

# ANALYSIS II

OLIVER C. SCHNÜRER

ZUSAMMENFASSUNG. Skript zu einer Vorlesung Analysis II.

## INHALTSVERZEICHNIS

|                                                       |     |
|-------------------------------------------------------|-----|
| 4. Differentiation in einer Variablen                 | 2   |
| 4.1. Differenzierbare Funktionen                      | 2   |
| 4.2. Die Regeln von de l'Hospital                     | 11  |
| 4.3. Differentiation von Funktionenfolgen             | 14  |
| 4.4. Elementare Funktionen                            | 16  |
| 4.5. Polynome ★                                       | 25  |
| 4.6. Taylorsche Formeln                               | 26  |
| 5. Integration in einer Variablen                     | 32  |
| 5.1. Das Riemannsche Integral                         | 32  |
| 5.2. Integrationsregeln                               | 36  |
| 5.3. Monotone und stetige Funktionen                  | 41  |
| 5.4. Hauptsatz der Differential- und Integralrechnung | 42  |
| 5.5. Integralsätze und Transformationsregeln          | 44  |
| 5.6. Integration rationaler Funktionen ★              | 46  |
| 5.7. Lebesguesches Integrabilitätskriterium ★         | 47  |
| 5.8. Uneigentliche Integrale                          | 50  |
| 5.9. Parameterabhängige Integrale                     | 55  |
| 6. Differentiation in Banachräumen                    | 63  |
| 6.1. Differenzierbarkeit                              | 63  |
| 6.2. Der Mittelwertsatz und Anwendungen               | 71  |
| 6.3. Differentiation von Funktionenfolgen             | 72  |
| 6.4. Partielle Ableitungen                            | 74  |
| 6.5. Ableitungen höherer Ordnung                      | 80  |
| 6.6. Taylorsche Formel                                | 87  |
| 6.7. Lokale Extrema                                   | 88  |
| 7. Existenzsätze                                      | 90  |
| 7.1. Banachscher Fixpunktsatz                         | 90  |
| 7.2. Satz über die Umkehrabbildung                    | 93  |
| 7.3. Satz von der impliziten Funktion                 | 99  |
| 7.4. Das Rangtheorem ★                                | 103 |
| 7.5. Extrema mit Nebenbedingungen                     | 105 |
| 8. Numerische Suche nach einem Minimum ★              | 108 |
| 9. Zusatzmaterial zur Differenzierbarkeit ★           | 110 |
| Literatur                                             | 110 |

---

*Date:* 28. März 2025.

*2000 Mathematics Subject Classification.* 26-01.

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

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 [6].

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

#### 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 4.1.1.** Sei  $\Omega \subset \mathbb{R}$  offen,  $F$  eine Banachraum und  $f: \Omega \rightarrow F$  eine Abbildung.

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

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

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

##### Bemerkung 4.1.2.

- (i) ★ Die Aussagen

$$f'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

und

$$\left\| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right\| \rightarrow 0$$

für  $x \rightarrow x_0$  mit  $x \neq x_0$  sind äquivalent.

- (ii) Direkt aus der Definition folgt: Ist  $f$  in  $x_0$  differenzierbar, so ist  $f$  in  $x_0$  auch stetig.  
 (iii) ★ Die in  $x_0 \in \Omega$ , die in  $\Omega$  und die in  $\Omega$  stetig differenzierbaren Funktionen bilden jeweils einen Unterraum des Vektorraumes aller Abbildungen  $\Omega \rightarrow F$ .

**Definition 4.1.3** (Landausymbole). Seien  $E, F$  Banachräume, sei  $\Omega \subset E$  offen und seien  $f: \Omega \rightarrow F$  und  $g: \Omega \rightarrow \mathbb{R}_{>0}$  Abbildungen. Sei  $x_0 \in \Omega$ . Bei den nachfolgenden Definitionen ist stets kenntlich zu machen, dass  $x_0$  der Bezugspunkt ist.

- (i) Gilt

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

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

- (ii) Gilt

$$\overline{\lim}_{\substack{x \rightarrow x_0 \\ \Omega \ni x \neq x_0}} \frac{\|f(x)\|}{g(x)} < \infty,$$

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

##### Bemerkung 4.1.4. ★

- (i) Sei  $\alpha > 0$ . Wir benutzen häufig  $f(x) \in o(\|x - x_0\|^\alpha)$  und sagen, dass  $f$  **in**  $x_0$  **von höherer Ordnung als  $\alpha$  verschwindet**.
- (ii) Sei  $\alpha > 0$ . Wir benutzen häufig  $f(x) \in O(\|x - x_0\|^\alpha)$ .
- (iii) Wir verwenden diese Definitionen auch, wenn  $f$  nur in  $\Omega \setminus \{x_0\}$  definiert ist.
- (iv) Mit  $o(\|x - x_0\|^\alpha), \dots$  sind impliziert Mengen von Funktionen definiert. Statt  $f(x) \in o(\|x\|^\alpha)$  oder  $f(x) \in O(\|x\|^\alpha)$  schreibt man auch  $f(x) = o(\|x\|^\alpha)$  bzw.  $f(x) = O(\|x\|^\alpha)$ .
- (v) Im Falle  $\Omega \subset \mathbb{R}$  benutzen wir diese Notation auch für  $x \rightarrow \infty$  oder  $x \rightarrow -\infty$ .

**Beispiele 4.1.5.**

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

$$f(x) = o(|x|^\alpha) \quad \text{für alle } 0 < \alpha < n \quad \text{und} \quad f(x) = O(|x|^n).$$
- (ii) ★ Für  $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  mit  $x \mapsto \sqrt{x}$  gelten für  $x_0 = 0$  und alle  $0 < \alpha < \frac{1}{2}$ 

$$f(x) = o(x^\alpha) \quad \text{und} \quad f(x) = O(\sqrt{x}).$$
- (iii) Für  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x) = x^1 \cdot x^2$  gilt  $f(x) = O(|x|^2)$  und daher folgt für  $x_0 = 0$  und alle  $0 < \alpha < 2$ 

$$f(x) = o(|x|^\alpha).$$
- (iv) Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  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 4.1.6.** Sei  $F$  ein Banachraum und sei  $\Omega \subset \mathbb{R}$  offen. Dann ist  $f: \Omega \rightarrow F$  genau dann in  $x_0 \in \Omega$  differenzierbar, wenn es ein  $a \in F$  mit

$$(4.1) \quad f(x) = f(x_0) + a \cdot (x - x_0) + o(|x - x_0|)$$

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  $x \in \Omega$

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

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

★ Da die linearen stetigen Abbildungen  $A: \mathbb{R} \rightarrow F$  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 4.1.7.** Sei  $F$  ein Banachraum und  $\Omega \subset \mathbb{R}$  offen. Dann ist  $f: \Omega \rightarrow F$  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  $A = Df(x_0)$ .

**Bemerkung 4.1.8.** ★

- (i) Proposition 4.1.7 lässt sich später gut auf Funktionen verallgemeinern, die auf offenen Teilmengen eines Banachraumes definiert sind.
- (ii) Ist  $f$  in  $\Omega$  differenzierbar, so bezeichnet  $Df$  die Ableitung von  $f$ .  $Df$  ist durch

$$Df: \Omega \rightarrow L(\mathbb{R}, F),$$

$$x \mapsto Df(x)$$

gegeben.

- (iii) Im reellwertigen Fall ist  $A\langle x - x_0 \rangle$  durch die Multiplikation der beiden reellen Zