beschränkt ist, wenn alle Komponenten f^i, $1 \le i \le n$, beschränkt sind.
(ii) Integrabilität: Aus der Abschätzung

$$\left|f^i(x) - f^i(y)\right| \le |f(x) - f(y)| \le \sum_{i=1}^n \left|f^i(x) - f^i(y)\right|$$

für alle $x,\, y \in I$ folgt

$$\sigma\left(f^i, P\right) \le \sigma(f,P) \le \sum_{i=1}^n \sigma\left(f^i, P\right)$$

und daher die Behauptung über die Integrabilität.
(iii) Wert des Integrals: Die Riemannsche Summe $S(f,P,Z)$ ist gleich dem Vektor, dessen Komponenten die Riemannschen Summen $S\left(f^i,P,Z\right)$ sind. Gehen wir nun zum Grenzwert über und berücksichtigen, dass im $\mathbb{R}^n$ der Grenzwert einer Folge gleich dem Vektor ist, dessen Komponenten die Grenzwerte der Komponentenfolgen sind, so erhalten wir die Behauptung. $\square$

Proposition 5.18. *Sei E ein Banachraum über dem Körper $\mathbb{K}$. Dann bilden die Riemann integrierbaren Funktionen $f\colon I = [a,b] \to E$ einen $\mathbb{K}$-Vektorraum: $R(I,E)$. Wir schreiben $R(I) \equiv R(I,\mathbb{R})$. Es gelten für alle f, $g \in R(I,E)$ und alle $\lambda \in \mathbb{K}$*

$$\int\limits_a^b (f+g) = \int\limits_a^b f + \int\limits_a^b g \qquad \textit{und} \qquad \int\limits_a^b \lambda f = \lambda f.$$

Dies bedeutet, dass das Integral mit $f \mapsto \int\limits_I f$ als Abbildung $R(I,E) \to E$ ein linearer Operator (= lineare Abbildung) ist.

Beweis. Seien f, $g \in R(I,E)$. Die Beschränktheit von $f+g$ bzw. λf ist klar. Sei $J \subset I$ beliebig. Aus

$$\operatorname{osc}(f|_J + g|_J) \le \operatorname{osc}(f|_J) + \operatorname{osc}(g|_J)$$

erhalten wir

$$\sigma(f+g,P) \le \sigma(f,P) + \sigma(g,P)$$

für beliebige Partitionen P von I. Somit ist $f+g$ wieder Riemann integrabel.

Der Rest des Beweises ist einfach: Übung. □

Mo 25.05.2020

Proposition 5.19. *Seien E, F Banachräume. Sei $I = [a,b]$ und sei $f \in R(I,E)$. Sei $\varphi\colon E \to F$ Lipschitzstetig auf $R(f)$. Dann gilt $\varphi \circ f \in R(I,F)$.*

Beweis.

(i) Beschränktheit: Wegen $f \in R(I,E)$ ist $R(f)$ beschränkt, d. h. es gibt $x_0 \in E$ und $r > 0$ mit $f(I) \subset B_r(x_0)$. Für beliebige $x \in B_r(x_0)$ folgt nun $\|\varphi(x) - \varphi(x_0)\| \le L \cdot \|x - x_0\| \le Lr$. Also ist $\varphi \circ f(I) \subset B_{Lr}(\varphi(x_0))$. Somit ist $\varphi \circ f$ beschränkt.

(ii) Integrierbarkeit: Sei L die Lipschitzkonstante von f auf $R(f)$. Dann folgt $\|\varphi(x) - \varphi(y)\| \le L\cdot\|x-y\|$ für alle x, $y \in R(f)$. Sei $J \subset I$ beliebig. Wir erhalten dann $\operatorname{osc}(\varphi \circ f|_J) \le L \cdot \operatorname{osc}(f|_J)$ und folgern, dass $\sigma(\varphi \circ f, P) \le L \cdot \sigma(f,P)$ für alle Partitionen P von I gilt. Wir erhalten die Behauptung. □

Korollar 5.20. *Sei $I = [a,b]$. Ist $f \in R(I,E)$ und ist $p \in [1,\infty)$, so folgt $\|f\|^p \in R(I,\mathbb{R})$.*

Proposition 5.21. *Sei $I = [a,b]$ und seien $f, g \in R(I,\mathbb{K})$, so folgen*

$$fg \in R(I,\mathbb{K}) \qquad \textit{und} \qquad \frac{f}{g} \in R(I,\mathbb{K}), \qquad \textit{falls } \inf_I |g| > 0.$$

Beweis.

(i) Die Beschänktheit von fg und f/g ist jeweils klar.

(ii) Sei $J \subset I$ beliebig und seien x, $y \in J$. Dann folgt aus

$$f(x)g(x) - f(y)g(y) = f(x)(g(x) - g(y)) + (f(x) - f(y))g(y)$$

die Oszillationsabschätzung

$$\operatorname{osc}((fg)|_J) \le \|f\|_{L^\infty} \cdot \operatorname{osc}(g|_J) + \operatorname{osc}(f|_J) \cdot \|g\|_{L^\infty}.$$

Somit ist $fg \in R(I,\mathbb{K})$.

(iii) Gehe analog vor. Übung. □

Proposition 5.22. *Sei $I = [a,b]$ und seien f, $g \in R(I,E)$. Stimmen f und g auf einer dichten Teilmenge von I überein, so gilt*

$$\int\limits_a^b f = \int\limits_a^b g.$$

Beweis. Wähle in den approximierenden Riemannschen Summen die Zwischenpunkte aus der dichten Teilmenge, in der f und g übereinstimmen. □

Integration von Folgen und Reihen. Das Riemannsche Integral verhält sich unter gleichmäßiger Konvergenz gutartig. Bei punktweiser Konvergenz werden wir im dritten Semester sehen, dass sich das Lebesguesche Integral besser verhält.

Theorem 5.23. *Sei $I = [a,b]$. Sei $(f_n)_n$ eine Folge von Riemann integrierbaren Funktionen $f_n \in R(I,E)$, die gleichmäßig gegen eine Funktion $f\colon I \to E$ konvergieren. Dann gilt auch $f \in R(I,E)$ und für das Integral erhalten wir*

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f.$$

Beweis.

(i) Beschränktheit: Aus
$$\|f(x)\| \le \|f_n(x)\| + \|f(x) - f_n(x)\| \le \|f\|_{L^\infty} + \|f - f_n\|_{L^\infty}$$
folgt die Beschränktheit von f.

(ii) Integrabilität: $f_n \rightrightarrows f$ ist äquivalent zu $\|f - f_n\|_{L^\infty} \to 0$ für $n \to \infty$. Sei $A \subset I$ beliebig und seien $x,\, y \in A$. Sei $n \in \mathbb{N}$ fest. Dann erhalten wir
$$\begin{aligned}\|f(x) - f(y)\| &\le \|f(x) - f_n(x)\| + \|f_n(x) - f_n(y)\| + \|f_n(y) - f(y)\|\\ &\le \|f - f_n\|_{L^\infty} + \operatorname{osc}(f_n|_A) + \|f_n - f\|_{L^\infty}\end{aligned}$$
und somit
$$\operatorname{osc}(f|_A) \le \operatorname{osc}(f_n|_A) + 2\|f - f_n\|_{L^\infty}$$
sowie für eine beliebige Partition P von I
$$\sigma(f,P) \le \sigma(f_n,P) + 2(b-a)\|f - f_n\|_{L^\infty}.$$
Für festes $\varepsilon > 0$ wählen wir nun zunächst n so groß, dass $2(b-a)\|f - f_n\|_{L^\infty} < \varepsilon$ gilt. Nun wählen wir die Partition P so, dass auch $\sigma(f_n,P) < \varepsilon$ ist. Somit gibt es eine Partition P von I mit $\sigma(f,P) < 2\varepsilon$ und f ist integrabel.

(iii) Wert des Integrals: Mit Dreiecksungleichung und der Monotonie des Integrals erhalten wir
$$\left\|\int_a^b f - \int_a^b f_n\right\|_E \le \int_a^b \|(f - f_n)(x)\|_E\, dx \le \int_a^b \|f - f_n\|_{L^\infty} \le (b-a)\cdot\|f - f_n\|_{L^\infty}.$$
Dies liefert die Behauptung. □

5.3. **Monotone und stetige Funktionen.** Bisher wissen wir nur, dass konstante Funktionen Riemann integrierbar sind. In diesem Kapitel werden wir sehen, dass auch alle monotonen Funktionen und alle stetigen Funktionen Riemann integrierbar sind.

Proposition 5.24. *Sei $I = [a,b]$. Sei $f\colon I \to E$ stetig, also $f \in C^0(I,E)$. Dann ist $f \in R(I,E)$.*

Beweis. Da I kompakt ist, ist f auf I gleichmäßig stetig. Zu $\varepsilon > 0$ gibt es also ein $\delta > 0$, so dass
$$|x - y| < \delta \quad\Longrightarrow\quad \|f(x) - f(y)\| < \varepsilon$$
für alle $x, y \in I$ gilt. Insbesondere gilt für alle Teilmengen $J \subset I$ mit $\operatorname{diam} J = \sup\limits_{x,\,y\in J} |x-y| < \delta$ die Abschätzung $\operatorname{osc}(f|_J) < \varepsilon$.

Sei nun P eine beliebige Partition von I mit $\eta(P) < \delta$. Dann erhalten wir

$$\sigma(f, P) < \varepsilon \cdot (b - a).$$

Die Beschränktheit von f ist klar. Somit ist $f \in R(I, E)$. □

Proposition 5.25. *Sei $f\colon I = [a, b] \to \mathbb{R}$ monoton, so ist f Riemann integrabel.*

Beweis. Wir dürfen ohne Einschränkung annehmen, dass f monoton wächst. Sei $J = [\alpha, \beta] \subset I$ ein beliebiges Teilintervall. Dann erhalten wir

$$\operatorname{osc}(f|_J) = f(\beta) - f(\alpha).$$

Wir zeigen nun das Riemannsche Integrabilitätskriterium mit Hilfe der äquidistanten Partition P_n, $n \in \mathbb{N}_{>0}$, von I, die durch $x_{i+1} - x_i = \frac{b-a}{n}$ für $0 \le i \le n - 1$ charakterisiert ist. Für die Partition P_n bekommen wir eine Teleskopsumme und erhalten

$$\sigma(f, P_n) = \frac{b-a}{n} \sum_{i=0}^{n-1} (f(x_{i+1}) - f(x_i)) = \frac{b-a}{n} (f(b) - f(a)) \to 0$$

für $n \to \infty$. Die Beschränktheit von f ist aufgrund der Monotonie klar. Somit ist $f \in R(I, \mathbb{R})$. □

Definition 5.26.

(i) Eine Funktion $f\colon I = [a, b] \to E$ heißt *Treppenfunktion,* falls es eine Partition $P = \{x_0, x_1, \ldots, x_n\}$ von I gibt, so dass f auf den Teilintervallen $I_i = [x_i, x_{i+1})$ konstant ist. Es gilt dann

$$f = \sum_{i=0}^{n-1} a_i \chi_{I_i}$$

für geeignete $a_i \in E$, $0 \le i \le n - 1$, (außer in x_n), wobei χ_A die *charakteristische Funktion* einer Menge A mit

$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A \end{cases}$$

bezeichnet. Wir sprechen auch von einer Treppenfunktion, wenn die Werte in den Teilpunkten x_i nicht zu dieser Definition passen.

(ii) Eine Funktion $f\colon I = [a, b] \to E$ heißt *stückweise stetig* bzw. *stückweise von der Klasse C^m*, $m \ge 0$, falls eine Partition P von I existiert, so dass f bzw. alle Ableitungen von f bis zur Ordnung m auf allen Intervallen $\mathring{I}_i = (x_i, x_{i+1})$, $0 \le i \le n - 1$, stetig sind und sich stetig auf jedes $\overline{I_i}$ fortsetzen lassen.

(iii) Stückweise affin lineare Funktionen, stückweise quadratische Funktionen sowie stückweise polynomiale Funktionen und stückweise konstante Funktionen (= Treppenfunktionen) sind analog definiert.

Proposition 5.27. *Stückweise stetige Funktionen und daher auch Treppenfunktionen sind Riemann integrabel.*

Beweis. Benutze, dass die Einschränkungen auf die Intervalle I_i zu einer geeigneten Partition stetig und daher auch Riemann integrabel sind. Nach Proposition 5.11 sind die Funktionen damit auch auf dem gesamten Definitionsbereich Riemann integrabel. □

5.4. **Hauptsatz der Differential- und Integralrechnung.** Für das Integral gilt das folgende Resultat, das gelegentlich auch als Mittelwertsatz bezeichnet wird.

Proposition 5.28 (Mittelwertsatz)**.** *Sei $I = [a, b]$ und sei $f \in C^0(I)$. Dann gibt es ein $\xi \in I$ mit*

$$\int_a^b f = f(\xi)(b-a).$$

Beweis. Nach Korollar 5.16 erhalten wir $(b-a)\inf f \le \int f \le (b-a)\sup f$. Jede Zahl in diesem Bereich lässt sich aber aufgrund des Zwischenwertsatzes für stetige Funktionen in der Form $f(\xi)\cdot(b-a)$ schreiben. □

Für vektorwertige Funktionen erhalten wir ebenfalls eine Variante des Mittelwertsatzes.

Proposition 5.29 (Mittelwertsatz)**.** *Sei $I = [a, b]$ und sei $f \in R(I, E)$, so gilt*

$$\left\| \int_a^b f \right\| \le \sup_{x \in I} \|f(x)\| \cdot (b-a) \equiv \|f\|_{L^\infty} \cdot (b-a).$$

Beweis. Nach Dreiecksungleichung gilt

$$\left\| \int_a^b f \right\| \le \int_a^b \|f\| \le \int_a^b \|f\|_{L^\infty} = \|f\|_{L^\infty} \cdot (b-a)$$

wie behauptet. □

Do 28.05.2020

Proposition 5.30. *Sei $I = [a, b]$ und sei $f \in R(I, E)$. Dann definieren wir eine Abbildung*

$$F\colon I \to E \qquad \text{durch} \qquad x \mapsto \int_a^x f.$$

Die Abbildung F ist Lipschitzstetig mit Lipschitzkonstante $\le \|f\|_{L^\infty}$.

Beweis. Seien $x, y \in I$ und gelte ohne Einschränkung $x < y$. Dann erhalten wir

$$\|F(y) - F(x)\| = \left\| \int_a^y f - \int_a^x f \right\| = \left\| \int_x^y f \right\| \le \int_x^y \|f\| \le \|f\|_{L^\infty} \cdot |y-x|$$

wie behauptet. □

Theorem 5.31 (Hauptsatz der Differential- und Integralrechnung)**.** *Sei $I = [a, b]$ und sei $f \in C^0(I, E)$. Dann ist die Funktion F aus Proposition 5.30 in (a, b) (stetig) differenzierbar und es gilt dort $F' = f$.*

Grundidee des Beweises ist, dass $\int\limits_{x_0}^{x_0+\varepsilon} f \approx f(x_0)\cdot\varepsilon$ gilt.

Beweis. Seien $x, x_0 \in (a, b)$. Dann folgt

$$\|F(x) - F(x_0) - f(x_0)\cdot(x-x_0)\| = \left\| \int_{x_0}^x (f - f(x_0)) \right\| \le \sup_{\xi \in [x_0, x]} \|f(\xi) - f(x_0)\| \cdot |x - x_0|.$$

Wenn wir nun durch $|x - x_0|$ dividieren, konvergiert die rechte Seite für $x \to x_0$ aufgrund der Stetigkeit von f immer noch gegen Null. Dies besagt aber nach Definition der Differenzierbarkeit gerade, dass F an der Stelle x_0 differenzierbar ist und die Ableitung $f(x_0)$ ist. Die Stetigkeit der Ableitung folgt aus der Stetigkeit von f. □

Korollar 5.32. *Sei $\Omega \subset \mathbb{R}$ ein offenes Intervall. Sei $f \in C^1(\Omega, E)$. Dann gilt*

$$f(x) = f(x_0) + \int\limits_{x_0}^{x} f'(t)\,dt$$

für alle x, $x_0 \in \Omega$.

Beweis. Sei $x_0 \in \Omega$ fest. Betrachte beide Seiten als Funktionen von x. Nach dem Hauptsatz der Differential- und Integralrechnung stimmen die Ableitungen der rechten und der linken Seite nach x in Ω überein. Somit ist die Differenz nach Korollar 4.24 eine Konstante in E. Für $x = x_0$ gilt Gleichheit, somit ist die Konstante $0 \in E$ und die Behauptung folgt. □

Aus Korollar 5.32 erhalten wir das folgende häufig verwendete Resultat.

Korollar 5.33. *Sei $\Omega \subset \mathbb{R}$ ein offenes Intervall. Sei $f \in C^1(\Omega, E)$. Dann gilt*

$$f(x) - f(x_0) = \int\limits_0^1 \frac{d}{dt} f(tx + (1-t)x_0)\,dt = \int\limits_0^1 Df(tx + (1-t)x_0)\langle x - x_0\rangle\,dt$$

für alle $x, x_0 \in \Omega$.

Beweis. Wir definieren $\varphi(t) := f(tx + (1-t)x_0)$ und erhalten aufgrund der Kettenregel $\varphi'(t) = Df(tx + (1-t)x_0)\langle x - x_0\rangle$. Das Resultat folgt nun aus Korollar 5.32. □

Definition 5.34. Sei $\Omega \subset \mathbb{R}$ eine offenes Intervall. Sei $f\colon \Omega \to E$ stetig. Eine Funktion $F\colon \Omega \to E$ mit $F' = f$ heißt *Stammfunktion* von f.

Bemerkung 5.35.

(i) Ist f stetig, so definiert $F(x) = \int\limits_{x_0}^{x} f$ eine Stammfunktion von f.

(ii) Ist F eine Stammfunktion von f und $c \in E$, so ist auch $x \mapsto F(x) + c$ eine Stammfunktion von f.

(iii) Seien F und G zwei Stammfunktionen von f. Dann stimmen F und G bis auf eine additive Konstante überein.

Beweis. Es gilt $F' = G'$. Somit folgt die Behauptung aus Korollar 4.24. □

(iv) Sei F eine Stammfunktion von f. Dann folgt nach Korollar 5.32 für beliebige x_1, $x_2 \in (a, b)$

$$\int\limits_{x_1}^{x_2} f = F(x_2) - F(x_1) \equiv F|_{x_1}^{x_2}.$$

(v) Als *unbestimmtes Integral* $\int f$ bezeichnen wir die Menge aller Stammfunktionen von f.

Als Vorbereitung auf das folgende Kapitel halten wir noch einen Satz über die Ableitung des Integrals in dem Fall, dass die obere Integrationsgrenze variabel ist, fest.

Proposition 5.36. *Sei $I = [a, b]$ und sei $f \in C^0(I, E)$. Sei $\Omega \subset \mathbb{R}$ offen und sei $\varphi \in C^1(\Omega, \mathbb{R})$ mit $\operatorname{im} \varphi \subset (a, b)$. Dann ist die Funktion*

$$G(x) := \int\limits_a^{\varphi(x)} f(t)\, dt$$

in Ω (stetig) differenzierbar und es gilt für $x_0 \in \Omega$

$$G'(x_0) = f(\varphi(x_0)) \cdot \varphi'(x_0).$$

Beweis. In der Notation des Beweises des Hauptsatzes der Differential- und Integralrechnung gilt $G(x) = F(\varphi(x))$. Daher folgt die Behauptung aus der Kettenregel.

Die Stetigkeit der Ableitung ist wieder klar. □

5.5. **Integralsätze und Transformationsregeln.**
Die partielle Integration ist eine integrierte Version der Produktregel.

Theorem 5.37 (Partielle Integration)**.** *Sei $I = [a, b]$ und seien f, $g \in C^1(I, \mathbb{K})$. Dann gilt*

$$\int\limits_{x_0}^{x} f'g = (fg)|_{x_0}^{x} - \int\limits_{x_0}^{x} fg'$$

für alle x, $x_0 \in I$.

Eine der beiden Funktionen kann auch banachraumwertig sein.

Beweis. In (a, b) gilt aufgrund der Produktregel $(fg)' = f'g + fg'$. Wir nehmen zunächst an, dass x, $x_0 \in (a, b)$ gelten. Wir integrieren dies $\int\limits_{x_0}^{x}$ und benutzen $\int\limits_{x_0}^{x} (fg)' = (fg)|_{x_0}^{x}$. Somit folgt die Behauptung in diesem Fall. Durch Grenzübergang erhalten wir aufgrund der Stetigkeit aller Summanden die Behauptung auch für beliebige x, $x_0 \in I$. □

Beispiele 5.38.

(i) Für a, $b > 0$ folgt mit $f'(t) = t$ und $g(t) = \log t$

$$\begin{aligned} \int\limits_a^b t \log t\, dt &= \frac{1}{2} t^2 \log t \bigg|_a^b - \int\limits_a^b \frac{1}{2} t^2 \frac{1}{t}\, dt \\ &= \left(\frac{1}{2} t^2 \log t - \frac{1}{4} t^2 \right) \bigg|_a^b. \end{aligned}$$

(ii) Für a, $b \in \mathbb{R}$ folgt mit $f'(t) = \sin t$ und $g(t) = t$

$$\begin{aligned} \int\limits_a^b t \sin t\, dt &= t(-\cos t)|_a^b - \int\limits_a^b 1 \cdot (-\cos t)\, dt \\ &= (-t \cos t + \sin t)|_a^b. \end{aligned}$$

Theorem 5.39 (Transformationsformel)**.** *Seien $I = [a, b]$, $J = [\alpha, \beta]$, $\varphi \in C^1(J)$ mit $\varphi(J) \subset I$, $a = \varphi(\alpha)$ und $b = \varphi(\beta)$. Sei $f \in C^0(I, E)$, so gilt*

$$\int\limits_a^b f(x)\, dx = \int\limits_\alpha^\beta f \circ \varphi(t) \varphi'(t)\, dt.$$

Die Transformationsformel ist letztlich eine integrierte Form der Kettenregel.

Beweis. Wir betrachten beide Seiten als Funktion der oberen Grenze β, also insbesondere auch $b = \varphi(\beta)$, benutzen γ statt β und $c = \varphi(\gamma)$ statt b, da die obere Grenze nun variabel sein soll und erhalten

$$\frac{d}{d\gamma}\int\limits_a^c f(x)\,dx = \frac{d}{dc}\int\limits_a^c f(x)\,dx \cdot \frac{d\varphi(\gamma)}{d\gamma} = f(c)\varphi'(\gamma)$$

sowie mit Proposition 5.36

$$\frac{d}{d\gamma}\int\limits_\alpha^\gamma f\circ\varphi(t)\varphi'(t)\,dt = f\circ\varphi(\gamma)\varphi'(\gamma).$$

Daher stimmen die Ableitungen beider Seiten überein.

Für $\gamma = \alpha$ sind beide Seiten Null. Somit folgt die Behauptung. □

Beispiel 5.40. Seien $f(t) = e^{-t}$ und $\varphi(x) = x^2$. Mit der Transformationsformel können wir nun das folgende in der Stochastik wichtige Integral bestimmen. Sei $a > 0$. Es gilt

$$\begin{aligned}\int\limits_0^a e^{-x^2} x\,dx &= \frac{1}{2}\int\limits_0^a f(\varphi(x))\varphi'(x)\,dx = \frac{1}{2}\int\limits_{\varphi(0)}^{\varphi(a)} e^{-t}\,dt\\ &= \frac{1}{2}\int\limits_0^{a^2} e^{-t}\,dt = -\frac{1}{2}e^{-t}\Big|_0^{a^2} = \frac{1}{2}\left(1-e^{-a^2}\right).\end{aligned}$$

5.6. **Uneigentliche Integrale.** In diesem Kapitel werden wir Funktionen integrieren, die am Rand des Definitionsintervalles „singulär“ werden.

Definition 5.41. Sei E ein metrischer Raum und $V \subset E$. Dann heißt $A \subset V$ *kompakt enthalten* in V, $A \Subset V$, falls $\overline{A}$ kompakt ist und $\overline{A} \subset V$ gilt.

Bemerkung 5.42.

(i) Ist $E = V$, so ist $A \Subset V$ gleichbedeutend zur relativen Kompaktheit von A in E.
(ii) Ist V offen in $\mathbb{R}^n$ und $A \Subset V$, so gilt insbesondere $\operatorname{dist}(\overline{A}, \partial V) > 0$.
(iii) Wir erinnern daran, dass $\nearrow$ bzw. $\searrow$ monotone Konvergenz bezeichnen.

Definition 5.43. Sei $(x_n)_n \subset \mathbb{R}$ eine Folge und $a \in \mathbb{R}$.

(i) Für $x_n \to a$ für $n \to \infty$ mit $x_n < a$ für alle $n \in \mathbb{N}$ schreiben wir $x_n \uparrow a$ für $n \to \infty$. Man findet auch $x_n \to a^-$.
(ii) Für $x_n \to a$ für $n \to \infty$ mit $x_n > a$ für alle $n \in \mathbb{N}$ schreiben wir $x_n \downarrow a$ für $n \to \infty$. Man findet auch $x_n \to a^+$.

Definition 5.44. Sei E ein Banachraum und sei $\emptyset \neq I \subset \mathbb{R}$ ein beliebiges nichtleeres Intervall.

(i) Eine Funktion $f\colon I \to E$ heißt *lokal integrierbar,* $f \in R_{\text{loc}}(I, E)$, falls $f|_J \in R(J, E)$ für alle Intervalle $J \Subset I$ gilt.
(ii) Sei $I = [a, b)$ und gelte $f \in R_{\text{loc}}(I, E)$. Dann sagen wir, dass das *uneigentliche Integral* $\int\limits_a^b f$ existiert oder konvergiert, falls

$$\lim_{\delta\downarrow 0}\int\limits_a^{b-\delta} f$$

existiert. Wir definieren dann $\int_a^b f$ als diesen Grenzwert und schreiben $f \in R(I,E)$.

Beispiel 5.45. Es gilt

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{\delta\downarrow 0}\int_\delta^1 \frac{1}{\sqrt{x}}\,dx = \lim_{\delta\downarrow 0} 2\sqrt{x}\Big|_\delta^1 = 2.$$

Do 04.06.2020

Ganz analog definieren wir für unbeschränkte Integrationsbereiche

Definition 5.46.

(i) Sei $I = [a,\infty)$ und $f \in R_{\text{loc}}(I,E)$. Dann sagen wir, dass das *uneigentliche Integral* $\int_0^\infty f$ existiert oder konvergiert, falls

$$\lim_{x\to\infty}\int_a^x f$$

existiert und definieren in diesem Fall $\int_a^\infty f$ als diesen Grenzwert.

Beispiel 5.47. Es gilt

$$\int_0^\infty e^{-t}\,dt = \lim_{x\to\infty}\int_0^x e^{-t} = \lim_{x\to\infty} -e^{-t}\Big|_0^x = 1.$$

Lemma 5.48. *Ist $f \in R_{loc}((a,b),\mathbb{R})$ mit $f \geq 0$ in (a,b), a, $b \in \mathbb{R}$ bzw. $(a,b) = (-\infty,\infty)$, so gelten*

$$\int_a^b f = \lim_{\delta\downarrow 0}\int_{a+\delta}^{b-\delta} f \qquad \textit{bzw.} \qquad \int_{-\infty}^\infty f = \lim_{x\to\infty}\int_{-x}^x f,$$

wobei möglicherweise beide Seiten gleich $+\infty$ sind.

Beweis. Übung. □

Bemerkung 5.49. Ist

$$f(x) = \begin{cases} +1, & x \geq 0, \\ -1, & x < 0, \end{cases}$$

so existiert $\int_{-\infty}^\infty f$ nicht, es gilt jedoch

$$\lim_{x\to\infty}\int_{-x}^x f = 0.$$

Definition 5.50.

(i) Sei $f \in R_{\text{loc}}((a,b),E)$, $a,b \in \mathbb{R}\cup\{\pm\infty\}$. Dann heißt das uneigentliche Integral $\int_a^b f$ *absolut konvergent* oder *absolut integrierbar,* falls $\int_a^b \|f\|$ konvergiert.

(ii) $f \in R_{\text{loc}}(\mathbb{R},E)$ heißt *integrierbar* (über $\mathbb{R}$), falls $\int_{-\infty}^\infty f$ existiert.

Da uneigentliche Integrale als Grenzwerte definiert sind, existieren sie genau dann, wenn die zugehörigen Cauchykriterien erfüllt sind. Im Fall $\int\limits_a^\infty f$ geben wir es hier an, in den anderen Fällen lauten sie analog.

Proposition 5.51 (Cauchykriterium)**.** *Das uneigentliche Integral* $\int\limits_a^\infty f$ *existiert genau denn, wenn*

$$\underset{\varepsilon>0}{\forall}\ \underset{R>a}{\exists}\ \underset{x,y\geq R}{\forall} \left\| \int_a^x f - \int_a^y f \right\| = \left\| \int_x^y f \right\| < \varepsilon$$

gilt.

Beweis. Dies folgt direkt nach Definition. □

Damit erhalten wir direkt

Korollar 5.52. *Ein absolut konvergentes Integral ist konvergent.*

Beweis. Dies folgt direkt aus dem Cauchykriterium und der folgenden Folgerung aus der Dreiecksungleichung

$$\left\| \int_x^y f \right\| \leq \left| \int_x^y \|f\| \right|.$$

□

Proposition 5.53 (Majorantenkriterium)**.**
Sei $I = [a, \infty)$ *und gelten* $g \in R_{loc}(I, \mathbb{R}_+)$, $f \in R_{loc}(I, E)$ *sowie* $\|f(x)\| \leq g(x)$ *für alle* $x \in I$*. Existiert* $\int\limits_a^\infty g$*, so ist auch* f *absolut integrierbar.*

Beweis. Für alle $x > a$ gilt

$$\int_a^x \|f\| \leq \int_a^x g \leq \int_a^\infty g.$$

Hieraus folgt die Behauptung aufgrund der Monotonie der linken Seite in x. □

Bemerkung 5.54. Ist das Integrationsintervall nichtkompakt, so können unerwartete Effekte auftraten. Auch bei gleichmäßiger Konvergenz gilt dann im Allgemeinen nicht mehr

$$\int_0^\infty \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_0^\infty f_n.$$

Dies sehen wir an dem folgenden Beispiel. Definiere für $n \in \mathbb{N}$ Funktionen $f_n \colon \mathbb{R}_+ \to \mathbb{R}$ durch

$$f_n(x) := \begin{cases} \frac{1}{n}, & n \leq x \leq 2n, \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt einerseits $f_n \rightrightarrows 0$, andererseits gilt aber auch $\int\limits_0^\infty f_n = 1 \not\to 0$ für $n \to \infty$.

Ist der Grenzwert $\lim\limits_{x\to\infty} \int\limits_a^x f_n$ in der Definition des uneigentlichen Integrales jedoch gleichmäßig in n, so erhalten wir

Theorem 5.55. *Sei $I = [a, \infty)$ und sei $f_n \in R_{loc}(I, E)$ eine Folge von Funktionen, die lokal gleichmäßig (und daher auch kompakt, d. h. gleichmäßig auf kompakten Teilmengen) gegen eine Funktion $f\colon I \to E$ konvergieren. Konvergieren die uneigentlichen Integrale $\int\limits_a^\infty f_n$ gleichmäßig, so ist f lokal Riemann integrabel, $\int\limits_a^\infty f$ existiert und es gilt*

$$\lim_{n\to\infty} \int\limits_a^\infty f_n = \int\limits_a^\infty f.$$

Beweis.

(i) Die Behauptung $f \in R_{\text{loc}}(I, E)$ folgt direkt aus Theorem 5.23.

(ii) Wir schreiben die gleichmäßige Konvergenz der Integrale explizit als gleichmäßige Cauchybedingung auf. Es gilt

$$\underset{\varepsilon>0}{\forall}\ \underset{R>a}{\exists}\ \underset{x,\,y\ge R}{\forall}\ \underset{n\in\mathbb{N}}{\forall} \quad \left\| \int\limits_x^y f_n \right\| < \varepsilon.$$

Mit Theorem 5.23 können wir zum Grenzwert $n \to \infty$ übergehen und erhalten für das ε und das R von oben

$$\left\| \int\limits_x^y f \right\| \le \varepsilon \qquad \text{für alle } x,\, y \ge R.$$

Diese Cauchybedingung besagt aber gerade, dass $\int\limits_a^\infty f$ konvergiert.

(iii) Zur Behauptung über den Wert des Integrals: Sei $\varepsilon > 0$. Dann gibt es aufgrund der obigen Integralabschätzungen für f und f_n ein $R > a$ mit

$$\left\| \int\limits_R^\infty f_n \right\| + \left\| \int\limits_R^\infty f \right\| \le 2\varepsilon$$

für alle $n \in \mathbb{N}$, da wir in den obigen Abschätzungen nun zum Grenzwert $y \to \infty$ übergehen dürfen.

Da $f_n \rightrightarrows f$ auf $[a, R]$ gilt, gibt es $n_0 \in \mathbb{N}$, so dass für alle $n \ge n_0$

$$\left\| \int\limits_a^R (f - f_n) \right\| \le \varepsilon$$

gilt. Zusammengenommen erhalten wir somit

$$\begin{aligned} \left\| \int\limits_a^\infty (f - f_n) \right\| &\le \left\| \int\limits_a^R (f - f_n) \right\| + \left\| \int\limits_R^\infty (f - f_n) \right\| \\ &\le \varepsilon + \left\| \int\limits_R^\infty f \right\| + \left\| \int\limits_R^\infty f_n \right\| \le 3\varepsilon. \end{aligned}$$

Die Behauptung folgt. □

Bemerkung 5.56. Eine Folge von uneigentlichen Integralen $\int\limits_a^\infty f_n$ ist gleichmäßig konvergent, falls es eine Funktion $g \in R_{\text{loc}}([a, \infty), \mathbb{R})$ mit $\int\limits_a^\infty g < \infty$ und $\|f_n(x)\| \le g(x)$ für alle $n \in \mathbb{N}$ und alle $x \in [a, \infty)$ gibt.

Beachte, dass somit automatisch $g(x) \geq 0$ für alle $x \in [a, \infty)$ gilt.

Beweis. Übung. □

5.7. **Parameterabhängige Integrale.** Sei $f\colon I \times \Omega \to E$ eine Funktion. Dabei ist $I = [a, b]$ ein kompaktes Intervall, Ω ein metrischer Raum, eine offene Teilmenge von $\mathbb{K}$ oder ein Intervall, dann auch mit $J = [\alpha, \beta]$ bezeichnet, und E ein Banachraum.

Proposition 5.57. *Sei Ω ein metrischer Raum und $f \in C^0(I \times \Omega, E)$. Dann ist*

$$F(x) := \int\limits_a^b f(t, x)\, dt$$

in Ω stetig.

Beweis.

(i) Sei $x_n \to x_0$ eine konvergente Folge in Ω. Dann ist $K := \{x_i : i \in \mathbb{N}\} \ni x_0$ kompakt. Die Einschränkung von f auf die kompakte Menge $I \times K$ ist somit gleichmäßig stetig. Sei nun f ohne Einschränkung auf $I \times K$ definiert und daher gleichmäßig stetig.

(ii) Wir definieren nun Funktionen $f_n := f(\cdot, x_0)$ und $g := g(\cdot, x_0)$. Aufgrund der gleichmäßigen Stetigkeit folgt $f_n \rightrightarrows g$ für $n \to \infty$. Da wir den Limes bei gleichmäßiger Konvergenz mit dem Integral vertauschen dürfen, erhalten wir

$$\int\limits_a^b f(t, x_0)\, dt = \int\limits_a^b g = \lim_{n\to\infty} \int\limits_a^b f_n = \lim_{n\to\infty} \int\limits_a^b f(t, x_n)\, dt$$

wie behauptet. □

Eine Definition, die uns auch in Kapitel 6 noch intensiver beschäftigen wird, ist

Definition 5.58 (Partielle Ableitungen)**.**

(i) Seien $\Omega_i \subset \mathbb{R}$, $1 \leq i \leq n$, offene nichtleere Intervalle. Sei $\Omega := \prod\limits_{i=1}^n \Omega_i \subset \mathbb{R}^n$ und sei $f\colon \Omega \to E$ eine Funktion. Dann heißt die Funktion f im Punkt $x_0 = \left(x_0^i\right)$ nach x^k oder in Richtung k *partiell differenzierbar,* falls die Funktion

$$g\colon \Omega_k \to E \quad \text{mit} \quad g(t) := f\left(x_0^1, \ldots, x_0^{k-1}, t, x_0^{k+1}, \ldots, x_0^n\right) \equiv f\left(\hat{x}_0^k, t\right)$$

im Punkt $t = x_0^k \in \Omega_k$ differenzierbar ist. Wir bezeichnen die Ableitung von g mit $\frac{\partial f}{\partial x^k}$, $\frac{\partial}{\partial x^k} f$ oder $D_k f$.

(ii) Ist f in ganz Ω partiell nach x^k differenzierbar, so heißt f in Ω nach x^k differenzierbar. Wir erhalten in diesem Fall eine Abbildung $D_k f\colon \Omega \to E$.

(iii) Ist die *partielle Ableitung* $D_k f$ stetig, so heißt f *stetig partiell* nach x^k *differenzierbar.*

(iv) f heißt *partiell differenzierbar,* falls alle Ableitungen $D_k f$, $1 \leq k \leq n$, existieren und *stetig partiell differenzierbar,* falls diese zusätzlich stetig sind.

(v) Ist $D_k f$ wieder partiell differenzierbar, so schreiben wir für die entsprechenden partiellen Ableitungen $D_k D_k f = \frac{\partial^2 f}{(\partial x^k)^2}$, $D_l D_k f = \frac{\partial}{\partial x^l} \frac{\partial}{\partial x^k} f = \frac{\partial^2 f}{\partial x^l \partial x^k}$ oder entsprechende Ausdrücke für höhere Ableitungen.

Mo 08.06.2020

Theorem 5.59. *Sei $\Omega \subset \mathbb{K}$ nichtleer und offen, seien f, $D_2 f \equiv \frac{\partial f}{\partial x} \in C^0(I \times \Omega, E)$. Dann ist*

$$F(x) = \int\limits_a^b f(t, x)\, dt$$

stetig differenzierbar und es gilt

$$F'(x) = \int\limits_a^b \frac{\partial f}{\partial x}(t,x)\,dt.$$

Beweis. Sei $x_0 \in \Omega$. Sei $\delta > 0$, so dass $B_\delta(x_0) \subset \Omega$ gilt. Dann erhalten wir für $x \in B_\delta(x_0)$

$$\frac{F(x)-F(x_0)}{x-x_0} - \int\limits_a^b \frac{\partial f}{\partial x}(t,x_0)\,dt = \int\limits_a^b \frac{f(t,x)-f(t,x_0)}{x-x_0} - \frac{\partial f}{\partial x}(t,x_0)\,dt \equiv I.$$

Wir schätzen dies mit der Dreiecksungleichung für Integrale und dem Mittelwertsatz in der Form von Korollar 4.23 ab und erhalten

$$\begin{aligned}\|I\| &\le \int\limits_a^b \left\| f(t,x) - f(t,x_0) - \frac{\partial f}{\partial x}(t,x_0)\cdot(x-x_0)\right\| dt\, \frac{1}{|x-x_0|}\\ &\le \int\limits_a^b \sup_{\xi\in[x_0,x]} \left\|\frac{\partial f}{\partial x}(t,\xi) - \frac{\partial f}{\partial x}(t,x_0)\right\|,\end{aligned}$$

wobei das ξ von t abhängen darf. Aufgrund der Stetigkeit von $\frac{\partial f}{\partial x}$, konvergiert die rechte Seite für $x \to x_0$ gegen Null und wir erhalten aus $0 = \lim\limits_{x\to x_0} I$ die Behauptung.

Die Stetigkeit der Ableitung folgt nun aus Proposition 5.57. □

Im Fall, dass auch noch die Grenzen des Intervalles von x abhängen, erhalten wir

Korollar 5.60. *Sei $\Omega \subset \mathbb{R}$ nichtleer und offen, seien $f, \frac{\partial f}{\partial x} \in C^0(I\times\Omega, E)$. Seien $\varphi, \psi\colon \Omega \to \mathring{I}$ in $x_0 \in \Omega$ differenzierbar, so ist*

$$F(x) = \int\limits_{\psi(x)}^{\varphi(x)} f(t,x)\,dt$$

in x_0 differenzierbar und es gilt

$$F'(x_0) = f(\varphi(x_0),x_0)\cdot\varphi'(x_0) - f(\psi(x_0),x_0)\cdot\psi'(x_0) + \int\limits_{\psi(x_0)}^{\varphi(x_0)} \frac{\partial f}{\partial x}(t,x_0)\,dt.$$

Beweis.

(i) Gelte ohne Einschränkung $0 \in \Omega$. Vermöge

$$F(x) = \int\limits_{\psi(x)}^{\varphi(x)} f(t,x)\,dt = \int\limits_0^{\varphi(x)} f(t,x)\,dt - \int\limits_0^{\psi(x)} f(t,x)\,dt$$

genügt es, die Behauptung im Fall $\psi(x) \equiv 0$ zu zeigen.

(ii) Es gilt

$$\begin{aligned}&F(x)-F(x_0)\\ &= \int\limits_0^{\varphi(x)} f(t,x)\,dt - \int\limits_0^{\varphi(x_0)} f(t,x_0)\,dt\end{aligned}$$

$$= \int\limits_0^{\varphi(x)} f(t,x) - f(t,x_0)\,dt + \int\limits_0^{\varphi(x)} f(t,x_0)\,dt - \int\limits_0^{\varphi(x_0)} f(t,x_0)\,dt$$

$$= \int\limits_0^{\varphi(x_0)} f(t,x) - f(t,x_0)\,dt + \int\limits_{\varphi(x_0)}^{\varphi(x)} f(t,x) - f(t,x_0)\,dt + \int\limits_{\varphi(x_0)}^{\varphi(x)} f(t,x_0)\,dt.$$

Wir dividieren nun durch $x - x_0$ und lassen $x \to x_0$. Dann konvergiert der erste Term auf der rechten Seite wegen Theorem 5.59 gegen den Integralterm aus der Behauptung und der dritte Term konvergiert nach Proposition 5.36 gegen den ersten Term auf der rechten Seite in der Behauptung.

(iii) Es genügt also, die folgende Behauptung zu zeigen:

$$0 \overset{!}{=} \lim_{x\to x_0} \left\| \frac{1}{x - x_0} \int\limits_{\varphi(x_0)}^{\varphi(x)} f(t,x) - f(t,x_0)\,dt \right\| \equiv \lim_{x\to x_0} I.$$

Es gilt aufgrund des vektorwertigen Mittelwertsatzes

$$\begin{aligned} I &\le \left| \int\limits_{\varphi(x_0)}^{\varphi(x)} \left\| \frac{f(t,x) - f(t,x_0)}{x - x_0} \right\| dt \right| \\ &\le |\varphi(x) - \varphi(x_0)| \cdot \sup_{t\in[\varphi(x),\varphi(x_0)]} \left\| \frac{f(t,x) - f(t,x_0)}{x - x_0} \right\| \\ &\le |\varphi(x) - \varphi(x_0)| \cdot \sup_{t\in[\varphi(x),\varphi(x_0)]} \cdot \sup_{\xi\in[x,x_0]} \left\| \frac{\partial f}{\partial x}(t,\xi) \right\|. \end{aligned}$$

Aufgrund der Stetigkeit von $\frac{\partial f}{\partial x}$ ist der zweite Faktor beschränkt. Aufgrund der Stetigkeit von φ konvergiert der erste Faktor für $x \to x_0$ gegen Null. Somit folgt die Behauptung. □

Theorem 5.61 (Doppelintegrale)**.** *Seien $I = [a,b]$ und $J = [\alpha,\beta]$ kompakte Intervalle und $f \in C^0(I \times J, E)$, so gilt*

$$\int\limits_\alpha^\beta \left(\int\limits_a^b f(t,x)\,dt \right) dx = \int\limits_a^b \left(\int\limits_\alpha^\beta f(t,x)\,dx \right) dt,$$

d. h. es kommt nicht auf die Integrationsreihenfolge an.

Beweis. Nach Proposition 5.57 ist das Ergebnis nach der jeweils inneren Integration eine stetige und daher auch integrierbare Funktion. Somit sind die Doppelintegrale auf beiden Seiten wohldefiniert.

Wir definieren nun für $y \in [\alpha,\beta]$ die Funktionen

$$\varphi(y) := \int\limits_\alpha^y \left(\int\limits_a^b f(t,x)\,dt \right) dx \quad \text{und} \quad \psi(y) := \int\limits_a^b \left(\int\limits_\alpha^y f(t,x)\,dx \right) dt.$$

Es gilt $\varphi(\alpha) = 0 = \psi(\alpha)$. Weiterhin sind φ und ψ in (α,β) nach Theorem 5.59 (ψ) (die dafür nötigen Voraussetzungen wollen wir nachfolgend prüfen), Proposition 5.57 (φ und ψ) und dem Hauptsatz der Differential- und Integralrechnung (φ und ψ) differenzierbar und es gilt dort

$$\varphi'(y) = \int\limits_a^b f(t,y)\,dt = \psi'(y).$$

Somit stimmen φ und ψ überein.

Für die Anwendbarkeit von Theorem 5.59 müssen wir noch die Stetigkeit von $g\colon I \times J \to E$ mit

$$I \times J \ni (t, y) \mapsto \int\limits_{\alpha}^{y} f(t, x)\, dx$$

und die Stetigkeit von $D_2 g = f$, die klar ist, überprüfen. Sei $(t_0, y_0) \in I \times J$ und $(t, y) \in I \times J$ mit $|t - t_0| + |y - y_0| < \delta$. Sei $\varepsilon > 0$. Wir wollen zeigen, dass für hinreichend kleines $\delta > 0$ auch $\|g(t, y) - g(t_0, y_0)\| < \varepsilon$ gilt. Da $t \mapsto g(t, y_0)$ nach Proposition 5.57 stetig ist, gibt es ein $\delta > 0$, so dass $\|g(t, y_0) - g(t_0, y_0)\| < \varepsilon$ gilt. Wir erhalten

$$\begin{aligned}\|g(t, y) - g(t_0, y_0)\| &\le \|g(t, y) - g(t, y_0)\| + \|g(t, y_0) - g(t_0, y_0)\| \\ &\le \left\| \int\limits_{y}^{y_0} f(t, x)\, dx \right\| + \varepsilon \le |y - y_0| \cdot \|f\|_{L^\infty} + \varepsilon < 2\varepsilon\end{aligned}$$

für $\delta < \frac{\varepsilon}{1 + \|f\|_{L^\infty}}$. Da $\varepsilon > 0$ beliebig war, folgt hieraus die Behauptung. $\square$

6. Differentiation in Banachräumen

Sie sollten versuchen, die Resultate in diesem Abschnitt zumindest in dem Fall zu verstehen, dass die Banachräume euklidische Räume $\mathbb{R}^n$, $n \in \mathbb{N}$, sind.

6.1. **Differenzierbarkeit.**

Definition 6.1.

(i) Seien E, F Banachräume. Sei $\Omega \subset E$ offen. Dann heißt eine Abbildung $f\colon \Omega \to F$ in $x_0 \in \Omega$ *differenzierbar,* falls es eine stetige lineare Abbildung $A \in L(E, F)$ mit
$$f(x) = f(x_0) + A\langle x - x_0 \rangle + o(\|x - x_0\|)$$
für alle $x \in \Omega$ gibt.

(ii) Wir nennen dann A die *Ableitung* von f im Punkt x_0 und bezeichnen sie mit $Df(x_0)$, $df(x_0)$ oder $f'(x_0)$.

(iii) Ist f in jedem Punkt in Ω differenzierbar, so heißt f in Ω *differenzierbar.* Wir nennen dann die Abbildung
$$\begin{aligned} Df\colon \Omega &\to L(E, F), \\ x &\mapsto Df(x) \end{aligned}$$
Ableitung oder *Differential* von f.

(iv) Ist $Df\colon \Omega \to L(E, F)$ stetig, so heißt f in Ω *stetig differenzierbar.*

Bemerkung 6.2. Seien E, F Banachräume, $\Omega \subset E$ offen und $f\colon \Omega \to F$ in $x_0 \in \Omega$ differenzierbar.

(i) Ist $\dim E < \infty$, so ist jede lineare Abbildung $A\colon E \to F$ automatisch stetig, siehe Analysis I.

(ii) f ist im Punkt x_0 stetig.

(iii) Die Ableitung $Df(x_0)$ ist eindeutig bestimmt.

(iv) Ist $E = \mathbb{K}$, so stimmt diese Definition mit der Definition für Funktionen in einer Variablen überein.

Bemerkung 6.3. Seien E, F Banachräume und sei $\Omega \subset E$ offen. Seien f, $g\colon \Omega \to F$ in $x_0 \in \Omega$ differenzierbar. Dann sind auch $f + g$ und λf für beliebige $\lambda \in \mathbb{K}$ in x_0 differenzierbar und es gelten

$$D(f + g)(x_0) = Df(x_0) + Dg(x_0) \quad \text{und} \quad D(\lambda f)(x_0) = \lambda Df(x_0).$$

Somit bilden die in einer vorgegebenen Menge $V \subset \Omega$, z. B. $V = \{x_0\}$, differenzierbaren Funktionen einen Unterraum des Vektorraumes aller Abbildungen $\Omega \to F$.

Beweis. Klar. □

Mo 15.06.2020

Beispiele 6.4.

(i) Konkretes Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ mit
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 \\ x^2y \\ y^3 \end{pmatrix}.$$
Wir benutzen die Definition der Ableitung
$$f((x,y)+(u,v)) - f((x,y)) = Df((x,y))\langle (u,v)\rangle + o(\|(u,v)\|)$$
und erhalten
$$\begin{aligned} & f\left(\begin{pmatrix} x+u \\ y+v \end{pmatrix}\right) - f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \\ = & \begin{pmatrix} (x+u)^2 - x^2 \\ (x+u)^2(y+v) - x^2y \\ (y+v)^3 - y^3 \end{pmatrix} \\ = & \begin{pmatrix} x^2 + 2xu + u^2 - x^2 \\ x^2y + 2xuy + u^2y + x^2v + 2xuv + u^2v - x^2y \\ y^3 + 3y^2v + 3yv^2 + v^3 - y^3 \end{pmatrix} \\ = & \begin{pmatrix} 2xu \\ 2xyu + x^2v \\ 3y^2v \end{pmatrix} + o(\|(u,v)\|), \end{aligned}$$
da beispielsweise $u^2 \in o(\|(u,v)\|)$ wegen
$$\left|\frac{u^2}{\|(u,v)\|}\right| = \left|\frac{u^2}{\max\{|u|,|v|\}}\right| = \left|\frac{u}{\max\{|u|,|v|\}}\right| \cdot |u| \le |u|$$
gilt.

Wir bemerken, dass wir unter Verwendung einer äquivalenten Norm auf $\mathbb{R}^2$ dasselbe Ergebnis erhalten.

Somit gilt
$$Df((x,y))\langle (u,v)\rangle = \begin{pmatrix} 2x & 0 \\ 2xy & x^2 \\ 0 & 3y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$
Es ist kein Zufall, dass in dieser 3×2-Matrix gerade die partiellen Ableitungen der Komponenten stehen. Wir werden dies später allgemeiner zeigen.

(ii) Sei $f\colon \Omega \to F$ konstant. Dann ist f in jedem Punkt aus Ω differenzierbar und es gilt $Df(x) = 0 \in L(E,F)$ für alle $x \in \Omega$.

(iii) Ist $f = A \in L(E,F)$, dann ist f differenzierbar und es gilt $Df(x) = A$ für alle $x \in E$, da
$$Ax = Ax_0 + A\langle x - x_0\rangle + 0$$
gilt.

Achtung, es sieht so aus, als ob die Ableitung von f wieder dieselbe Funktion wäre und somit alle Ableitungen von f gleich A wären. Dies scheint der eindimensionalen Erfahrung zu widersprechen. Es gelten jedoch $f(x) = Ax$ und $Df(x) = A$. Somit ist Df eine konstante Abbildung während f nichtkonstant ist.

Beim ersten Lesen betrachte man diese Definition nur für den Fall $n = 2$ und $E_i = \mathbb{R}$.

Definition 6.5. Seien $E_1, E_2, \ldots, E_n, F$ Banachräume. Sei

$$A\colon E_1 \times E_2 \times \ldots \times E_n \equiv \prod_{j=1}^{n} E_j \to F$$

eine Abbildung.

(i) Sind die Abbildungen $E_i \ni z \mapsto A\left(x^1, \ldots, x^{i-1}, z, x^{i+1}, \ldots, x^n\right)$ für alle $\left(x^1, \ldots, \widehat{x^i}, \ldots, x^n\right) \in \prod\limits_{\substack{j=1\\ j\neq i}}^{n} E_j$ und alle $1 \le i \le n$ linear, so heißt A *multilinear.* Wir verwenden auch in diesem Fall spitze Klammern, $A\langle\ldots\rangle$.

(ii) Wir bezeichnen den Banachraum aller stetigen multilinearen Abbildungen

$$\prod_{i=1}^{n} E_i \to F$$

als $L(E_1, E_2, \ldots, E_n; F)$ und versehen ihn mit der *Operatornorm*

$$\|A\| := \sup_{0\neq x^i\in E_i} \frac{\left\|A\left\langle x^1, \ldots, x^n\right\rangle\right\|}{\|x^1\| \cdot \ldots \cdot \|x^n\|}.$$

(iii) Gilt $E_1 = E_2 = \ldots = E_n = E$, so schreiben wir

$$L_n(E; F) \equiv L(E_1, E_2, \ldots, E_n; F).$$

Bemerkung 6.6.

(i) Sei $A\colon E_1\times\ldots\times E_n \to F$ eine multilineare Abbildung. Dann sind die folgenden Aussagen äquivalent:
 (a) A ist stetig.
 (b) A ist in 0 stetig.
 (c) Die oben definierte Operatornorm $\|A\|$ ist endlich.
 (d) Es gibt ein $c \ge 0$, so dass die Abschätzung

$$\left\|A\left\langle x^1, \ldots, x^n\right\rangle\right\| \le c \cdot \left\|x^1\right\| \cdot \ldots \cdot \|x^n\|$$

 für alle $\left(x^1, \ldots, x^n\right) \in \prod\limits_{i=1}^{n} E_i$ gilt.

(ii) $L(E_1, \ldots, E_n; F)$ ist mit der angegebenen Norm ein Banachraum.

(iii) Ist $A\colon E_1 \times E_2 \to F$ multilinear, so gilt $A\langle\lambda x, \lambda y\rangle = \lambda^2 A\langle x, y\rangle$.
Ist $A\colon E_1 \times E_2 \to F$ linear, so gilt $A\langle(\lambda x, \lambda y)\rangle = \lambda A\langle(x, y)\rangle$.

(iv) Es gibt einen normerhaltenden Isomorphismus

$$L(E_1; L(E_2; F)) \to L(E_1, E_2; F).$$

Er ist durch $\alpha \mapsto ((x, y) \mapsto (\alpha(x))(y))$ für $\alpha \in L(E_1; L(E_2; F))$ gegeben.

Beweis. Wichtige Übung. □

Beispiel 6.7. Seien E_1, E_2, F Banachräume. Setze $E := E_1 \times E_2$ mit der Produktmetrik. Sei $A \in L(E_1, E_2; F)$ eine stetige bilineare Abbildung. Dann ist A in einem beliebigen Punkt $(x_0, y_0) \in E_1 \times E_2$ differenzierbar und es gilt $DA(x_0, y_0)\langle(u, v)\rangle = A(x_0, v) + A(u, y_0)$.

Beweis. Es gilt

$$\begin{aligned} &A(x_0 + u, y_0 + v) - A(x_0, y_0)\\ =\,&A(x_0, y_0) + A(x_0, v) + A(u, y_0) + A(u, v) - A(x_0, y_0)\\ =\,&A(x_0, v) + A(u, y_0) + A(u, v)\end{aligned}$$

$$= DA(x_0, y_0)\langle(u,v)\rangle + o(\|u\| + \|v\|).$$

Für die letzte Gleichheit rechnen wir noch nach, dass $A(u, v) \in o(\|u\| + \|v\|)$ gilt, wobei wir daran erinnern, dass $\|u\| + \|v\|$ die Norm auf dem Produktraum $E_1 \times E_2$ ist. Es gilt

$$\|A(u, v)\| \le \|A\| \cdot \|u\| \cdot \|v\| \le \|A\| \cdot (\|u\| + \|v\|)^2 \in o(\|u\| + \|v\|).$$

Schließlich müssen wir noch zeigen, dass $DA(x_0, y_0) \in L(E_1 \times E_2; F)$ ist. Die Linearität ist klar. Seien $\|u\| + \|v\| \le 1$. Dann folgt

$$\begin{aligned}\|DA(x_0, y_0)\langle(u, v)\rangle\| &\le \|A(x_0, v)\| + \|A(u, y_0)\|\\ &\le \|A\| \cdot \|x_0\| \cdot \|v\| + \|A\| \cdot \|u\| \cdot \|y_0\|\\ &\le \|A\| \cdot \|x_0\| + \|A\| \cdot \|y_0\|\end{aligned}$$

und somit die Stetigkeit von $DA(x_0, y_0)$. □

Ein entsprechendes Resultat gilt auch für multilineare Abbildungen.

Ohne den Darstellungssatz von Fréchet-Riesz können wir den Gradienten zwar definieren, dessen Existenz aber im Allgemeinen nicht zeigen.

Definition 6.8 (Gradient)**.** Seien H ein Hilbertraum, $\Omega \subset H$ offen und $f\colon \Omega \to \mathbb{K}$ in $x_0 \in \Omega$ differenzierbar. Dann ist $Df(x_0) \in L(H; \mathbb{K}) \equiv H^*$. Aufgrund der allgemeinen Hilbertraumtheorie (Darstellungssatz von Fréchet-Riesz) gibt es einen Vektor $v \in H$ mit $\langle\cdot, v\rangle_H = Df(x_0)\langle\cdot\rangle$. Dann heißt v der *Gradient* von f in x_0. Wir schreiben dafür $\operatorname{grad} f(x_0)$ oder $\nabla f(x_0)$. Es gilt dann für alle $u \in H$

$$\langle u, \nabla f(x_0)\rangle = Df(x_0)\langle u\rangle.$$

Beispiel 6.9.

(i) Sei H ein reeller Hilbertraum und $f\colon H \to \mathbb{R}$ durch $f(x) = \|x\|^2 = \langle x, x\rangle$ definiert. Dann gilt $\nabla f(x_0) = 2x_0$ für alle $x_0 \in H$.

Beweis. Es gilt für alle $x \in H$

$$f(x_0 + x) = f(x_0) + 2\langle x_0, x\rangle + f(x).$$

Hieraus folgt die Behauptung. □

Theorem 6.10 (Kettenregel)**.** *Seien E, F, G Banachräume und seien $\Omega \subset E$ und $V \subset F$ offen. Seien $g\colon \Omega \to F$ in x_0 und $f\colon V \to G$ in $g(x_0) \in V$ differenzierbar. Dann ist die Verknüpfung $f \circ g$ in einer Umgebung $B_\delta(x_0)$ von x_0 wohldefiniert und in x_0 differenzierbar. Es gilt*

$$(f \circ g)'(x_0) = f'(g(x_0)) \circ g'(x_0).$$

Wir folgen im Wesentlichen dem eindimensionalen Beweis der Kettenregel und wiederholen den Beweis, da die Kettenregel wichtig ist.

Beweis.

(i) Wohldefiniertheit: Wir haben angenommen, dass $g(x_0) \in V$ gilt. Da V offen ist, gibt es eine Umgebung $B_r(g(x_0))$ von $g(x_0)$ mit $B_r(g(x_0)) \subset V$. Aufgrund der Stetigkeit von g in x_0 gibt es ein $\delta > 0$ mit $B_\delta(x_0) \subset \Omega$ und $g(B_\delta(x_0)) \subset B_r(g(x_0))$. Somit ist die Verknüpfung $f \circ g$ in $B_\delta(x_0)$ wohldefiniert.

(ii) Aufgrund der Differenzierbarkeit von f in $g(x_0)$ erhalten wir

$$f \circ g(x) - f \circ g(x_0) = f'(g(x_0))\langle g(x) - g(x_0)\rangle + o(\|g(x) - g(x_0)\|).$$

Aufgrund der Differenzierbarkeit von g in x_0 erhalten wir

$$g(x) - g(x_0) = g'(x_0)\langle x - x_0\rangle + o(\|x - x_0\|).$$

Zusammengenommen erhalten wir also

$$f \circ g(x) - f \circ g(x_0) = f'(g(x_0))\langle g'(x_0)\langle x - x_0\rangle\rangle + f'(g(x_0))\langle o(\|x - x_0\|)\rangle \\ + o(\|g(x) - g(x_0)\|).$$

(iii) Die beiden Restterme liegen in $o(\|x - x_0\|)$: Dies ist für den ersten Restterm klar. Den zweiten Restterm schreiben wir als

$$o(\|g(x) - g(x_0)\|) = \varepsilon(\|g(x) - g(x_0)\|) \cdot \|g(x) - g(x_0)\|$$

mit einer Funktion ε, die $\varepsilon \to 0$ für $g(x) \to g(x_0)$ erfüllt. Wir erhalten nach Anwendung der Norm und der Dreiecksungleichung auf die Definition von Differenzierbarkeit für g

$$\|g(x) - g(x_0)\| \le \|g'(x_0)\| \cdot \|x - x_0\| + o(\|x - x_0\|).$$

Aufgrund dieser Abschätzung bleibt $\frac{\|g(x)-g(x_0)\|}{\|x-x_0\|}$ im Grenzwert $x \to x_0$ beschränkt und da $x \to x_0$ aufgrund der Stetigkeit von g auch $g(x) \to g(x_0)$ impliziert, folgt $\varepsilon(\|g(x) - g(x_0)\|) \to 0$. Somit erhalten wir

$$o(\|g(x) - g(x_0)\|) \in o(\|x - x_0\|)$$

und die Behauptung folgt. □

Do 18.06.2020

Korollar 6.11. *Seien E, E' Banachräume, seien $\Omega \subset E$ und $\Omega' \subset E'$ offene Teilmengen und sei $f\colon \Omega \to \Omega'$ invertierbar und $g\colon \Omega' \to \Omega$ die Inverse. Sei f in $x_0 \in \Omega$ differenzierbar und g in $f(x_0)$ differenzierbar, so gilt*

$$g'(f(x_0)) = (f'(x_0))^{-1}.$$

Beweis. Es gelten $\mathrm{id}_\Omega = g \circ f$ und $\mathrm{id}_{\Omega'} = f \circ g$. Setze $y_0 := f(x_0)$. Wir erhalten aufgrund der Kettenregel

$$\mathrm{id}_E = g'(f(x_0)) \circ f'(x_0) \quad \text{und} \quad \mathrm{id}_{E'} = f'(g(y_0)) \circ g'(y_0) = f'(x_0) \circ g'(f(x_0)).$$

Somit ist $g'(f(x_0))$ beidseitige Inverse von $f'(x_0)$ und nach Voraussetzung stetig, also in $L(E', E)$. □

Theorem 6.12. *Seien E, E' Banachräume. Seien $\Omega \subset E$ und $\Omega' \subset E'$ offene Teilmengen. Sei $f\colon \Omega \to \Omega'$ ein Homöomorphismus mit Inverse $g\colon \Omega' \to \Omega$. Falls f in x_0 differenzierbar ist und $f'(x_0)$ ein Homöomorphismus ist, ist g in $f(x_0)$ differenzierbar und es gilt*

$$g'(f(x_0)) = (f'(x_0))^{-1}.$$

Beachte, dass wir hier nicht mehr voraussetzen, dass g in $f(x_0)$ differenzierbar ist.

Beweis.

(i) Für einen linearen Homöomorphismus A gilt allgemein folgendes: A^{-1} ist stetig (und linear). Somit gibt es ein $c > 0$, so dass $\|A^{-1}y\| \le c^{-1} \cdot \|y\|$ für alle y gilt. Wir wählen speziell $y = Ax$ und erhalten $c \cdot \|x\| \le \|Ax\|$. Somit erhalten wir in unserer Situation ein $c > 0$ mit

$$c\|x\| \le \|f'(x_0)\langle x\rangle\| \quad \text{für alle } x \in E.$$

(ii) Aufgrund der Differenzierbarkeit von f in x_0 erhalten wir mit der schon bekannten Schreibweise $\varepsilon(\|a\|)$ für eine Funktion mit $\varepsilon(\|a\|) \to 0$ für $a \to 0$

$$\|f(x) - f(x_0)\| \ge \|f'(x_0)\langle x - x_0\rangle\| - \|o(\|x - x_0\|)\| \\ \ge c\|x - x_0\| - \varepsilon(\|x - x_0\|) \cdot \|x - x_0\|$$

$$\geq \tfrac{c}{2}\|x - x_0\|$$

für alle $x \in B_\delta(x_0)$, falls wir $\delta > 0$ so klein wählen, dass für $x \in B_\delta(x_0)$ bereits $\varepsilon(\|x - x_0\|) \leq \frac{c}{2}$ gilt.

(iii) Wir kürzen nun $y = f(x)$ und $y_0 = f(x_0)$ ab und erhalten

$$\begin{aligned} & g(y) - g(y_0) - (f'(x_0))^{-1}\langle y - y_0\rangle \\ = & \, x - x_0 - (f'(x_0))^{-1}\langle f(x) - f(x_0)\rangle \\ = & \, (f'(x_0))^{-1}\langle f'(x_0)\langle x - x_0\rangle - f(x) + f(x_0)\rangle \\ = & \, (f'(x_0))^{-1}\langle -o(\|x - x_0\|)\rangle \overset{!}{=} o(\|y - y_0\|). \end{aligned}$$

Zur Begründung der letzten Behauptung genügt aufgrund der Stetigkeit von $(f'(x_0))^{-1}$ der Nachweis, dass $o(\|x - x_0\|) \in o(\|y - y_0\|)$ gilt. Es gilt

$$o(\|x - x_0\|) = \varepsilon(\|x - x_0\|) \cdot \|x - x_0\| \overset{\text{(ii)}}{\leq} \varepsilon(\|x - x_0\|) \cdot \tfrac{2}{c}\|y - y_0\|.$$

Wegen $x \to x_0 \iff y \to y_0$ folgt damit $o(\|x - x_0\|) \in o(\|y - y_0\|)$ und somit auch die Behauptung über die Differenzierbarkeit von g in $f(x_0)$. $\square$

Proposition 6.13. *Seien E ein Banachraum, $\Omega \subset E$ offen und $F = \prod\limits_{i=1}^{n} F_i$ ein Produktraum von Banachräumen. Sei $f\colon \Omega \to F$ mit $f = \left(f^i\right)_{1\leq i\leq n}$ eine Abbildung. Dann ist f genau dann in x_0 differenzierbar, falls jede Komponente f^i, $1 \leq i \leq n$, in x_0 differenzierbar ist und es gilt*

$$Df(x_0) = \left(Df^1(x_0), \ldots, Df^n(x_0)\right),$$

wobei wir $L(E, F)$ und $\prod\limits_{i=1}^{n} L(E, F_i)$ (mit Maximumsnorm) wie nachfolgend angegeben identifizieren.

Beweis.

(i) Zur Identifikation:
  (a) Zu $A \in L(E, F)$ definieren wir $A_i = \pi_i \circ A \in L(E, F_i)$.
  (b) Umgekehrt ordnen wir den Abbildungen $A_i \in L(E, F_i)$, $1 \leq i \leq n$, die Abbildung $A\colon x \mapsto (A_1x, \ldots, A_nx)$ zu.
  (c) Nach Definition der Maximumsnorm auf dem Produktraum gilt $\|Ax\| = \max\limits_{1\leq i\leq n} \|A_ix\|$ und wir erhalten

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} \max_{1\leq i\leq n} \|A_ix\| = \max_{1\leq i\leq n} \sup_{\|x\|=1} \|A_ix\| = \max_{1\leq i\leq n} \|A_i\|.$$

  Damit sind die angegebenen Abbildungen normtreue Isomorphismen, wir dürfen die beiden Räume also wie angegeben identifizieren.

(ii) Differenzierbarkeit:
  (a) Die kanonische Projektion $\pi_i\colon F \to F_i$ ist eine lineare stetige Abbildung. Ist f in x_0 differenzierbar, so auch $f^i = \pi_i \circ f$ und nach Kettenregel gilt $Df^i(x_0) = \pi_i \circ Df(x_0)$.
  (b) Seien umgekehrt die Abbildungen f^i, $1 \leq i \leq n$, in x_0 differenzierbar. Dann erhalten wir für jedes i

$$f^i(x) = f^i(x_0) + Df^i(x_0)\langle x - x_0\rangle + o(\|x - x_0\|).$$

Diese Komponenten fassen wir zusammen und erhalten mit Hilfe der obigen Identifikation

$$f(x) = f(x_0) + \left(Df^1(x_0), \dots, Df^n(x_0)\right) \langle x - x_0 \rangle \\ + \underbrace{\left(o(\|x - x_0\|), \dots, o(\|x - x_0\|)\right)}_{\in o(\|x-x_0\|)},$$

wobei „$\in$“ nach Definition der Produktnorm folgt. □

Proposition 6.14 (Produkt- und Quotientenregel)**.** *Seien E ein Banachraum, $\Omega \subset E$ offen und f, $g\colon \Omega \to \mathbb{K}$ in $x_0 \in \Omega$ differenzierbar.*

(i) Dann ist auch $h := fg$ in x_0 differenzierbar und es gilt

$$Dh(x_0) = Df(x_0) \cdot g(x_0) + f(x_0) \cdot Dg(x_0).$$

(Dabei ist $Dh(x_0) \in L(E; \mathbb{K})$ durch

$$E \ni v \mapsto Df(x_0)\langle v \rangle \cdot g(x_0) + f(x_0) \cdot Dg(x_0)\langle v \rangle \in \mathbb{K}$$

gegeben.)

(ii) Ist $f \neq 0$ nahe x_0, so ist auch $h := \frac{1}{f}$ in x_0 differenzierbar und es gilt dort

$$Dh(x_0) = -\frac{1}{f^2(x_0)} Df(x_0).$$

Beweis.

(i) Statt dies von Hand zu zeigen, können wir wie folgt Bekanntes verwenden: Sei $a \in L_2(\mathbb{K}; \mathbb{K})$ definiert durch $a(s,t) := s \cdot t$. Sei weiter $\varphi\colon \Omega \to \mathbb{K}^2$ die Abbildung mit $\varphi(x) := (f(x), g(x))$. Dann gilt $h = a \circ \varphi$, h ist nach Kettenregel differenzierbar und es gilt $Dh(x_0) = Da(\varphi(x_0))\langle D\varphi(x_0)\langle \cdot \rangle \rangle$. Man überzeugt sich anhand der bereits hergeleiteten Formeln für Da und $D\varphi$, dass dies tatsächlich die Behauptung zeigt.

(ii) Übung. □

Theorem 6.15. *Seien E ein Banachraum, $\Omega \subset E$ offen und $f\colon \Omega \to \mathbb{R}$ eine in $x_0 \in \Omega$ differenzierbare Funktion, die in x_0 ein lokales Maximum annimmt. Dann gilt $Df(x_0) = 0$.*

Definition: *Punkte x_0 mit $Df(x_0) = 0$ nennen wir* kritische Punkte.

Beweis. Sei $e \in E$ ein beliebiger Vektor. Dann gibt es $\varepsilon > 0$, so dass $x_0 + te \in \Omega$ für alle $t \in (-\varepsilon, \varepsilon)$ gilt. Nach Kettenregel ist die Funktion $\varphi\colon \mathbb{R} \to \mathbb{R}$ mit $\varphi(t) := f(x_0 + te)$ in $t = 0$ differenzierbar und nimmt in $t = 0$ ein lokales Maximum an. Somit gilt nach Proposition 4.15

$$0 = \varphi'(0) = Df(x_0)\langle e \rangle.$$

Da $e \in E$ beliebig war, folgt $Df(x_0) = 0 \in L(E, \mathbb{R})$. □

Definition 6.16. Seien E, F Banachräume und $\Omega \subset E$ offen. Dann definieren wir die folgenden Funktionenräume:

(i) $C^1(\Omega, F) := \{f\colon \Omega \to F \,:\, f \text{ ist in } \Omega \text{ stetig differenzierbar}\}$.

(ii) $C^1\left(\overline{\Omega}, F\right) := \{f \in C^1(\Omega, F)\colon f$ und Df lassen sich stetig und beschränkt auf $\overline{\Omega}$ fortsetzen$\}$.

(iii) Wir definieren die C^1-Norm auf $C^1\left(\overline{\Omega}, F\right)$ durch

$$\|f\|_{C^1\left(\overline{\Omega}, F\right)} \equiv \|f\|_{C^1(\Omega, F)} \equiv \|f\|_{C^1} := \sup_{x \in \Omega} \|f(x)\|_F + \sup_{x \in \Omega} \|Df(x)\|_{L(E,F)}.$$

(iv) Ist $F = \mathbb{R}$, so schreiben wir auch $C^1(\Omega) \equiv C^1(\Omega, \mathbb{R})$, $C^1\left(\overline{\Omega}\right) \equiv C^1\left(\overline{\Omega}, \mathbb{R}\right)$ oder gelegentlich auch $|f|_{C^1} \equiv \|f\|_{C^1}$.

Theorem 6.17. *Seien E, F Banachräume und sei $\Omega \subset E$ offen. Dann ist $C^1\left(\overline{\Omega}, F\right)$ mit der C^1-Norm ein Banachraum.*

Beweis. Übung. □

Mo 22.06.2020

Eulersche Homogenitätsrelation.

Definition 6.18.

(i) Sei E ein $\mathbb{R}$-Vektorraum. Dann heißt $\Gamma \subset E$ ein *Kegel,* falls
$$x \in \Gamma \quad \implies \quad tx \in \Gamma$$
für alle $t > 0$ gilt.

(ii) Seien E, F zwei $\mathbb{R}$-Vektorräume und $\Gamma \subset E$ ein Kegel. Dann heißt eine Funktion $\varphi\colon \Gamma \to F$ *positiv homogen* vom Grade α mit $\alpha \in \mathbb{R}$, falls
$$\varphi(tx) = t^\alpha \varphi(x)$$
für alle $t > 0$ und alle $x \in \Gamma$ gilt.

Proposition 6.19 (Eulersche Homogenitätsrelation)**.** *Seien E, F Banachräume, $\Omega \subset E$ ein offener Kegel und sei $\varphi\colon \Omega \to F$ differenzierbar. Dann ist φ genau dann positiv homogen vom Grade α, wenn die* Eulersche Homogenitätsrelation
$$D\varphi(x)\langle x\rangle = \alpha\varphi(x)$$
für alle $x \in \Omega$ gilt.

Beweis.

„$\Rightarrow$“: Sei φ positiv homogen vom Grade α. Fixiere $x \in \Omega$ beliebig und definiere $f(t) := \varphi(tx)$ für $t > 0$. Wegen $f(t) = t^\alpha\varphi(x)$ können wir f auf zwei Arten differenzieren und erhalten
$$f'(t) = D\varphi(tx)\langle x\rangle = \alpha t^{\alpha-1}\varphi(x).$$
Wir werten dies im Punkt $t = 1$ aus und erhalten die Behauptung.

„$\Leftarrow$“: Zusatzmaterial. □

6.2. **Der Mittelwertsatz und Anwendungen.** Den schon bekannten Mittelwertsatz können wir wie folgt verallgemeinern.

Theorem 6.20 (Mittelwertsatz (MWS))**.** *Seien E, F Banachräume, $\Omega \subset E$ offen und sei $f\colon \Omega \to F$ differenzierbar. Seien $x, y \in \Omega$ mit $[x, y] \subset \Omega$. Dann folgt*
$$\|f(x) - f(y)\| \le \sup_{0\le t\le 1} \|Df(tx + (1-t)y)\| \cdot \|x - y\| = \sup_{z\in[x,y]} \|Df(z)\| \cdot \|x - y\|.$$

Beweis. Wir definieren $\varphi(t) := f(tx + (1-t)y)$. (Manche Autoren bevorzugen die äquivalente Definition $\varphi(t) := f(y + t(x - y))$. Auf jeden Fall werden solche *Konvexkombinationen* häufiger verwandt.) Da Ω offen ist, ist φ nicht nur auf $[0, 1]$ sondern auch noch auf $(-\varepsilon, 1 + \varepsilon)$ für ein kleines $\varepsilon > 0$ definiert und aufgrund der Kettenregel differenzierbar. Mit Hilfe des Mittelwertsatzes für auf Intervallen definierte Funktionen und der Kettenregel erhalten wir
$$\begin{aligned}\|f(x) - f(y)\| = \|\varphi(1) - \varphi(0)\| &\le \sup_{0\le t\le 1} \|\varphi'(t)\| \cdot |1 - 0| \\ &= \sup_{0\le t\le 1} \|Df(tx + (1-t)y)\langle x - y\rangle\| \\ &\le \sup_{0\le t\le 1} \|Df(tx + (1-t)y)\| \cdot \|x - y\|\end{aligned}$$
wie behauptet. □

Analog zum Eindimensionalen erhalten wir auch hier das folgende

Korollar 6.21. *Seien E, F Banachräume, $\Omega \subset E$ offen und $f\colon \Omega \to F$ differenzierbar. Seien $x, y, x_0 \in \Omega$ beliebig mit $[x, y] \subset \Omega$. Dann gilt*

$$\|f(x) - f(y) - Df(x_0)\langle x - y\rangle\| \le \sup_{z\in[x,y]} \|Df(z) - Df(x_0)\| \cdot \|x - y\|.$$

Beweis. Wende den Mittelwertsatz auf die Funktion

$$g(x) := f(x) - Df(x_0)\langle x\rangle$$

an. □

Korollar 6.22. *Seien E, F Banachräume, $\Omega \subset E$ ein* Gebiet, *d. h. eine offene und zusammenhängende Menge, und $f\colon \Omega \to F$ differenzierbar. Gilt $Df \equiv 0$ in Ω, so ist f konstant.*

Beweis. Sei $x_0 \in \Omega$. Aus Korollar 6.21 folgt, dass f auf Geradenstücken in Ω, d. h. Mengen der Form $[x, y] \subset \Omega$, konstant ist. Somit folgt, dass $\Lambda := \{x \in \Omega\colon f(x) = f(x_0)\}$ offen, abgeschlossen und nichtleer ist, genauer:

- offen: Sei $y \in \Lambda$. Dann gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(y) \subset \Omega$. Sei $z \in B_\varepsilon(y)$ beliebig. Wegen $[y, z] \in \Omega$ ist f auf $[y, z]$ konstant. Somit folgt insbesondere $z \in \Lambda$ und Λ ist offen.
- abgeschlossen: Aufgrund der Stetigkeit von f ist Λ (relativ) abgeschlossen. (Wie bei der Offenheit könnte man hier auch vom Grenzwert einer in Ω konvergenten Folge her argumentieren.)
- nichtleer: Es gilt $x_0 \in \Lambda$.

Da Ω zusammenhängend ist, folgt die Behauptung. □

6.3. **Differentiation von Funktionenfolgen.** Wir erhalten ähnliche Resultate wie im Eindimensionalen, z. B. wie in Theorem 4.32.

Theorem 6.23. *Seien E, F Banachräume, $\Omega \subset E$ offen. Sei $(f_n)_n \subset C^1(\Omega, F)$ eine Funktionenfolge mit $f_n \to f$ und $Df_n \rightrightarrows g$ für Funktionen $f\colon \Omega \to F$ und $g\colon \Omega \to L(E, F)$. Dann ist f ebenfalls stetig differenzierbar und es gilt $Df = g$. Ist Ω zusätzlich konvex und beschränkt, so folgt $f_n \rightrightarrows f$.*

Beweis.

(i) g ist als Grenzwert stetiger Funktionen unter gleichmäßiger Konvergenz selbst wieder stetig.

(ii) $Df = g$: Sei $x_0 \in \Omega$ beliebig. Für den Nachweis, dass f differenzierbar ist und $Df = g$ gilt, dürfen wir ohne Einschränkung annehmen, dass Ω ein Ball ist.
Seien $x, x_0 \in \Omega$. Der Mittelwertsatz, Korollar 6.21, impliziert

$$\|f_n(x) - f_n(x_0) - Df_n(x_0)\langle x - x_0\rangle\| \le \sup_{y\in[x,x_0]} \|Df_n(y) - Df_n(x_0)\| \cdot \|x - x_0\|.$$

Aufgrund der punktweisen Konvergenz $f_n \to f$ und $Df_n \to g$ konvergiert die linke Seite für $n \to \infty$ gegen $\|f(x) - f(x_0) - g(x_0)\langle x - x_0\rangle\|$. Für die rechte Seite liefert die Dreiecksungleichung

$$\begin{aligned} &\sup_{y\in[x,x_0]} \|Df_n(y) - Df_n(x_0)\| \\ \le &\sup_{y\in[x,x_0]} \|Df_n(y) - g(y)\| + \sup_{y\in[x,x_0]} \|g(y) - g(x_0)\| + \|g(x_0) - Df_n(x_0)\|. \end{aligned}$$

Auf der rechten Seite konvergiert der erste Term aufgrund der gleichmäßigen Konvergenz $Df_n \rightrightarrows g$ gegen Null und der dritte Term konvergiert aufgrund der punktweisen Konvergenz ebenfalls für $n \to \infty$ gegen Null. Somit folgt

$$\|f(x) - f(x_0) - g(x_0)\langle x - x_0\rangle\| \le \sup_{y\in[x,x_0]} \|g(y) - g(x_0)\| \cdot \|x - x_0\|.$$

Da g stetig ist, folgt $\sup\limits_{y\in[x,x_0]} \|g(y)-g(x_0)\| \to 0$ für $x\to x_0$. Somit liegt die rechte Seite in $o(\|x-x_0\|)$, f ist an der Stelle x_0 differenzierbar und es gilt $g = Df$.

(iii) $f_n \rightrightarrows f$: Sei $d := \operatorname{diam}\Omega$. Sei $\varepsilon > 0$. Sei $x \in \Omega$ beliebig und sei $y \in \Omega$ fest gewählt. Dann impliziert der Mittelwertsatz, angewandt auf die Funktion $f_n - f_m$,
$$\|(f_n(x)-f_m(x))-(f_n(y)-f_m(y))\| \le \sup_{z\in\Omega} \|Df_n(z)-Df_m(z)\|\cdot d.$$
Aufgrund der gleichmäßigen Konvergenz ist die rechte Seite für große n, m kleiner als ε und aufgrund der punktweisen Konvergenz gilt $\|f_n(y)-f_m(y)\| \le \varepsilon$ für große n, m. Somit folgt aufgrund der Dreiecksungleichung für beliebige $x\in\Omega$
$$\begin{aligned}&\|f_n(x)-f_m(x)\|\\ &\le \|(f_n(x)-f_m(x))-(f_n(y)-f_m(y))\| + \|f_n(y)-f_m(y)\| \le 2\varepsilon.\end{aligned}$$
Die Behauptung folgt. $\square$

Differentiation von Reihen.

Beispiel 6.24. Sei E ein Banachraum. Definiere
$$\begin{aligned}\exp\colon L(E) &\to L(E),\\ A &\mapsto \exp A \equiv e^A := \sum_{n\in\mathbb{N}} \frac{A^n}{n!}.\end{aligned}$$
Wir möchten nun die Ableitung der Exponentialfunktion ausrechnen. Dies wird im dritten Semester für lineare gewöhnliche Differentialgleichungen wichtig.

(i) Definiere $f_n\colon L(E)\to L(E)$ durch $A\mapsto A^n$.

(ii) Wir definieren weiterhin die lineare Abbildung $\varphi\colon L(E)\to L(E)^n$ durch $A\mapsto (A,\ldots,A)$ und erhalten $D\varphi(A)\langle B\rangle = (B,\ldots,B)$.

(iii) Definiere schließlich die multilineare Abbildung $a \in L_n(L(E);L(E))$ durch
$$a(A_1,\ldots,A_n) = A_1\circ\ldots\circ A_n.$$
Als multilineare Abbildung ist a differenzierbar und es gilt
$$Da(A_1,\ldots,A_n)\langle B,\ldots,B\rangle = \sum_{i=1}^n A_1\circ\ldots\circ A_{i-1}\circ B\circ A_{i+1}\circ\ldots\circ A_n.$$

(iv) Es gilt $f_n = a\circ\varphi$ und daher nach Kettenregel
$$Df_n(A)\langle B\rangle = Da(\varphi(A))\langle D\varphi(A)\langle B\rangle\rangle = \sum_{i=1}^n \underbrace{A\circ\ldots\circ A}_{i-1\text{ Stück}}\circ B\circ \underbrace{A\circ\ldots\circ A}_{n-i\text{ Stück}}.$$

(v) Es gilt $\|Df_n(A)\langle B\rangle\| \le n\|A\|^{n-1}\|B\|$ und somit folgt $\|Df_n(A)\| \le n\|A\|^{n-1}$. Also konvergiert die Reihe $\left(\left(\frac{1}{n!}Df_n(\cdot)\right)\right)_n$ auf beliebigen aber festen Bällen $B_r(0)\subset L(E)$ gleichmäßig (Majorantenkriterium mit $\exp'(r) = \exp(r)$). Somit ist die Exponentialfunktion differenzierbar und es gilt
$$D\exp(A) = \sum_{n\in\mathbb{N}} \frac{1}{n!} Df_n(A)\langle B\rangle.$$

(vi) Der Ausdruck für die Ableitung wird im Falle $[A,B]=0$, d. h. $AB = BA$ übersichtlicher. In diesem Fall gilt $Df_n(A)\langle B\rangle = nA^{n-1}B$ und wir erhalten
$$D\exp(A)\langle B\rangle = e^A B = Be^A.$$

Do 25.06.2020

6.4. **Partielle Ableitungen.** Wir beschränken uns hier auf den $\mathbb{R}^n$-Fall. Beim ersten Lesen genügt es, den Fall $\mathbb{R}$-wertiger Funktionen zu verstehen. Der Banachraumfall ist herauskommentiert als Anhang vorhanden.

Definition 6.25. Sei $\Omega \subset \mathbb{R}^n$ offen. Sei F ein Banachraum und sei $f\colon \Omega \to F$ eine Funktion. Sei $x_0 = \left(x_0^i\right)_{1\le i\le n} \in \Omega$ und sei $\varepsilon > 0$, so dass für alle $z = \left(z^i\right)_{1\le i\le n} \in \mathbb{R}^n$ mit $\left|x_0^i - z^i\right| < \varepsilon$ auch $z \in \Omega$ gilt.

(i) Dann heißt f in x_0 in Richtung i, $1 \le i \le n$, *partiell differenzierbar*, falls die Abbildung
$$\left(x_0^i - \varepsilon, x_0^i + \varepsilon\right) \ni x^i \mapsto f\left(x_0^1, \ldots, x_0^{i-1}, x^i, x_0^{i+1}, \ldots, x_0^n\right) \equiv f\left(\hat{x}_0^i, x^i\right)$$
in x_0^i differenzierbar ist. Die Ableitung dieser Funktion heißt *partielle Ableitung* von f in x_0 in Richtung i (oder bezüglich x^i). Wir schreiben $D_i f(x_0) \equiv \frac{\partial f}{\partial x^i}(x_0) \equiv f_i(x_0)$. Es gilt $D_i f(x_0) \in L(\mathbb{R}, F) \cong F$.

(ii) f heißt in x_0 *partiell differenzierbar,* falls f in x_0 in alle Richtungen i, $1 \le i \le n$, partiell differenzierbar ist.

(iii) Ist f in allen Punkten $x \in \Omega$ partiell differenzierbar, so ist $D_i f$ eine Abbildung $D_i f\colon \Omega \to F$ und f heißt in Ω *partiell differenzierbar.* Sind alle diese Abbildungen $D_i f$, $1 \le i \le n$, stetig, so heißt f *stetig partiell differenzierbar.*

Proposition 6.26. *Sei $\Omega \subset \mathbb{R}^n$ offen. Sei F ein Banachraum und sei $f\colon \Omega \to F$ in $x_0 \in \Omega$ differenzierbar. Dann ist f in y auch partiell differenzierbar.*

Ist $\chi_i\colon \mathbb{R} \hookrightarrow \mathbb{R}^n$ mit $x \mapsto (0, \ldots, 0, x, 0, \ldots, 0)$ die kanonische Einbettung (in die i-te Komponente), so folgt
$$D_i f(x_0) = Df(x_0) \circ \chi_i.$$

Beweis. Sei $\varepsilon > 0$ so klein, dass $\left(\hat{x}_0^i, x^i\right) \equiv \left(x_0^1, \ldots, x_0^{i-1}, x^i, x_0^{i+1}, \ldots, x_0^n\right) \in \Omega$ für alle $x^i \in \left(x_0^i - \varepsilon, x_0^i + \varepsilon\right)$ gilt. Wir definieren die Abbildung
$$\varphi\colon (x_0^i - \varepsilon, x_0^i + \varepsilon) \to \Omega,$$
$$x^i \mapsto \left(\hat{x}_0^i, x^i\right).$$
Da φ affin linear ist, folgt
$$D\varphi\left(x^i\right)\langle t\rangle = (0, \ldots, 0, t, 0, \ldots, 0) = \chi_i\langle t\rangle$$
für alle $x^i \in (x_0^i - \varepsilon, x_0^i + \varepsilon)$ und alle $t \in \mathbb{R}$, also $D\varphi\left(x^i\right) = \chi_i$. Nun gilt
$$f\left(\hat{x}_0^i, x^i\right) = f \circ \varphi\left(x^i\right).$$
Auf der linken Seite steht hier die Funktion aus der Definition der partielle Ableitung. Wenden wir die Kettenregel auf die rechte Seite an, so erhalten wir
$$D_i f(x_0) = Df\left(\varphi\left(x_0^i\right)\right)\left\langle D\varphi\left(x_0^i\right)\right\rangle = Df(x_0) \circ \chi_i$$
und daher gerade die Behauptung. □

Die Ableitung einer Funktion ergibt sich aus den partiellen Ableitungen.

Korollar 6.27. *Sei $\Omega \subset \mathbb{R}^n$ offen und sei F ein Banachraum. Sei $f\colon \Omega \to F$ in $x_0 \in \Omega$ differenzierbar. Dann folgt*
$$Df(x_0)\langle u\rangle = \sum_{i=1}^n D_i f(x_0)\left\langle u^i\right\rangle \quad \text{für } u = \left(u^i\right)_{1\le i\le n} \in \mathbb{R}^n,$$
d. h. es gilt
$$Df(x_0) = \sum_{i=1}^n D_i f(x_0) \circ \pi_i,$$
wobei $\pi_i\colon \mathbb{R}^n \to \mathbb{R}$ die kanonische Projektion auf die i-te Komponente bezeichnet.

Beweis. Es gilt nach Proposition 6.26 in der dortigen Notation und aufgrund der Linearität

$$\sum_{i=1}^{n} D_i f(x_0)\left\langle u^i\right\rangle = \sum_{i=1}^{n} Df(x_0)\left\langle \chi_i\left\langle u^i\right\rangle\right\rangle = Df(x_0)\left\langle \sum_{i=1}^{n}\chi_i\left\langle u^i\right\rangle\right\rangle = Df(x_0)\langle u\rangle$$

wie behauptet. □

Bemerkung 6.28. Sei $\Omega \subset \mathbb{R}^n$ offen und $f\colon \Omega \to \mathbb{R}$ in $x_0 \in \Omega$ differenzierbar, so haben wir $Df(x_0)$ mit einem Vektor in $\mathbb{R}^n$ identifiziert. Nun gilt

$$\nabla f(x_0) = (f_1, \ldots, f_n)^T.$$

Umgekehrt braucht eine partiell differenzierbare Funktion nicht differenzierbar zu sein.

Beispiel 6.29. Definiere $f\colon \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x,y) := \begin{cases} \frac{x^3}{x^2+y^2}, & (x,y) \neq 0, \\ 0, & (x,y) = 0. \end{cases}$$

Insbesondere gelten $f(x,0) = x$ und $f(0,y) = 0$. Daher ist f im Ursprung partiell differenzierbar mit $f_1(0) \equiv f_x(0) = 1$ und $f_2(0) \equiv f_y(0) = 0$.

Wäre f im Ursprung differenzierbar, so folgte insbesondere

$$o(|x|) = f(x,x) - f(0,0) - \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ x \end{pmatrix} \right\rangle = \frac{1}{2}x - x = -\frac{1}{2}x.$$

Widerspruch.

Für stetig differenzierbare Funktionen gilt die Umkehrung jedoch

Theorem 6.30. *Sei $\Omega \subset \mathbb{R}^n$ offen und sei F ein Banachraum. Dann ist $f\colon \Omega \to F$ genau dann in Ω stetig differenzierbar, wenn f in Ω stetig partiell differenzierbar ist.*

Beweis. Wegen Proposition 6.26 genügt es zu zeigen, dass f in Ω stetig differenzierbar ist, falls f in Ω stetig partiell differenzierbar ist, dass also „$\Longleftarrow$“ richtig ist.

Ist f differenzierbar, so können wir nach Korollar 6.27 die Ableitung Df mit Hilfe der partiellen Ableitungen $D_i f$ darstellen. Damit ist Df stetig, falls dies für $D_i f$ gilt und wir müssen nur noch die Differenzierbarkeit von f in Ω nachweisen. Sei dazu $x_0 \in \Omega$ beliebig und Ω ohne Einschränkung konvex. Wir wollen zeigen, dass

$$f(x) = f(x_0) + \sum_{i=1}^{n} D_i f(x_0)\left\langle x^i - x_0^i\right\rangle + o(\|x - x_0\|)$$

gilt.

Zur besseren Übersicht führen wir den Rest des Beweises nur im Falle $n = 2$ durch. Es gilt

$$\begin{aligned} f\left(x^1,x^2\right) - f\left(x_0^1,x_0^2\right) &= f\left(x^1,x^2\right) - f\left(x_0^1,x^2\right) + f\left(x_0^1,x^2\right) - f\left(x_0^1,x_0^2\right) \\ &= D_1 f\left(x_0^1,x^2\right)\left\langle x^1 - x_0^1\right\rangle + o_{x^2}\left(\left\|x^1 - x_0^1\right\|\right) \\ &\quad + D_2 f\left(x_0^1,x_0^2\right)\left\langle x^2 - x_0^2\right\rangle + o\left(\left\|x^2 - x_0^2\right\|\right). \end{aligned} \tag{6.1}$$

Auf der rechten Seite hat der erste Term noch das falsche Argument, der zweite Term hängt noch in unerwünschter Weise von x^2 ab und die beiden letzten Terme haben die gewünschte Form.

Für den zweiten Term benutzen wir die Definition dieses Terms und den Mittelwertsatz und erhalten

$$\left\|o_{x^2}\left(\left\|x^1 - x_0^1\right\|\right)\right\| = \left\|f\left(x^1,x^2\right) - f\left(x_0^1,x^2\right) - D_1 f\left(x_0^1,x^2\right)\left\langle x^1 - x_0^1\right\rangle\right\|$$

$$\leq \sup_{\xi \in \left[x^1, x_0^1\right]} \left\| D_1 f\left(\xi, x^2\right) - D_1 f\left(x_0^1, x^2\right)\right\| \cdot \underbrace{\left\|x^1 - x_0^1\right\|}_{\leq \|x - x_0\|}.$$

Aufgrund der Stetigkeit von $D_1 f$ sehen wir, dass der Term $o_{x^2}\left(\left\|x^1 - x_0^1\right\|\right)$ ein $o(\|x - x_0\|)$-Term ist. Insbesondere wird also der erste Faktor in einer kleinen Umgebung von $\left(x_0^1, x_0^2\right)$ unabhängig von x^2 klein.

Nun zum ersten Term der rechten Seite in (6.1). Es gilt

$$\begin{aligned} D_1 f\left(x_0^1, x^2\right)\left\langle x^1 - x_0^1\right\rangle = {} & D_1 f\left(x_0^1, x_0^2\right)\left\langle x^1 - x_0^1\right\rangle \\ & + \left(D_1 f\left(x_0^1, x^2\right) - D_1 f\left(x_0^1, x_0^2\right)\right)\left\langle x^1 - x_0^1\right\rangle. \end{aligned}$$

Da $D_1 f$ stetig ist, ist auch diese Differenz in $o(\|x - x_0\|)$. □

In Kombination mit der Kettenregel erhalten wir

Theorem 6.31. *Seien $\Omega \subset \mathbb{R}^n$ und $\Omega' \subset \mathbb{R}^m$ offen. Sei F ein Banachraum. Seien $f\colon \Omega \to F$ und $g = \left(g^i\right)_i : \Omega' \to \Omega \subset \mathbb{R}^n$ differenzierbar:*

$$\begin{tikzcd} \Omega' \arrow[r, "g"'] \arrow[rr, bend left, "h"] & \Omega \arrow[r, "f"'] & F \\[-20pt] \cap & \cap & \\[-20pt] \mathbb{R}^m & \mathbb{R}^n & \end{tikzcd}$$

Dann ist auch $h = f \circ g$ differenzierbar und es gilt

$$Dh = \sum_{i=1}^{n} D_i f(g)\left\langle Dg^i\right\rangle,$$

d. h. mit Argumenten $x' \in \Omega'$ und $u \in \mathbb{R}^m$ erhalten wir

$$Dh(x')\langle u\rangle = \sum_{i=1}^{n} D_i f(g(x'))\left\langle Dg^i(x')\langle u\rangle\right\rangle$$

und für $F = \mathbb{R}^l$ und $1 \leq j \leq m$ sowie $1 \leq k \leq l$

$$\frac{\partial h^k}{\partial x^j}(x') = \sum_{i=1}^{n} \frac{\partial f^k}{\partial x^i}(g(x'))\frac{\partial g^i}{\partial x'^j}(x').$$

Beweis. Nach Kettenregel gilt $Dh = Df(g)\langle Dg\rangle$. Explizit erhalten wir

$$\begin{aligned} Dh(x)\langle u\rangle &= Df(g(x))\langle Dg(x)\langle u\rangle\rangle = Df(g(x))\left\langle \left(Dg^1(x), \ldots, Dg^n(x)\right)\langle u\rangle\right\rangle \\ &= \sum_{i=1}^{n} D_i f(g(x))\left\langle \pi_i\left(Dg^1(x)\langle u\rangle, \ldots, Dg^n(x)\langle u\rangle\right)\right\rangle \\ &= \sum_{i=1}^{n} D_i f(g(x))\left\langle Dg^i(x)\langle u\rangle\right\rangle \end{aligned}$$

wie behauptet. □

Bemerkung 6.32 (Jacobimatrix)**.** Seien $\Omega \subset \mathbb{R}^n$ offen und $f\colon \Omega \to \mathbb{R}^m$ differenzierbar. Nach Korollar 6.27 gilt für $x_0 \in \Omega$ und $u \in \mathbb{R}^n$

$$Df(x_0)\langle u\rangle = \sum_{i=1}^{n} D_i f(x_0)u^i \equiv \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x_0)u^i.$$

Die j-te Komponente, $1 \leq j \leq m$, davon lautet $\sum\limits_{i=1}^{n} \frac{\partial f^j}{\partial x^i}(x_0)u^i$.

Die Ableitung können wir bezüglich der kanonischen Basis des $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ als Matrixmultiplikation darstellen. Die die lineare Abbildung $Df(x_0)$ darstellende

Matrix heißt *Jacobimatrix.* Wir bezeichnen sie mit $\left(\frac{\partial f^j}{\partial x^i}(x_0)\right)_{\substack{1\le j\le m\\1\le i\le n}}$ oder $J_f(x_0)$. In Matrizendarstellung erhalten wir

$$\left(\frac{\partial f^j}{\partial x^i}\right)_{i,j} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} \end{pmatrix}.$$

Dann erhalten wir $Df(x_0)\langle u\rangle$ in Koordinaten als Matrixprodukt zwischen der Jacobimatrix $J_f(x_0)$ und dem Spaltenvektor $u = \left(u^i\right)_i \in \mathbb{R}^n$.

Ist $n = m$, so heißt $\det J_f(x_0) \equiv |J_f(x_0)| \equiv \left|\frac{\partial f^j}{\partial x^i}\right| \equiv \left|\frac{df}{dx}\right|$ *Funktionaldeterminante.*

Aus Theorem 6.31 erhalten wir für die Jacobimatrizen

$$J_h(x) = J_f(g(x)) \cdot J_g(x).$$

Mo 29.06.2020

Wir zeigen einen eindimensionalen Hebbarkeitssatz.

Lemma 6.33. *Sei $\Omega \subset \mathbb{R}$ offen mit $0 \in \Omega$. Seien $f, g\colon \Omega \to \mathbb{R}$ stetig und f in $\Omega \setminus \{0\}$ stetig differenzierbar mit $f' = g$ in $\Omega \setminus \{0\}$. Dann ist f in ganz Ω stetig differenzierbar und es gilt dort $f' = g$.*

Beweis. Wir dürfen ohne Einschränkung annehmen, dass Ω ein Intervall ist. Es genügt zu zeigen, dass f in 0 differenzierbar ist und $f'(0) = g(0)$ erfüllt, dass also

$$f(x) - f(0) - g(0)x = o(|x|) \quad \text{für alle } x \in \Omega$$

gilt. Dazu zeigen wir, dass es zu beliebigen $\varepsilon > 0$ ein $\delta > 0$ mit

$$|f(x) - f(0) - g(0)x| \le \varepsilon \cdot |x|$$

für alle $|x| < \delta$ gibt.

Definiere dazu $\varphi\colon [0, 1] \to \mathbb{R}$ durch

$$\varphi(t) := f(x) - f(tx) - g(0)(1 - t)x.$$

Dann ist φ auf $[0, 1]$ stetig und in $(0, 1)$ differenzierbar. Der Mittelwertsatz ist also anwendbar. Es gelten

$$\begin{aligned}
\varphi(1) &= 0, \\
|\varphi(0)| &= |f(x) - f(0) - g(0)x| \overset{!}{\le} \varepsilon \cdot |x|, \\
|\varphi(0)| &= |\varphi(0) - \varphi(1)| \\
&\le \sup_{t\in(0,1)} |\varphi'(t)| \cdot 1 \\
&= \sup_{t\in(0,1)} | - f'(tx)x - g(0)(-1)x| \\
&= \sup_{t\in(0,1)} |f'(tx) - g(0)| \cdot |x| \\
&\le \sup_{|y|<\delta} |f'(y) - g(0)| \cdot |x| \quad \text{für } |x| \le \delta, \\
&= \sup_{|y|<\delta} |g(y) - g(0)| \cdot |x|.
\end{aligned}$$

Die Behauptung folgt nun aus der Stetigkeit von g. ☐

Lemma 6.34. *Sei $\Omega \subset \mathbb{R}^n$ offen. Seien $f\colon \Omega \to \mathbb{R}$ und $g\colon \Omega \to \mathbb{R}^n$ stetig. Sei $H \subset \mathbb{R}^n$ eine Hyperebene, also ein $n-1$-dimensionaler affiner Unterraum. Sei f in $\Omega \setminus H$ stetig differenzierbar mit $Df = g$. Dann ist f in ganz Ω stetig differenzierbar und es gilt dort $Df = g$.*

Beweis. Nach Theorem 6.30 genügt es zu zeigen, dass f in Ω partiell differenzierbar ist und dort $f_i = g\langle e_i\rangle$ gilt. Dies folgt in einem gedrehten Koordinatensystem, in dem keiner der Basisvektoren e_i in H liegt, aus Lemma 6.33. □

Richtungsableitung.

Definition 6.35. Seien E, F Banachräume und $\Omega \subset E$ offen. Sei $e \in E$ mit $\|e\| = 1$. Sei $f\colon \Omega \to F$ eine Funktion. Dann heißt f in x_0 *in Richtung e differenzierbar,* falls die Funktion $\varphi(t) := f(x_0 + te)$, die für kleine $\varepsilon > 0$ in $(-\varepsilon, \varepsilon)$ definiert ist, in $t = 0$ differenzierbar ist. Dann heißt $\dot\varphi(0)$ *Richtungsableitung* von f im Punkt x_0 in Richtung e. Wir schreiben dafür $D_e f(x_0)$, $\nabla_e f(x_0)$ oder $\frac{\partial f}{\partial e}$.

Bemerkung 6.36.

(i) Ist $E = \mathbb{R}^n$, so sind die partiellen Ableitungen gerade die Richtungsableitungen in Richtung e_i, $f_i(x_0) = D_{e_i} f(x_0)$.
(ii) Ist f in x_0 differenzierbar, so gilt nach Kettenregel $D_e f(x_0) = Df(x_0)\langle e\rangle$ für alle $e \in E$ mit $\|e\| = 1$.
(iii) Gelegentlich erlaubt man in der Definition der Richtungsableitung auch beliebige $e \neq 0$.
(iv) Die Funktion aus Beispiel 6.29 ist im Ursprung in jede Richtung differenzierbar, jedoch nicht differenzierbar.

6.5. **Ableitungen höherer Ordnung.** In der folgenden Definition werden wir höhere Ableitungen induktiv definieren und zur besseren Anschauung die zweite und dritte Ableitung zusätzlich explizit definieren. Beachte dabei, dass wir die bisherige Definition von Differenzierbarkeit anwenden können, da auch $L(E, F)$, $L(E, L(E, F))$, $\ldots$ mit der Operatornorm wieder Banachräume sind.

Definition 6.37. Seien E, F Banachräume, $\Omega \subset E$ offen und $f\colon \Omega \to F$ eine Abbildung. Sei $x_0 \in \Omega$ beliebig.

(i) Wir definieren $D^0 f := f\colon \Omega \to F$ als die 0-te Ableitung von f. Ist f in $U \subset \Omega$ differenzierbar, so definieren wir die *erste Ableitung* von f durch $D^1 f := Df\colon U \to L(E, F)$.
(ii) Ist Df in einer Umgebung U von x_0 definiert und die Abbildung $Df\colon U \to L(E, F)$ in x_0 differenzierbar, so heißt $D(Df)(x_0) \equiv D^2 f(x_0)\colon E \to L(E, F)$, also $D^2 f(x_0) \in L(E, L(E, F))$ *zweite Ableitung* von f in x_0.
(iii) Ist $D^2 f\colon U \to L(E, L(E, F))$ in einer Umgebung U von x_0 definiert und in x_0 differenzierbar, so heißt die Ableitung dieser Funktion *dritte Ableitung* von f in x_0: $D(D^2 f)(x_0) \equiv D^3 f(x_0) \equiv DDDf(x_0) \in L(E, L(E, L(E, F)))$.
(iv) Ist die k-te Ableitung $D^k f$ von f in einer Umgebung U von x_0 definiert und ist $D^k f\colon U \to L(E, \ldots, L(E, F)\ldots)$ in x_0 differenzierbar, so heißt die Ableitung dieser Funktion, $D\left(D^k f\right)(x_0) \equiv D^{k+1} f(x_0)$, die *$(k+1)$-ste Ableitung* von f in x_0.
(v) Existiert die k-te Ableitung von f in allen $x \in \Omega$, so heißt f in Ω *k-mal differenzierbar.*
(vi) Existiert die k-te Ableitung von f in allen $x \in \Omega$ und ist sie in Ω stetig, so heißt f in Ω *k-mal stetig differenzierbar* oder von der *Klasse* $C^k(\Omega, F)$.
(vii) Sind $k \in \mathbb{N}$, $f \in C^k(\Omega, F)$ und alle Ableitungen von f bis zur Ordnung k auf $\overline{\Omega}$ stetig und beschränkt fortsetzbar, so ist f von der Klasse $C^k\left(\overline{\Omega}, F\right)$, $f \in C^k\left(\overline{\Omega}, F\right)$.
(viii) Ist $f \in C^k(\Omega, F)$ für alle $k \in \mathbb{N}$, so heißt f *glatt* oder von der *Klasse* $C^\infty(\Omega, F)$.

Bemerkung 6.38.

(i) Da $L(E, F)$ selbst wieder ein Banachraum ist, können wir die Definition von Differenzierbarkeit auf die Abbildung $Df\colon U \to L(E, F)$ anwenden.

(ii) Aus der linearen Algebra, siehe auch Bemerkung 6.6, kennen wir den (normtreuen) Isomorphismus
$$L(E, L(E, F)) \cong L_2(E; F),$$
$A \mapsto a$ mit
$$a(u, v) := \underbrace{(A\langle u\rangle)}_{\in L(E,F)}\langle v\rangle \quad \text{für alle } u, v \in E.$$

Damit erhalten wir direkt

Proposition 6.39. *Sei $\Omega \subset E$ im Banachraum E offen. Sei $f\colon \Omega \to F$ in x_0 zweimal differenzierbar. Dann ist, vermöge der Identifikation*
$$L(E, L(E, F)) \cong L_2(E; F),$$
$D^2f(x_0) \in L_2(E; F)$ und es gilt
$$D^2f(x_0)(u, v) \equiv D^2f(x_0)\langle u, v\rangle = \left(D^2f(x_0)\langle u\rangle\right)\langle v\rangle.$$

Definition 6.40 (C^k-Normen)**.** Seien E, F Banachräume und $\Omega \subset E$ offen. Dann definieren wir auf $C^k\left(\overline{\Omega}, F\right)$ eine Norm, die C^k-Norm, durch
$$\|f\|_{C^k(\Omega,F)} := \sum_{i=0}^{k} \left\|D^i f\right\|_{C^0(\Omega, L_i(E;F))}$$
mit $L_0(E, F) \equiv F$.

Genauso wie im eindimensionalen Fall sehen wir, dass die Räume $C^k\left(\overline{\Omega}, F\right)$ Banachräume sind.

Theorem 6.41. *Seien E, F Banachräume und sei $\Omega \subset E$ offen. Sei $k \in \mathbb{N}$. Dann ist $C^k\left(\overline{\Omega}, F\right)$ mit der C^k-Norm ein Banachraum.*

Beweis. Übung. □

Lemma 6.42. *Seien E, F Banachräume und $\Omega \subset E$ offen. Sei $f\colon \Omega \to F$ in $x_0 \in \Omega$ zweimal differenzierbar. Sei U eine Umgebung von x_0, in der f einmal differenzierbar ist. (U existiert nach Definition von zweimaliger Differenzierbarkeit.) Sei $v \in E$ fest. Definiere $\varphi\colon U \to F$ durch $\varphi(x) = Df(x)\langle v\rangle$. Dann ist φ in x_0 differenzierbar und es gilt für alle $u \in E$*
$$D\varphi(x_0)\langle u\rangle = D^2f(x_0)\langle v, u\rangle.$$

Beweis. Wir schreiben φ als Komposition der linearen Abbildung
$$\Phi\colon L(E, F) \to F \quad \text{mit} \quad A \mapsto Av$$
und der Abbildung $Df\colon U \to L(E, F)$: $\varphi = \Phi \circ Df$. Aus $\Phi(A) - \Phi(A_0) = (A - A_0)\langle v\rangle = D\Phi(A_0)\langle A - A_0\rangle$ wie bei jeder stetigen linearen Abbildung folgt hier $D\Phi(A_0)\langle B\rangle = B\langle v\rangle$. Nach Kettenregel ist φ somit in x_0 differenzierbar und es gilt $D\varphi = D\Phi(Df) \circ D^2f$ und in expliziterer Form
$$\begin{aligned} D\varphi(x_0)\langle u\rangle &= D\Phi(Df(x_0))\langle D(Df(\cdot)\langle\cdot\rangle)(x_0)\langle u\rangle\rangle \\ &= D\Phi(Df(x_0))\left\langle D^2f(x_0)\langle\cdot\rangle\langle u\rangle\right\rangle = D^2f(x_0)\langle v, u\rangle, \end{aligned}$$
wobei $Df(\cdot)\langle\cdot\rangle$ eine Funktion $x \mapsto L(E, F)$ bezeichnet. □

Do 02.07.2020

Das folgende Resultat ist eine Variante des noch folgenden Satzes von Schwarz. Beachte, das hier die Existenz der zweiten Ableitung in x_0 und nicht nur die partieller Ableitungen gefordert wird.

Theorem 6.43. *Seien E, F Banachräume, $\Omega \subset E$ offen und sei $f\colon \Omega \to F$ in $x_0 \in \Omega$ zweimal differenzierbar. Dann ist die bilineare Abbildung $D^2 f(x_0)$ symmetrisch, d. h. es gilt für alle $u, v \in E$*

$$D^2 f(x_0)\langle u, v\rangle = D^2 f(x_0)\langle v, u\rangle.$$

Beweis. Wir behaupten, dass

$$f(x_0+u+v)-f(x_0+u)-f(x_0+v)+f(x_0) = D^2 f(x_0)\langle u, v\rangle + o\left((\|u\|+\|v\|)^2\right) \tag{6.2}$$

gilt, falls u, v so gewählt sind, dass alle Ausdrücke wohldefiniert sind, also z. B. $u, v \in B_r(0)$ mit $r > 0$, so dass $B_{2r}(x_0) \subset \Omega$ gilt, was wir im Folgenden annehmen werden.

Aus (6.2) folgt die Behauptung: Die linke Seite ist in u und v symmetrisch. Somit gilt

$$\left\|D^2 f(x_0)\langle u, v\rangle - D^2 f(x_0)\langle v, u\rangle\right\| = o\left((\|u\|+\|v\|)^2\right)$$

und für alle *Einheitsvektoren* (= Vektoren der Länge eins) $\overline{u}, \overline{v} \in E$ und alle $|t| < r$ erhalten wir daraus

$$t^2 \left\|D^2 f(x_0)\langle \overline{u}, \overline{v}\rangle - D^2 f(x_0)\langle \overline{v}, \overline{u}\rangle\right\| = o\left(t^2\right).$$

Wir dividieren nun durch $t^2 \neq 0$ und erhalten im Grenzwert $t \to 0$ die Behauptung.

Somit ist noch (6.2) zu zeigen. Definiere für $0 \le t \le 1$ die Funktion

$$\varphi(t) := f(x_0 + u + tv) - f(x_0 + tv).$$

Die behauptete Gleichheit werden wir dann aus einer Abschätzung für $\varphi(1)-\varphi(0)-\varphi'(0)$ erhalten, vergleiche das Ende dieses Beweises.

Es gilt der Mittelwertsatz für auf Teilmengen von $\mathbb{R}$ definierte Funktionen in der Form

$$\|\varphi(1) - \varphi(0) - \varphi'(0)\| \le \sup_{0\le t\le 1} \|\varphi'(t) - \varphi'(0)\|.$$

Die rechte Seite wollen wir weiter abschätzen. Nach Kettenregel gilt

$$\varphi'(t) = Df(x_0 + u + tv)\langle v\rangle - Df(x_0 + tv)\langle v\rangle.$$

Wir addieren eine Null und schreiben die beiden Differenzen aufgrund der Differenzierbarkeit mit Hilfe der zweiten Ableitung, benutzen die Linearität der Ableitungen im „$\langle\cdot\rangle$"-Argument und erhalten

$$\begin{aligned}
\varphi'(t) &= \{Df(x_0 + u + tv) - Df(x_0) - [Df(x_0 + tv) - Df(x_0)]\}\langle v\rangle \\
&= \left\{D^2 f(x_0)\langle u + tv\rangle + o(\|u + tv\|) - D^2 f(x_0)\langle tv\rangle - o(\|tv\|)\right\}\langle v\rangle \\
&= D^2 f(x_0)\langle u\rangle\langle v\rangle + o(\|u + tv\|)\|v\| + o(\|tv\|)\|v\|.
\end{aligned}$$

Wie man sich beispielsweise mit Hilfe der ε-Notation leicht überzeugt, sind beide Fehlerterme in $o\left((\|u\|+\|v\|)^2\right)$. Hieraus folgt einerseits

$$\sup_{0\le t\le 1} \|\varphi'(t) - \varphi'(0)\| \le o\left((\|u\|+\|v\|)^2\right),$$

andererseits können wir den Ausdruck für $\varphi'(t)$, ausgewertet an der Stelle $t = 0$, auch in der Folgerung aus dem Mittelwertsatz einsetzen und erhalten

$$\varphi(1) - \varphi(0) = D^2 f(x_0)\langle u, v\rangle + o\left((\|u\|+\|v\|)^2\right).$$

Nach Definition von φ ist dies aber gerade die Behauptung. □

Auch für zweite Ableitungen bekommen wir einen Zusammenhang zwischen $D^2 f$ und zweiten partiellen Ableitungen.

Proposition 6.44. *Sei $\Omega \subset \mathbb{R}^n$ offen und sei F ein Banachraum. Sei $f\colon \Omega \to F$ in $x_0 \in \Omega$ zweimal differenzierbar und seien $1 \le i, j \le n$. Dann gelten*

(i)

$$D_iD_jf(x_0) = D^2f(x_0)\langle e_j, e_i\rangle$$

(ii) und

$$D_iD_jf(x_0) = D_jD_if(x_0),$$

d. h. die zweiten partiellen Ableitungen kommutieren.

Die Reihenfolge der linearen Argumente ist eine Konvention und kommt in der Literatur auch anders herum vor.

Beweis.

(i) Für die ersten partiellen Ableitungen von f haben wir in Proposition 6.26 bereits $D_jf(\cdot) = Df(\cdot) \circ \chi_j$, $1 \le j \le n$, mit der Standardeinbettung $\mathbb{R} \to \mathbb{R}^n$ in die j-te Komponente gezeigt. Daraus erhalten wir $D_jf(\cdot) = Df(\cdot)\langle e_j\rangle$ durch Anwenden auf 1. Wir differenzieren dies nochmals und benutzen dabei nochmals Proposition 6.26 und Lemma 6.42

$$D_iD_jf(x_0) = D(Df(\cdot)\langle e_j\rangle)(x_0)\langle e_i\rangle = D^2f(x_0)\,\langle e_j, e_i\rangle\,.$$

(ii) Dies folgt nun unmittelbar aus Theorem 6.43. Es gilt nämlich

$$D_iD_jf(x_0) = D^2f(x_0)\langle e_j, e_i\rangle = D^2f(x_0)\langle e_i, e_j\rangle = D_jD_if(x_0). \qquad \square$$

Existieren die zweiten partiellen Ableitungen und sind stetig, so existieren auch die zweiten Ableitungen. Als Vorbereitung dafür zeigen wir

Lemma 6.45. *Sei F ein Banachraum. Dann sind $L(\mathbb{R}^n, F)$ und $(L(\mathbb{R}, F))^n \cong F^n$ isomorphe Banachräume, d. h. es gibt eine beschränkte lineare Abbildung zwischen ihnen mit beschränkter Inversen.*

Ist $\chi_i\colon \mathbb{R} \hookrightarrow \mathbb{R}^n$ die kanonische Einbettung in die i-te Komponente und $\pi_i\colon \mathbb{R}^n \to \mathbb{R}$ die Projektion auf den i-ten Faktor, so ist ein möglicher solcher Banachraumisomorphismus durch $\Phi\colon A \mapsto (A_i)_{1\le i\le n}$ mit $A_i = A \circ \chi_i$ und die Inverse dazu ist durch $A = \sum\limits_{i=1}^{n} A_i \circ \pi_i$ gegeben.

Beweis. Direktes Nachrechnen. Beachte, dass wir nicht behaupten, dass der Banachraumisomorphismus normerhaltend ist. $\square$

Theorem 6.46. *Sei $\Omega \subset \mathbb{R}^n$ offen und sei F ein Banachraum. Sei $f\colon \Omega \to F$ eine Abbildung. Dann ist f genau dann zweimal stetig in Ω differenzierbar, wenn f in Ω zweimal stetig partiell differenzierbar ist.*

Beweis. Nach Theorem 6.30 (stetige und stetige partielle Differenzierbarkeit sind äquivalent) ist die Behauptung für die ersten Ableitungen bekannt.

„$\Rightarrow$“: Sei f zunächst zweimal stetig differenzierbar. Nach Proposition 6.44 gilt $D_iD_jf(x_0)\langle\cdot,\cdot\rangle = D^2f(x_0)\langle e_i, e_j\rangle$ und wir erhalten die Behauptung für die zweiten Ableitungen.

„$\Leftarrow$“: Sei nun f umgekehrt zweimal stetig partiell differenzierbar, D_if und D_jD_if, $1 \le i, j \le n$ existieren also und sind stetig. Nach Theorem 6.30 sind somit f und D_if in Ω stetig differenzierbar.

Wir verwenden nun Φ aus Lemma 6.45 und erhalten eine Abbildung $\Phi \circ Df\colon \Omega \to F^n$. Nach Proposition 6.26 sind die Komponenten dieser Abbildung gerade die partiellen Ableitungen D_if.

Nach Voraussetzung sind die Funktionen D_if, die ja gerade die Komponenten der Abbildung $\Phi \circ Df$ sind, in Ω differenzierbar. Somit ist nach Proposition 6.13 auch die Abbildung $\Phi \circ Df$ in Ω differenzierbar. Wir verknüpfen diese Funktion noch mit der linearen Abbildung Φ^{-1} und erhalten, dass $Df = \Phi^{-1} \circ \Phi \circ Df$ in Ω differenzierbar ist.

Zur Stetigkeit der Abbildung $x \mapsto D^2 f(x)$: Seien $u = \left(u^i\right)_{1\le i\le n}, v = \left(v^j\right)_{1\le j\le n} \in \mathbb{R}^n$. Es gilt

$$\begin{aligned} D^2 f(x)\langle u, v\rangle &= D^2 f(x)\left\langle \sum_{i=1}^n u^i e_i, \sum_{j=1}^n v^j e_j\right\rangle \\ &= \sum_{i,j=1}^n u^i \cdot v^j \cdot D^2 f(x)\left\langle e_i, e_j\right\rangle = \sum_{i,j=1}^n u^i \cdot v^j \cdot D_j D_i f(x) \end{aligned}$$

und somit erhalten wir

$$D^2 f(x) = \sum_{i,j=1}^n \pi_i(\cdot) \cdot \pi_j(\cdot) \cdot D_j D_i f(x).$$

Hieraus folgt die Stetigkeit. □

Korollar 6.47 (Satz von Schwarz). *Sei $\Omega \subset \mathbb{R}^n$ offen und sei F ein Banachraum. Sei $f\colon \Omega \to F$ zweimal stetig partiell differenzierbar, so sind die zweiten partiellen Ableitungen symmetrisch, d. h. für alle $1 \le i, j \le n$ und alle $x \in \Omega$ gilt*

$$D_i D_j f(x) = D_j D_i f(x).$$

Beweis. Nach Theorem 6.46 ist f in Ω zweimal differenzierbar. Nun können wir Proposition 6.44 anwenden und erhalten die behauptete Vertauschbarkeit der partiellen Ableitungen. □

Bemerkung 6.48.

(i) Sei $\Omega \subset \mathbb{R}^n$ offen und sei $f\colon \Omega \to \mathbb{R}$ zweimal differenzierbar. Dann erhalten wir für alle $x \in \Omega$ und alle $u, v \in \mathbb{R}^n$ nach Proposition 6.44 und der Linearität oder aus dem Beweis von Theorem 6.46

$$D^2 f(x)\langle u, v\rangle = \sum_{i,j=1}^n D_i D_j f(x) u^j v^i.$$

(ii) Wie in der linearen Algebra können wir auch hier die symmetrische Bilinearform $D^2 f(x)$ durch eine Matrix darstellen, nämlich durch $(D_i D_j f(x))_{1\le i,j\le n}$, die *Hessematrix* oder *Hessesche* von f.

Beispiele 6.49.

(i) Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ durch $(x, y) \mapsto x^2 y$ gegeben. Dann ist

$$f_{12}((x, y)) = D_2(2xy) = 2x = D_1\left(x^2\right) = f_{21}((x, y)).$$

(ii) Die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x, y) = \begin{cases} \dfrac{xy\left(x^2 - y^2\right)}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0), \end{cases}$$

ist im Ursprung zweimal partiell differenzierbar, jedoch sind die zweiten partiellen Ableitungen dort weder stetig noch symmetrisch.

Beweis. Übung. □

(iii) Sei $B\colon \mathbb{R}^n \times \mathbb{R}^m \equiv \mathbb{R}^{n+m} \to \mathbb{R}^k$ bilinear. B ist dann automatisch auch stetig. Dann ist $B \in C^\infty$ und $D^3 B \equiv 0$.

Beweis. Es gilt

$$DB(u,v)\langle(a,b)\rangle = B(a,v) + B(u,b).$$

Die Abbildung $DB(\cdot,\cdot)\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ ist linear. Somit gilt

$$\begin{aligned} D(DB(\cdot,\cdot)\langle(a,b)\rangle)(u,v)\langle(c,d)\rangle &= D^2B(u,v)\langle(a,b),(c,d)\rangle \\ &= B(a,d) + B(c,b). \end{aligned}$$

Dies ist von (u,v) unabhängig. Also ist $D^3B \equiv 0$. □

Mo 06.07.2020

(iv) Sei E ein Banachraum und $GL(E) \subset L(E)$ die (offene) Menge der Operatoren in $L(E)$ mit Inverser Abbildung in $L(E)$. Dann ist

$$\Phi\colon GL(E) \ni A \mapsto A^{-1} \in GL(E) \subset L(E)$$

in $GL(E)$ differenzierbar und es gilt $D\Phi(A)\langle B\rangle = -A^{-1}BA^{-1}$. (Vergleiche dies mit $\left(g^{-1}\right)' = -\frac{g'}{g^2}$.) Es gilt $\Phi \in C^\infty$.

Beweis. Wir zeigen nur die Behauptung über $D\Phi$ und lassen es als Übung, die Glattheit von $D\Phi$ mit Hilfe der Glattheit der in (A,C) bilinearen Abbildung

$$\Psi(A,C)\langle B\rangle := -ABC,$$

der Kettenregel und Induktion zu zeigen.

Differenzierbarkeit: Statt $\Phi(A+B) - \Phi(A) - D\Phi(A)\langle B\rangle \in o(\|B\|)$ zu zeigen, multiplizieren wir die Gleichung von links mit A und erhalten

$$\begin{aligned} o(\|B\|) &\overset{!}{=} A(B+A)^{-1} - AA^{-1} + AA^{-1}BA^{-1} \\ &= \left((B+A)A^{-1}\right)^{-1} - \mathbb{1} + BA^{-1} \\ &= \left(BA^{-1} + \mathbb{1}\right)^{-1} - \mathbb{1} + BA^{-1}. \end{aligned}$$

Die Behauptung folgt nun mit Hilfe der Neumannschen Reihe, denn es gilt

$$\left(\mathbb{1} + BA^{-1}\right)^{-1} = \mathbb{1} - BA^{-1} + \left(BA^{-1}\right)^2 - \left(BA^{-1}\right)^3 \pm \ldots$$

für $\left\|BA^{-1}\right\| < 1$. □

6.6. Taylorsche Formel.

Definition 6.50. Seien E,F Banachräume, $\Omega \subset E$ offen und $f\colon \Omega \to F$ eine k-mal differenzierbare Abbildung. Seien $u \in E$ und $x \in \Omega$. Dann setzen wir

$$D^kf(x)\langle u\rangle^k := D^kf(x)\underbrace{\langle u,\ldots,u\rangle}_{k\text{ Stück}}$$

für $k \in \mathbb{N}_{>0}$ und $D^0f(x)\langle u\rangle^0 = f(x)$.

Lemma 6.51. *Seien E,F Banachräume, $\Omega \subset E$ offen und $f\colon \Omega \to F$ eine k-mal differenzierbare Abbildung. Seien $x,u \in E$ mit $[x,x+u] \subset \Omega$. Dann ist $[0,1] \ni t \mapsto \varphi(t) \equiv f(x+tu) \in F$ ebenfalls k-mal differenzierbar und es gilt*

$$\varphi^{(k)}(t) = D^kf(x+tu)\langle u\rangle^k.$$

Beweis. Kettenregel und Induktion. □

Theorem 6.52 (Taylorsche Formel). *Seien E,F Banachräume, $\Omega \subset E$ offen und $f \in C^p(\Omega,F)$, $p \in \mathbb{N}$. Sei $x \in \Omega$. Dann gilt für alle $u \in E$ mit $x+u \in \Omega$ die Darstellung*

$$f(x+u) = \sum_{k=0}^{p} \frac{1}{k!} D^kf(x)\langle u\rangle^k + o(\|u\|^p).$$

Beweis. Da Ω offen ist, dürfen wir ohne Einschränkung annehmen, dass $\delta > 0$ so gewählt ist, dass $B_\delta(x) \subset \Omega$ gilt. Da die Behauptung nur im Falle, dass $\|u\|$ klein ist, nichttrivial ist, dürfen wir ohne Einschränkung $\|u\| < \delta$ annehmen. Wir betrachten nun φ mit $[0,1] \ni t \mapsto \varphi(t) \equiv f(x+tu) \in F$ wie in Lemma 6.51. Dann ist φ sogar noch in einer kleinen Umgebung von $[0,1] \subset \mathbb{R}$ von der Klasse C^p. Wir wenden nun den eindimensionalen Satz von Taylor, Theorem 4.69, an und erhalten

$$\varphi(1) = \sum_{k=0}^{p} \frac{1}{k!} \varphi^{(k)}(0) + R_p(\varphi, 0)(1)$$

sowie die Restgliedabschätzung

$$\|R_p(\varphi, 0)(1)\| \le \frac{1}{(p-1)!} \sup_{0<t<1} \left\| \varphi^{(p)}(t) - \varphi^{(p)}(0) \right\|.$$

Die Behauptung folgt nun direkt mit Hilfe der Formel für die Ableitungen von φ aus Lemma 6.51. □

Bemerkung 6.53 (Lagrangesche Restgliedabschätzung)**.** Ist $f \in C^{p+1}(\Omega, F)$, so erhalten wir mit einem analogen Vorgehen, diesmal unter Verwendung von Theorem 4.74, die Abschätzung

$$f(x+u) = \sum_{k=0}^{p} \frac{1}{k!} D^k f(x)\langle u\rangle^k + O\left(\|u\|^{p+1}\right).$$

Stellen wir totale Ableitungen mit Hilfe der partiellen Ableitungen dar, so erhalten wir aus der Taylorschen Formel

Korollar 6.54. *Sei $\Omega \subset \mathbb{R}^n$ offen, F ein Banachraum und $f \in C^p(\Omega, F)$ mit $p \ge 1$. Dann gilt für alle $x, x+u \in \Omega$*

$$f(x+u) = \sum_{k=0}^{p} \frac{1}{k!} \sum_{i_1,\ldots,i_k=1}^{n} D_{i_1} \ldots D_{i_k} f(x) u^{i_1} \ldots u^{i_k} + o\left(\|u\|^p\right).$$

6.7. **Lokale Extrema.** Hinreichende und notwendige Bedingungen für lokale Extrema benutzen die aus der linearen Algebra bekannte Definitheit.

Definition 6.55. Sei E ein Banachraum und $a \in L_2(E; \mathbb{R})$ eine symmetrische Bilinearform. Dann heißt a

(i) *gleichmäßig positiv definit,* falls es ein $c > 0$ mit $a(x,x) \ge c$ für alle $x \in E$ mit $\|x\| = 1$ gibt.
(ii) Ist H ein reeller Hilbertraum und $A \in L(H)$ selbstadjungiert, so heißt A *positiv semidefinit, ...,* falls dies für die symmetrische Bilinearform $(x,y) \mapsto \langle Ax, y\rangle$ gilt. Wir schreiben hier ebenso $A \succcurlyeq 0, \ldots$.

Theorem 6.56. *Sei E ein Banachraum, $\Omega \subset E$ offen und $f \in C^2(\Omega, \mathbb{R})$. Sei $x_0 \in \Omega$ eine lokale Extremalstelle von f.*

(i) Dann ist x_0 ein kritischer Punkt von f, d. h. es gilt $Df(x_0) = 0$.
(ii) Weiterhin gilt $D^2 f(x_0) \succcurlyeq 0$ in einem lokalen Minimum bzw.
$D^2 f(x_0) \preccurlyeq 0$ in einem lokalen Maximum.

Beweis.

(i) Sei $u \in E$ beliebig. Dann ist 0 eine lokale Extremalstelle der Funktion φ mit $(-\varepsilon, \varepsilon) \ni t \mapsto \varphi(t) \equiv f(x_0 + tu) \in \mathbb{R}$, die für ein kleines $\varepsilon > 0$ definiert ist. Da φ auf einer offenen Teilmenge von $\mathbb{R}$ definiert ist, folgt $0 = \varphi'(0)$, also aufgrund der Kettenregel $Df(x_0)\langle u\rangle = 0$. Da $u \in E$ beliebig war, folgt $Df(x_0) = 0$ und x_0 ist ein kritischer Punkt von f.

(ii) Aus der Taylorschen Formel folgt nun

$$f(x_0+u) = f(x_0) + \frac{1}{2}D^2 f(x_0)\langle u\rangle^2 + o\left(\|u\|^2\right)$$

für alle u in einem hinreichend kleinen Ball $B_r(0)$, so dass $B_r(x_0) \subset \Omega$ gilt. Wir betrachten ab jetzt ohne Einschränkung den Fall eines lokalen Minimums. Durch Verkleinerung von $r > 0$ können wir erreichen, dass $f(x_0+u) \ge f(x_0)$ für alle $u \in B_r(0)$ gilt. Setzen wir in die Taylorsche Formel nun speziell $u = te$ mit $0 < t < r$ und ein beliebiges $e \in E$ mit $\|e\| = 1$ ein, so folgt

$$0 \le f(x_0+te) - f(x_0) = \frac{1}{2}t^2 D^2 f(x_0)\langle e, e\rangle + o\left(t^2\right).$$

Wir dividieren nun durch t^2, lassen $t \to 0$ und erhalten, da $e \in E$ mit $\|e\| = 1$ beliebig war, die Behauptung. □

Theorem 6.57. *Sei E ein Banachraum. Sei $\Omega \subset E$ offen und $f \in C^2(\Omega)$. Sei $x_0 \in \Omega$ ein kritischer Punkt von f.*

(i) Ist $D^2 f(x_0)$ gleichmäßig positiv definit, so nimmt f in x_0 ein lokales Minimum an.

(ii) Ist $D^2 f(x_0)$ gleichmäßig negativ definit, so nimmt f in x_0 ein lokales Maximum an.

Ist $\dim E < \infty$, also z. B. $E = \mathbb{R}^n$, so genügt die positive bzw. negative Definitheit als Voraussetzung.

Beweis. Sei $D^2 f(x_0)$ ohne Einschränkung gleichmäßig positiv definit, d. h. es gibt ein $c > 0$, so dass $D^2 f(x_0)\langle u, u\rangle \ge 2c\|u\|^2$ für alle $u \in E$ gilt. Sei $r > 0$, so dass $B_r(x_0) \subset \Omega$ gilt. Sei $\|u\| < r$. Dann folgt aus der Taylorschen Formel

$$f(x_0+u) \ge f(x_0) + c\|u\|^2 + o\left(\|u\|^2\right).$$

Ist r klein genug, so erhalten wir $\left|o\left(\|u\|^2\right)\right| \le \frac{c}{2}\|u\|^2$ und es folgt

$$f(x_0+u) \ge f(x_0) + \frac{c}{2}\|u\|^2$$

für alle $u \in B_r(0)$. Dies impliziert die Behauptung. □

7. Existenzsätze

7.1. **Banachscher Fixpunktsatz.** Bereits bekannt.

Für später benötigen wir jedoch noch die folgende Variante des Banachschen Fixpunktsatzes.

Theorem 7.1. *Sei E ein Banachraum und $f\colon \bar{B}_r(0) \to E$ eine Kontraktion mit Kontraktionskonstante κ. Gelte $\|f(0)\| \le r(1-\kappa)$. Dann besitzt f einen Fixpunkt.*

Im Unterschied zum Banachschen Fixpunktsatz nehmen wir hier nicht an, dass $f\colon \bar{B}_r(0) \to \bar{B}_r(0)$ eine Selbstabbildung ist. Trotzdem gehen wir ähnlich wie beim Beweis des Banachschen Fixpunktsatzes vor und iterieren die Abbildung f.

Beweis. Wir definieren induktiv $x_0 := 0$ und $x_{n+1} := f(x_n)$ für alle $n \in \mathbb{N}$. Solange $x_n \in \bar{B}_r(0)$ gilt, ist auch x_{n+1} wohldefiniert. Sei $N \in \mathbb{N}$, so dass $x_n \in \overline{B_r(0)}$ für alle $0 \le n \le N$ gilt. Wie beim Beweis des Banachschen Fixpunktsatzes erhalten wir für alle $n \le N+1$

$$\begin{aligned}
\|x_n - x_{n-1}\| &= \|f(x_{n-1}) - f(x_{n-2})\| \le \kappa \cdot \|x_{n-1} - x_{n-2}\| \\
&= \kappa \cdot \|f(x_{n-2}) - f(x_{n-3})\| \le \kappa^2 \cdot \|x_{n-2} - x_{n-3}\| \le \ldots \\
&\le \kappa^{n-1} \cdot \|x_1 - x_0\| = \kappa^{n-1} \cdot \|f(0)\|, \\
\|x_{N+1}\| &= \|x_0 - x_{N+1}\| \le \|x_0 - x_1\| + \ldots + \|x_N - x_{N+1}\|
\end{aligned}$$

$$\leq \sum_{i=0}^{N} \kappa^i \cdot \|f(0)\| \leq \frac{1}{1-\kappa} \cdot \|f(0)\| \leq r.$$

Hieraus folgt, dass auch noch $x_{N+1} \in \bar{B}_r(0)$ gilt. Induktiv folgt also $x_n \in \bar{B}_r(0)$ für alle $n \in \mathbb{N}$.

Mit einer Abschätzung ähnlich wie oben oder wie im Beweis des Banachschen Fixpunktsatzes folgt nun, dass $(x_n)_n$ eine Cauchyfolge ist und der Limes ist der gesuchte Fixpunkt in $\bar{B}_r(0)$. □

Do 09.07.2020

7.2. **Satz über die Umkehrabbildung.** Wir beginnen mit etwas Vorbereitung. Wichtig wird in diesem Kapitel die Menge der Diffeomorphismen.

Definition 7.2 (Diffeomorphismus)**.** Seien E, F Banachräume und $U \subset E$ sowie $V \subset F$ offen.

(i) Eine Abbildung $f\colon U \to V$ heißt ein C^k-*Diffeomorphismus,* $k \in \mathbb{N}_{>0} \cup \{\infty\}$, falls f bijektiv ist und f und f^{-1} von der Klasse C^k sind. Wir bezeichnen die Menge aller C^k-Diffeomorphismen von U nach V mit $\operatorname{Diff}^k(U, V)$.
(ii) Eine Abbildung $f\colon U \to F$ heißt *lokaler* C^k-*Diffeomorphismus,* falls f eine offene Abbildung ist (dies ist nötig um später die Definition von Diff^k anwenden zu können) und jeder Punkt $x_0 \in U$ eine offene Umgebung $B_r(x_0) \subset U$ besitzt, so dass $f \in \operatorname{Diff}^k(B_r(x_0), f(B_r(x_0)))$ gilt. Wir bezeichnen die Menge aller lokalen C^k-Diffeomorphismen von U nach F mit $\operatorname{Diff}^k_{\text{loc}}(U, F)$.
(iii) Die (möglicherweise leere) Menge aller topologischen (linearen) Isomorphismen (oder linearen Homöomorphismen) von E nach F bezeichnen wir mit $L_{\text{top}}(E, F)$. (Eine Abbildung $\varphi \in L_{\text{top}}(E, F)$ ist linear, stetig und bijektiv, ebenso φ^{-1}. Weiterhin bemerken wir, dass $L_{\text{top}}(E, E) = GL(E)$ gilt.)

Theorem 7.3 (Satz über die Umkehrfunktion, Satz von der inversen Abbildung, "Inverse function theorem")**.**
Seien E, F Banachräume, sei $\Omega \subset E$ offen und $f \in C^k(\Omega, F)$ mit $k \geq 1$. Sei $x_0 \in E$. Ist $Df(x_0) \in L_{top}(E, F)$, so gibt es eine offene Umgebung $U = B_r(x_0) \subset \Omega$, so dass $V = f(U)$ offen ist und $f \in \operatorname{Diff}^k(U, V)$ gilt.

Gilt also $Df(x) \in L_{top}(E, F)$ für alle $x \in \Omega$, so folgt $f \in \operatorname{Diff}^k_{loc}(\Omega, F)$.

Beweis.

(i) Wir nehmen ohne Einschränkung an, dass $E = F$ gilt. Sonst wenden wir das Resultat auf die Abbildung $g := (Df(x_0))^{-1} \circ f$ an. Es gelten $g \in C^k(\Omega, E)$ sowie $Dg(x_0) = \text{id}$. Ist g ein C^k-Diffeomorphismus, so auch f.
(ii) Sei $f \in C^k(\Omega, E)$ und $Df(x_0) \in GL(E)$. Wir erinnern daran, dass $GL(E) \subset L(E)$ eine offene Teilmenge ist und dass die Abbildung $\Phi\colon GL(E) \to GL(E)$ mit $A \mapsto A^{-1}$ von der Klasse C^∞ ist.

   Wegen $f \in C^1$, also $Df \in C^0$, gibt es eine Umgebung $U = B_r(x_0)$, so dass für alle $x \in B_r(x_0)$ auch $Df(x) \in GL(E)$ gilt.

   Wir nehmen nun an, dass $f\colon U \to V := f(B_r(x_0))$ ein Homöomorphismus ist (und zeigen dies später noch). Sei $g := (f|_U)^{-1}$. Nach Theorem 6.12, dessen Voraussetzungen in jedem Punkt $x \in U$ erfüllt sind, ist g differenzierbar und für alle $x \in U$ gilt

$$Dg(f(x)) = (Df(x))^{-1} = \Phi \circ Df(x),$$

   also $Dg = \Phi \circ Df \circ g$. Da wir vorausgesetzt haben, dass $f|_U$ ein Homöomorphismus ist, ist $g = (f|_U)^{-1}$ stetig und aufgrund der Formel für Dg gilt $g \in C^1$.

   Da $Df \in C^{k-1}$ ist, erhalten wir per Induktion $g \in C^k(V, E)$.

Bis auf den Beweis der nachfolgenden Behauptung beweist dies unser Theorem.

- Behauptung: Es gibt ein $r > 0$, so dass $f|_U : U \to V$ mit $U = B_r(x_0)$ und $V = f(B_r(x_0))$ ein Homöomorphismus ist. Dies werden wir in den noch folgenden Punkten zeigen.

(iii) Surjektivität und Stetigkeit sind klar.

(iv) Injektivität: Wir gehen ähnlich wie beim Beweis von Theorem 6.12 vor. Wir behaupten, dass $f|_U$ für hinreichend kleine $r > 0$ injektiv ist. Seien $x, y \in B_r(x_0)$ beliebig. Dann folgt

$$f(x) - f(y) = Df(y)\langle x - y\rangle + o(\|x - y\|).$$

In einer hinreichend kleinen Umgebung von x_0 gibt es eine Konstante $c > 0$, so dass für y aus dieser Umgebung $\|Df(y)\langle v\rangle\| \ge c\|v\|$ gilt. (Erklärung dazu: Aufgrund der Stetigkeit von $(Df(x_0))^{-1}$ gibt es eine Konstante $c > 0$, so dass $\left\|(Df(x_0))^{-1}\langle u\rangle\right\| \le \frac{1}{2c}\|u\|$ für alle $u \in E$ gilt. Wir setzen nun insbesondere $u = Df(x_0)\langle v\rangle$ und erhalten $2c \cdot \|v\| \le \|Df(x_0)\langle v\rangle\|$ für alle $v \in E$. Indem wir eine hinreichend kleine Umgebung von x_0 betrachten, dürfen wir für alle y in dieser Umgebung annehmen, dass $\|Df(y) - Df(x_0)\| \le c$ in der Operatornorm gilt. Daher folgt für solche y

$$\begin{aligned}\|Df(y)\langle v\rangle\| &\ge \|Df(x_0)\langle v\rangle\| - \|Df(x_0)\langle v\rangle - Df(y)\langle v\rangle\|\\ &\ge 2c \cdot \|v\| - \|Df(x_0) - Df(y)\| \cdot \|v\| \ge c \cdot \|v\|\end{aligned}$$

wie oben behauptet.)

Indem wir die Umgebung gegebenenfalls weiter verkleinern, können wir erreichen, dass $\|o(\|x - y\|)\| \le \frac{c}{2}\|x - y\|$ für alle x, y in dieser Umgebung gilt, da wir aufgrund des Mittelwertsatzes die Abschätzung

$$\begin{aligned}\|o(\|x - y\|)\| &= \|f(x) - f(y) - Df(y)\langle x - y\rangle\|\\ &\le \|x - y\| \cdot \sup_{\xi \in [x,y]} \|Df(\xi) - Df(y)\|\end{aligned}$$

haben und Df stetig ist. Wir erhalten somit für $x, y \in B_r(x_0)$ für geeignete kleine $r > 0$

$$\|f(x) - f(y)\| \ge \frac{c}{2}\|x - y\|$$

und somit die Injektivität.

(v) $f|_U$ ist eine offene Abbildung: Sei $\tilde{U} \subset U$ offen und $\bar{x} \in \tilde{U}$ beliebig. Wir möchten nachweisen, dass es ein $\varepsilon > 0$ gibt, so dass $B_\varepsilon(f(\bar{x})) \subset f\big(\tilde{U}\big)$ gilt.

Ohne Einschränkung dürfen wir (nach Anwendung von Translationen) annehmen, dass $\bar{x} = 0$ und $f(\bar{x}) = 0$ gelten.

Sei κ mit $0 < \kappa < 1$ beliebig. Da Df stetig ist, gibt es ein $\rho > 0$, so dass $B_\rho(0) \subset \tilde{U}$ und

$$\left\|\mathrm{id} - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y)\, dt\right\| \le \kappa$$

für alle $x, y \in \bar{B}_\rho(0)$ gelten. (Erklärung dazu: Es gilt

$$\begin{aligned}&\left\|\mathrm{id} - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y)\, dt\right\|\\ =&\left\|(Df(0))^{-1} \circ (Df(0)) - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y)\, dt\right\|\end{aligned}$$

$$
\begin{aligned}
&= \left\| (Df(0))^{-1} \circ \left(Df(0) - \int_0^1 Df(tx + (1-t)y)\,dt \right) \right\| \\
&\le \left\| (Df(0))^{-1} \right\| \cdot \left\| \int_0^1 Df(0)\,dt - \int_0^1 Df(tx + (1-t)y)\,dt \right\|,
\end{aligned}
$$

da $Df(0)$ von t unabhängig ist,

$$
\begin{aligned}
&= \left\| (Df(0))^{-1} \right\| \cdot \left\| \int_0^1 Df(0) - Df(tx + (1-t)y)\,dt \right\| \\
&\le \left\| (Df(0))^{-1} \right\| \cdot \sup_{z \in [x,y]} \|Df(0) - Df(z)\| \cdot 1 \\
&\le \left\| (Df(0))^{-1} \right\| \cdot \sup_{z \in B_\rho(0)} \|Df(0) - Df(z)\|
\end{aligned}
$$

und dieses letzte Supremum können wir aufgrund der Stetigkeit von Df durch Verkleinern von $\rho > 0$ beliebig klein machen.)

Sei $u \in B_\varepsilon(0)$ beliebig, wobei wir $\varepsilon > 0$ gegebenenfalls noch verkleinern dürfen. Wir definieren die Abbildung $\varphi \colon \bar{B}_\rho(0) \to E$ für $\rho > 0$ wie oben durch

$$\varphi(x) := x + (Df(0))^{-1} \langle u - f(x) \rangle$$

und wollen zeigen, dass φ für kleine $\varepsilon > 0$ die Voraussetzungen von Theorem 7.1 erfüllt. Dann erhalten wir hieraus nämlich einen Fixpunkt, was gleichbedeutend mit $u - f(x)$ ist und die Offenheit liefert, da $u \in B_\varepsilon(0)$ beliebig war.

Seien also $x, y \in \bar{B}_\rho(0)$. Dann folgt unter Verwendung des Hauptsatzes der Differential- und Integralrechnung

$$
\begin{aligned}
\varphi(x) - \varphi(y) &= x - y - (Df(0))^{-1} \langle f(x) - f(y) \rangle \\
&= x - y - (Df(0))^{-1} \left\langle \int_0^1 \frac{d}{dt} f(tx + (1-t)y)\,dt \right\rangle \\
&= x - y - (Df(x))^{-1} \left\langle \int_0^1 Df(tx + (1-t)y) \langle x - y \rangle \,dt \right\rangle \\
&= \left(\mathrm{id} - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y)\,dt \right) \langle x - y \rangle.
\end{aligned}
$$

Somit erhalten wir aufgrund der Wahl von ρ

$$\|\varphi(x) - \varphi(y)\| \le \kappa \cdot \|x - y\|$$

für alle $x, y \in \bar{B}_\rho(0)$. Somit ist φ wie gewünscht eine Kontraktion.

Mit Theorem 7.1 folgt nun die Behauptung, falls wir noch $\|\varphi(0)\| \le \rho(1-\kappa)$ beweisen können. Aufgrund unserer Annahme $f(0) = 0$ erhalten wir aber

$$\|\varphi(0)\| = \left\| (Df(0))^{-1} \langle u \rangle \right\| \le \left\| (Df(0))^{-1} \right\| \cdot \|u\| \le \left\| (Df(0))^{-1} \right\| \cdot \varepsilon.$$

Somit erhalten wir die behauptete Abschätzung, indem wir $\varepsilon > 0$ falls nötig verkleinern, so dass

$$\varepsilon \le \left\| (Df(0))^{-1} \right\|^{-1} \rho(1 - \kappa)$$

gilt. $\square$

Mo 13.07.2020

Als weiteres Korollar zum Satz über die Umkehrfunktion erhalten wir die folgende globale Variante

Korollar 7.4. *Sei $\Omega \subset E$ offen und $f \in C^k(\Omega, F)$, $k \geq 1$, injektiv. Sei weiterhin $Df(x) \in L_{top}(E, F)$ für alle $x \in \Omega$. Dann ist $f(\Omega)$ offen und es gilt $f \in \mathrm{Diff}^k(\Omega, f(\Omega))$.*

Beweis. Übung. □

Sind E oder F endlichdimensional, so folgt aus der Bijektivität von $Df(x_0)$, dass E und F endlichdimensional sind. Wir betrachten also den Fall einer Abbildung $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$. Dann folgt aus der Bijektivität von $Df(x_0)$ bereits $n = m$ und wir erhalten

Korollar 7.5. *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^k(\Omega, \mathbb{R}^n)$, $k \geq 1$. Gelte für die Jacobimatrix $|J_f(x_0)| \neq 0$ in $x_0 \in \Omega$. Dann gibt es eine offene Umgebung $U = B_r(x_0)$, so dass auch $V := f(U)$ offen ist und es gilt $f \in \mathrm{Diff}^k(U, V)$.*

Also folgt aus $|J_f(x)| \neq 0$ für alle $x \in \Omega$, dass $f \in \mathrm{Diff}^k_{loc}(\Omega, f(\Omega))$ gilt.

Theorem 7.6. *Seien E, F Banachräume. Sei $\Omega \subset E$ offen und $f \in C^1(\Omega, F)$. Sei $x_0 \in \Omega$ ein Punkt, so dass $Df(x_0)\colon E \to F$ surjektiv ist und eine der folgenden Bedingungen erfüllt sind:*

(i) $E = \mathbb{R}^n$,

Dann gibt es ein $\varepsilon > 0$, so dass $B_\varepsilon(f(x_0)) \subset f(\Omega)$ gilt.

Beweis. Aufgrund einer der drei vorausgesetzten Bedingungen gibt es zu $E_1 := \ker(Df(x_0)) \subset E$ einen abgeschlossenen Unterraum $E_2 \subset E$, so dass $E = E_1 \times E_2$ mit einer zur Produktnorm äquivalenten Metrik ist. Entsprechend schreiben wir Vektoren in E als $x = \left(x^1, x^2\right)$.

Im Falle $E = \mathbb{R}^n$ erreichen wir dies beispielsweise, indem wir eine Orthonormalbasis von E_1 mit $(a_k, \ldots, a_n)$ zu einer Orthonormalbasis von E ergänzen und $E_2 := \langle a_k, \ldots, a_n\rangle$ definieren.

Nun gilt für $u \in E$

$$Df(x_0)\langle u\rangle = D_1 f(x_0)\left\langle u^1\right\rangle + D_2 f(x_0)\left\langle u^2\right\rangle.$$

Nach Voraussetzung gilt $D_1 f(x_0)\left\langle u^1\right\rangle = Df(x_0)\left\langle\left(u^1, 0\right)\right\rangle = 0$ und $D_2 f(x_0)\colon E_2 \to F$ ist surjektiv, da $Df(x_0)$ nach Voraussetzung surjektiv ist.

Dann ist die Abbildung $D_2 f(x_0)$ offen, da

(i) $D_2 f(x_0)$ als injektive und surjektive Abbildung bijektiv ist.

Somit folgt in allen Fällen $D_2 f(x_0) \in L_{\mathrm{top}}(E_2, F)$. Die Abbildung $g := f\left(x_0^1, \cdot\right)$ erfüllt daher die Voraussetzungen des Satzes von der inversen Abbildung. Somit gibt es eine Umgebung $U \in \mathcal{U}\left(x_0^2\right)$ in E_2, so dass $g\colon U \to g(U)$ ein C^1-Diffeomorphismus ist. Insbesondere ist $g(U) \subset f(\Omega)$ offen. □

7.3. Satz von der impliziten Funktion.

Bemerkung 7.7. Der Satz von der impliziten Funktion verallgemeinert das folgende Resultat aus der linearen Algebra: Sei $A = (A_1\, A_2)$ eine Blockmatrix, so dass A_2 bijektiv ist. Dann gibt es für jedes x^1 ein $x^2 = \varphi\left(x^1\right)$, so dass $A\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = 0$ gilt.

Bemerkung 7.8.

(i) Den folgenden Satz über implizite Funktionen betrachten wir aufgrund der einfacheren Notierbarkeit in einer Banachraumversion. Die Variante im Euklidischen erhält man mit offenen $\Omega_1 \subset \mathbb{R}^n$, $\Omega_2 \subset \mathbb{R}^m$, für eine Funktion

$f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}^m$ und

$$D_1 f(x,y) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \dots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f^k}{\partial x^1} & \dots & \frac{\partial f^k}{\partial x^n} \end{pmatrix} (x,y)$$

sowie

$$D_2 f(x,y) = \begin{pmatrix} \frac{\partial f^1}{\partial y^1} & \dots & \frac{\partial f^1}{\partial y^m} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial y^1} & \dots & \frac{\partial f^m}{\partial y^m} \end{pmatrix} (x,y).$$

Die Bedingung $D_2 f(x_0,y_0) \in L_{\text{top}}(\mathbb{R}^m,\mathbb{R}^m)$ wird dabei zu

$$\det D_2 f(x_0,y_0) \neq 0.$$

Theorem 7.9 (Satz von der impliziten Funktion, Satz über implizite Funktionen, “Implicit function theorem”)**.**
Seien E_1, E_2, F Banachräume, $\Omega_i \subset E_i$ offen, $i = 1,2$ und $f \in C^k(\Omega_1 \times \Omega_2, F)$, $k \geq 1$. Sei $(x_0,y_0) \in \Omega_1 \times \Omega_2$ und gelte

$$f(x_0,y_0) = 0 \quad \textit{sowie} \quad D_2 f(x_0,y_0) \in L_{top}(E_2,F).$$

Dann gibt es eine offene Umgebung U von x_0 und $\varphi \in C^k(U,E_2)$ mit $\varphi(x_0) = y_0$ und

$$f(x,\varphi(x)) = 0$$

für alle $x \in U$.

Die Abbildung φ ist eindeutig bestimmt und es gilt

$$D\varphi(x) = -(D_2 f(x,\varphi(x)))^{-1} \circ D_1 f(x,\varphi(x)).$$

Gilt $f \in C^{k,\alpha}$, $0 < \alpha \leq 1$, so folgt auch $\varphi \in C^{k,\alpha}$.

Beweis. Wir möchten den Satz von der Umkehrabbildung benutzen.

(i) Um die Existenz von φ zu zeigen, definieren wir

$$g\colon \Omega_1 \times \Omega_2 \to E_1 \times F \quad \text{durch} \quad (x,y) \mapsto (x, f(x,y)).$$

Es gilt $g \in C^k(\Omega_1 \times \Omega_2, E_1 \times F)$, weil die Komponenten selber in C^k sind und für die Ableitung erhalten wir

$$Dg = (D\,\mathrm{id}_{E_1}, Df).$$

Wir schreiben dies komponentenweise

$$\begin{aligned} Dg(x,y)\langle u,v\rangle &= D_1 g(x,y)\langle u\rangle + D_2 g(x,y)\langle v\rangle \\ &= (u, D_1 f(x,y)\langle u\rangle) + (0, D_2 f(x,y)\langle v\rangle) \\ &= (u, D_1 f(x,y)\langle u\rangle + D_2 f(x,y)\langle v\rangle). \end{aligned}$$

In Matrixnotation können wir dies als

$$Dg(x,y) = \begin{pmatrix} \mathrm{id}_{E_1} & 0 \\ D_1 f(x,y) & D_2 f(x,y) \end{pmatrix}$$

schreiben.

(ii) Wir möchten nun nachweisen, dass $Dg(x_0,y_0) \in L_{\text{top}}(E_1 \times E_2, E_1 \times F)$ gilt. Das Argument (x_0,y_0) ist hier nicht so relevant, da stets gleich.

(a) Die Injektivität von $Dg(x_0,y_0)$ folgt aus der Injektivität von $D_2 f(x_0,y_0)$.

(b) Surjektivität von $Dg(x_0, y_0)$: Sei $(\bar{u}, \bar{v}) \in E_1 \times F$ beliebig und definiere

$$(u, v) := \left(\bar{u}, (D_2 f(x_0, y_0))^{-1} \langle \bar{v} - D_1 f(x_0, y_0) \langle \bar{u} \rangle \rangle \right).$$

Durch Einsetzen in die obige Formel für g sieht man direkt, dass sich der Term $D_1 f(x_0, y_0)\langle \bar{u} \rangle$ weghebt und $Dg(x_0, y_0)\langle u, v \rangle = (\bar{u}, \bar{v})$ gilt. Wir haben damit als Formel für die Inverse

$$(Dg(x_0, y_0))^{-1} \langle \bar{u}, \bar{v} \rangle = \left(\bar{u}, (D_2 f(x_0, y_0))^{-1} \langle \bar{v} - D_1 f(x_0, y_0) \langle \bar{u} \rangle \rangle \right).$$

(c) Stetigkeit von $(Dg(x_0, y_0))^{-1}$: Aus der obigen Formel erhalten wir direkt mit von $\|(D_2 f(x_0, y_0))^{-1}\|$ und $\|D_1 f(x_0, y_0)\|$ abhängigen Konstanten

$$\begin{aligned}
\left\|(Dg(x_0, y_0))^{-1} \langle \bar{u}, \bar{v} \rangle\right\| &= \max\left(\|\bar{u}\|, \left\|(D_2 f(x_0, y_0))^{-1} \langle \bar{v} - D_1 f(x_0, y_0) \langle \bar{u} \rangle \rangle\right\|\right) \\
&\leq c \cdot \max\left(\|\bar{u}\|, \|\bar{v} - D_1 f(x_0, y_0)\langle \bar{u} \rangle\|\right) \\
&\leq c \cdot \max\left(\|\bar{u}\|, \|\bar{v}\| + \|D_1 f(x_0, y_0)\langle \bar{u} \rangle\|\right) \\
&\leq c \cdot \max\left(\|\bar{u}\|, \|\bar{v}\| + \|\bar{u}\|\right) \leq c \cdot \max\left(\|\bar{u}\|, \|\bar{v}\|\right).
\end{aligned}$$

Somit ist $Dg(x_0, y_0) \in L_{\text{top}}(E_1 \times E_2, E_1 \times F)$.

(iii) Nach dem Satz über die Umkehrfunktion gibt es daher ein $r > 0$, so dass mit $U_0 := B_r(x_0)$ und $V_0 := B_r(y_0)$ sowie der offenen Menge $W_0 := g(U_0 \times V_0) \subset E_1 \times F$ je nach Differenzierbarkeit von f

$$g \in \text{Diff}^k(U_0 \times V_0, W_0) \quad \text{oder} \quad g \in \text{Diff}^{k,\alpha}(U_0 \times V_0, W_0)$$

gilt. Wir bezeichnen die lokale Inverse von g mit h

$$h \colon W_0 \to U_0 \times V_0.$$

Wir schreiben $h = \left(h^1, h^2\right)$ und erhalten für $(x, z) \in W_0 \subset E_1 \times F$

$$(x, z) = g \circ h(x, z) = g\left(h^1(x, z), h^2(x, z)\right) = \left(h^1(x, z), f\left(h^1(x, z), h^2(x, z)\right)\right).$$

Somit gelten für alle $(x, z) \in W_0$ die Identitäten $h^1 = \pi_1$ (Projektion auf die erste Komponente) und $f\left(x, h^2(x, z)\right) = z$.

Wir definieren nun $U := \{x \in U_0 \colon (x, 0) \in W_0\}$. Wegen $g(x_0, y_0) = (x_0, f(x_0, y_0)) = (x_0, 0)$ folgt $x_0 \in U$ und U ist offen, da U_0 und W_0 offene Mengen sind. Weiterhin definieren wir $\varphi \colon U \to V_0$ durch $\varphi := h^2(\cdot, 0)$. Dann gilt nach Definition $f(x, \varphi(x)) = 0$ für alle $x \in U$ und es ist $\varphi \in C^k(U, E_2)$ bzw. $\varphi \in C^{k,\alpha}(U, E_2)$.

(iv) Eindeutigkeit: Wir zeigen, dass für die gewählte Menge U die Abbildung $\varphi \colon U \to E_2$ die einzige Abbildung ist, so dass $f(x, \varphi(x)) = 0$ gilt. Sei dazu $(x, y) \in U_0 \times V_0$ beliebig mit $f(x, y) = 0$. Dann folgt $g(x, y) = (x, f(x, y)) = (x, 0)$. Da $g \colon U_0 \times V_0 \to W_0$ injektiv ist, kann es zu festem x maximal ein solches y geben. Daher ist φ eindeutig bestimmt.

(v) Formel für die Ableitung: Da $Dg(x, y) \in L_{\text{top}}(E_1 \times E_2, E_1 \times F)$ für $(x, y) \in U_0 \times V_0$ gilt, lesen wir aus der Matrixdarstellung von Dg ab, dass $D_2 f(x, y) \in L_{\text{top}}(E_2, F)$ ebenfalls für $(x, y) \in U_0 \times V_0$ gilt.

Mit der Kettenregel erhalten wir aus $f(x, \varphi(x)) = 0$ die Identität

$$D_1 f(x, \varphi(x))\langle \cdot \rangle + D_2 f(x, \varphi(x))\langle D\varphi(x)\langle \cdot \rangle \rangle = 0.$$

Für $x \in U$ gilt $(x, \varphi(x)) \in U_0 \times V_0$. Somit ist $D_2 f(x, \varphi(x))$ invertierbar und durch direktes Auflösen der obigen Gleichheit nach $D\varphi$ erhalten wir die Behauptung.

(vi) Ist $f \in C^{k,\alpha}$, so liest man an der Formel für $D\varphi$ ab, dass dann auch $\varphi \in C^{k,\alpha}$ gilt. $\square$

Bemerkung 7.10.

(i) Im Satz von der impliziten Funktion kann man durch Betrachten von $f - z_0$ statt f die Bedingung $f(x_0, y_0) = 0$ auch durch $f(x_0, y_0) = z_0 \in F$ ersetzen.
(ii) Sei $A \in L(\mathbb{R}^n, \mathbb{R}^m)$. Dann gilt stets für den Rang von A die Abschätzung $\operatorname{rk} A \le \min(m, n)$. Gilt $\operatorname{rk} A = \min(m, n)$, so sagen wir, dass A *maximalen* oder *vollen Rang* besitzt.
(iii) Wir bemerken, dass aus der Unterhalbstetigkeit des Ranges oder der Stetigkeit jeder Unterdeterminante folgt, dass die Menge, in der eine Abbildung Df für ein $f \in C^1(\Omega, \mathbb{R}^m)$, $\Omega \subset \mathbb{R}^n$ offen, maximalen Rang besitzt, offen ist.
(iv) Ist $f : \mathbb{R}^n \times \mathbb{R}^m \equiv \mathbb{R}^{n+m} \to \mathbb{R}^m$ gegeben, so gilt die Bedingung $D_2 f(x_0, y_0) \in L_{\text{top}}(\mathbb{R}^m, \mathbb{R}^m)$ aus dem Satz von der impliziten Funktion genau dann, wenn $D_2 f(x_0, y_0)$ surjektiv (oder injektiv) ist.
(v) Ist $f : \mathbb{R}^{n+m} \to \mathbb{R}^m$ gegeben und $Df(x_0, y_0)$ surjektiv bzw. $J_f(x_0, y_0)$ hat vollen Rang, so können wir durch Umnummerierung der Koordinaten im $\mathbb{R}^{n+m}$ erreichen, dass wir eine Abbildung $\tilde{f} : \mathbb{R}^{n+m} \to \mathbb{R}^m$ mit bijektivem $D_2 \tilde{f}(x_0, y_0)$ erhalten.

Do 16.07.2020

Aufgabe 7.11. Zeige, dass es unter den Voraussetzungen des Satzes über implizite Funktionen eine Umgebung A von (x_0, y_0) mit

$$\{(x, y) \in A : f(x, y) = 0\} = \{(x, \varphi(x)) : x \in U\} \cap A \equiv \operatorname{graph} \varphi \cap A$$

gibt, d. h. dass sich lokal die gesamte Menge $\{f = 0\}$ als $\operatorname{graph} \varphi$ beschreiben läßt.

Hinweis. Benutze den Teil des Beweises des Satzes über implizite Funktionen, in dem wir die Eindeutigkeit von φ gezeigt haben. □

In diesem Zusammenhang verwendet man häufiger die folgende Definition. Sie verallgemeinert auch die bisherige Definition eines kritischen Punktes.

Definition 7.12. Sei E ein Banachraum und $\Omega \subset E$ offen. Sei $f \in C^1(\Omega, \mathbb{R}^m)$.
(i) Dann heißt $x \in \Omega$ ein *regulärer Punkt* von f, falls $Df(x)$ surjektiv ist. Sonst heißt x ein *kritischer Punkt.*
(ii) Ein Punkt $y \in \mathbb{R}^m$ heißt *regulärer Wert* von f, falls $f^{-1}(\{y\})$ nur aus regulären Punkten besteht. Sonst heißt y ein *singulärer Wert.*

In der folgenden Bemerkung führen wir die später in der Differentialgeometrie noch wichtigen Hyperflächen ein.

Bemerkung 7.13.
(i) Sei $f \in C^k(\Omega)$, $\Omega \subset \mathbb{R}^n$ offen und $k \ge 1$. Die Bedingung, dass ∇f bzw. J_f in einem Punkt $x_0 \in \Omega$ vollen Rang besitzt ist dann äquivalent zu $\nabla f(x_0) \neq 0$.
(ii) Sei $c \in \mathbb{R}$. Gilt $\nabla f \neq 0$ überall in der Menge Ω, so gibt es für jeden Punkt $x_0 \in M_c := f^{-1}(\{c\})$ eine Umgebung, so dass in dieser Umgebung von x_0 die Menge M_c mit dem Graphen einer auf einer offenen Teilmenge von $\mathbb{R}^{n-1}$ definierten C^k-Funktion übereinstimmt, wobei $\mathbb{R}^{n-1}$ senkrecht zu einer Richtung e_i steht, so dass $Df\langle e_i \rangle \neq 0$ gilt. Es genügt sogar, die Bedingung $\nabla f \neq 0$ nur auf M_c zu fordern.
(iii) Ist $f(x_0) = c$ und $Df(x_0)\langle e_n \rangle \neq 0$, so gibt es aufgrund des Satzes über implizite Funktionen eine Umgebung U von $\hat{x}_0 \equiv \left(x_0^1, \ldots, x_0^{n-1}\right)$ und ein $\varepsilon > 0$ sowie eine Funktion $\varphi \in C^k(U)$ mit $\varphi(\hat{x}_0) = x_0^n$ und

$$M_c \cap (U \times (x_0^n - \varepsilon, x_0^n + \varepsilon)) = \operatorname{graph} \varphi$$

bzw. $f(\hat{x}, \varphi(\hat{x})) = c$ für alle $\hat{x} \in U$.
(iv) Mengen M_c mit dieser Eigenschaft heißen *Hyperflächen* der Klasse C^k in $\Omega \subset \mathbb{R}^n$ oder $n-1$-dimensionale *Untermannigfaltigkeiten* von $\Omega \subset \mathbb{R}^n$. Im Falle $n = 2$ heißen sie auch *Niveaulinien* und im Falle $n = 3$ *Niveauflächen.*

Wir halten diese wichtige Definition noch einmal gesondert fest.

Definition 7.14. Seien $\Omega \subset \mathbb{R}^n$ offen, $f \in C^k(\Omega)$, $k \geq 1$, und c ein regulärer Wert von f. Dann heißt $M_c := f^{-1}(\{c\})$ eine C^k-Hyperfläche in $\Omega \subset \mathbb{R}^n$.

Beispiel 7.15. Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ durch $x \mapsto |x|^2$ gegeben. Dann sind die Mengen $M_c := f^{-1}(\{c\})$ für alle $c \neq 0$ glatte Hyperflächen in $\mathbb{R}^n$. Es gilt $M_c = \partial B_{\sqrt{c}}(0)$ für alle $c > 0$. Es gilt $M_c \cap \{x^n > 0\} = \operatorname{graph}\left(\hat{x} \mapsto \sqrt{c^2 - |\hat{x}|^2}\right)$.

Aufgabe 7.16. Sei $\Omega \subset \mathbb{R}^n$ offen. Sei $f \in C^2(\Omega, \mathbb{R})$. Sei c ein regulärer Wert von f. Dann gibt es für jedes $x_0 \in M_c = f^{-1}(\{c\})$ eine offene Umgebung U von x_0, eine Translation T, eine orthogonale Abbildung O und eine Funktion $\varphi \in C^2(\mathbb{R}^{n-1})$, so dass

(i) $M_c \cap U = (TO \operatorname{graph} \varphi) \cap U$,
(ii) $TO(0) = x_0$,
(iii) $\nabla\varphi(0) = 0$ gelten und die Hessematrix $(\varphi_{ij}(0))_{1 \leq i,j \leq n-1}$ diagonal ist.

7.4. **Extrema mit Nebenbedingungen.**

Theorem 7.17 (Lagrangesche Multiplikatorregel)**.** *Sei $\Omega \subset \mathbb{R}^n$ offen. Sei $f \in C^1(\Omega)$ und sei $\Phi = \left(\Phi^i\right) \in C^1(\Omega, \mathbb{R}^m)$ mit $m < n$. Erfülle $z_0 \in \Omega$ die Bedingungen*

(i) $\Phi(z_0) = 0$,
(ii) $f|_{\{\Phi=0\}}$ *nimmt in z_0 ein lokales Extremum an, d. h. wenn wir f nur auf der Menge $\bigcap\limits_i \{\Phi^i = 0\}$ betrachten, so nimmt f in z_0 ein lokales Extremum an, und*
(iii) $\operatorname{rk} D\Phi(z_0) = m$, *d. h. es gilt* $\dim \operatorname{im} D\Phi(z_0) = m$.

Dann gibt es $\lambda = (\lambda_i)_i \in (\mathbb{R}^m)^*$, so dass z_0 ein kritischer Punkt der (nun auf ganz Ω definierten) Funktion $g\colon \Omega \to \mathbb{R}$

$$g := f + \sum_{i=1}^m \lambda_i \Phi^i$$

ist. Insbesondere gilt also

$$0 = \nabla g(z_0) = \nabla f(z_0) + \sum_{i=1}^m \lambda_i \nabla\Phi^i(z_0).$$

Beweis. Falls nötig nummerieren wir die Koordinaten um und schreiben $\mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m$ mit $z = (x, y)$ sowie $z_0 = (x_0, y_0)$. (Insbesondere für diese Aufsplittung des Definitionsbereiches benötigt man in Banachräumen die Zusatzvoraussetzungen.) Damit können wir es erreichen, dass $D_2\Phi(z_0) \in GL(m)$ gilt. (Da $GL(m)$ offen ist, gilt $D_2\Phi(z) \in GL(m)$ auch noch für z in einer kleinen Umgebung von z_0.) Nach dem Satz über die inverse Funktion finden wir eine Umgebung U von x_0 in $\mathbb{R}^{n-m}$ und eine Funktion $\varphi \in C^1(U, \mathbb{R}^m)$ mit $\varphi(x_0) = y_0$, so dass $\Phi(x, \varphi(x)) = 0$ für alle $x \in U$ gilt und $(x, \varphi(x))$ lokal alle Punkte mit $\Phi = 0$ beschreibt, d. h. es gibt ein $r > 0$, so dass $\{z \in B_r(z_0) \colon \Phi(z) = 0\} = \{(x, \varphi(x)) \colon x \in U\} \cap B_r(z_0)$ gilt. Wie im Satz von der inversen Funktion gilt $D\varphi(x) = -(D_2\Phi(x, \varphi(x)))^{-1} \circ D_1\Phi(x, \varphi(x))$ für alle $x \in U$.

Nun haben wir $\{\Phi = 0\}$ lokal als $\operatorname{graph} \varphi$ geschrieben und erhalten daher, dass die Funktion $h\colon U \to \mathbb{R}$ mit $h(x) := f(x, \varphi(x))$ in $x_0 \in U$ ein lokales Extremum annimmt. Somit folgt nach Kettenregel

$$0 = Dh(x_0) = D_1 f(x_0, \varphi(x_0)) + D_2 f(x_0, \varphi(x_0)) \circ D\varphi(x_0).$$

Nun benutzen wir die Darstellung für $D\varphi$ und erhalten mit $\varphi(x_0) = y_0$

$$0 = D_1 f(x_0, y_0) - D_2 f(x_0, y_0) \circ (D_2\Phi(x_0, y_0))^{-1} \circ D_1\Phi(x_0, y_0).$$

Wir definieren nun $\lambda := -D_2 f(z_0) \circ (D_2\Phi(z_0))^{-1} \in (\mathbb{R}^m)^*$. Somit können wir aus $Dh = 0$ letztlich folgern, dass in z_0

$$0 = D_1 f + \sum_{j=1}^{m} \lambda_j D_1 \Phi^j$$

gilt. Weiterhin folgt durch Umsortieren der Definition von λ in z_0

$$0 = D_2 f + \sum_{j=1}^{m} \lambda_j D_2 \Phi^j.$$

Also ist z_0 ein kritischer Punkt der Funktion

$$g = f + \sum_{j=1}^{m} \lambda_j \Phi^j.$$

Dies war gerade die Behauptung. □

Als Beispiel erhalten wir eine Möglichkeit, Eigenwerte und Eigenvektoren symmetrischer Matrizen zu bestimmen.

Theorem 7.18 (Eigenwerte symmetrischer Matrizen)**.**
Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Nehme die Funktion $f\colon \mathbb{S}^{n-1} \to \mathbb{R}$ mit

$$f(x) := \langle Ax, x\rangle$$

ihr Minimum in $x_0 \in \mathbb{S}^{n-1}$ an. Dann ist x_0 ein Eigenvektor von A zum Eigenwert $f(x_0)$.

Beweis. Ohne Unterscheidung in der Bezeichnung können wir auch annehmen, dass $f\colon \mathbb{R}^n \to \mathbb{R}$ durch $f(x) := \langle Ax, x\rangle$ definiert ist. Dann ist x_0 ein Minimum von f unter der Nebenbedingung $\Phi = 0$ mit $\Phi\colon \mathbb{R}^n \to \mathbb{R}$ und $\Phi(x) = |x|^2 - 1$. Es gilt $D\Phi(x_0) \neq 0$. Nach der Lagrangeschen Multiplikatorregel folgt daher, dass es ein $\lambda \in \mathbb{R}^* \cong \mathbb{R}$ gibt, so dass x_0 ein kritischer Punkt von $g := f + \lambda\Phi$ ist. Wir erhalten somit für alle $y \in \mathbb{R}^n$ aufgrund der Symmetrie von A

$$\begin{aligned} 0 &= Dg(x_0)\langle y\rangle = Df(x_0)\langle y\rangle + \lambda D\Phi(x_0)\langle y\rangle \\ &= \langle Ay, x_0\rangle + \langle Ax_0, y\rangle + 2\lambda\langle x_0, y\rangle \\ &= 2\langle Ax_0 + \lambda x_0, y\rangle. \end{aligned}$$

Da $y \in \mathbb{R}^n$ beliebig ist, folgt $Ax_0 = -\lambda x_0$.

Wir müssen noch nachweisen, dass $-\lambda = f(x_0)$ gilt. Dazu setzen wir die Eigenwertgleichung $Ax_0 = -\lambda x_0$ in die Definition von f ein und erhalten $f(x_0) = \langle Ax_0, x_0\rangle = \langle -\lambda x_0, x_0\rangle = -\lambda$, da $|x_0|^2 = 1$ gilt. □

Korollar 7.19. *Sei X ein metrischer Raum und $A\colon X \to \mathbb{R}^{n\times n}$ eine stetige Funktion, die für jedes $x \in X$ eine symmetrische $n \times n$-Matrix $A(x)$ liefert. Dann hängt der kleinste Eigenwert von $A(x)$ stetig von x ab.*

Beweis. Nach Theorem 7.18 genügt der Nachweis, dass $x \mapsto \min\limits_{\xi\in\mathbb{S}^{n-1}} \langle A(x)\xi, \xi\rangle$ stetig ist.

Auf $\mathbb{R}^{n\times n}$ sind alle Normen äquivalent. Daher können wir die Operatornorm verwenden und erhalten für symmetrische $n \times n$-Matrizen mit $\|A - B\| \le \varepsilon$ für $|\xi| = 1$

$$\min_{\zeta\in\mathbb{S}^{n-1}} \langle A\zeta, \zeta\rangle \le \langle A\xi, \xi\rangle = \langle B\xi, \xi\rangle + \langle (A - B)\xi, \xi\rangle \le \langle B\xi, \xi\rangle + \underbrace{\|A - B\|}_{\le\varepsilon} \cdot \underbrace{|\xi|^2}_{=1}.$$

Nun folgt zunächst $\min\limits_{\zeta\in\mathbb{S}^{n-1}}\langle A\zeta,\zeta\rangle \le \langle B\xi,\xi\rangle+\varepsilon$ für alle $\xi\in\mathbb{S}^{n-1}$. Da $\xi\in\mathbb{S}^{n-1}$ beliebig war, erhalten wir auch $\min\limits_{\zeta\in\mathbb{S}^{n-1}}\langle A\zeta,\zeta\rangle \le \min\limits_{\zeta\in\mathbb{S}^{n-1}}\langle B\zeta,\zeta\rangle+\varepsilon$. Im obigen Beweis hätten wir die Rollen von A und B auch austauschen können. Somit folgt die behauptete Stetigkeit des Minimums mit $\delta=\varepsilon$. $\square$

Literatur

1. Jean Dieudonné, *Foundations of modern analysis*, Academic Press, New York, 1969, Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.
2. Claus Gerhardt, *Analysis. I*, German ed., International Series in Analysis, International Press, Somerville, MA, 2004.
3. Oliver C. Schnürer, *Analysis I*, 2019, Skript zur Vorlesung.
4. Oliver C. Schnürer, *Lineare Algebra I*, 2010, Skript zur Vorlesung.
5. Wikipedia, `http://www.wikipedia.org`.
6. Michel Willem, *Minimax theorems*, Progress in Nonlinear Differential Equations and their Applications, 24, Birkhäuser Boston Inc., Boston, MA, 1996.

Oliver C. Schnürer, Universität Konstanz, Germany
Email address: `Oliver.Schnuerer@uni-konstanz.de`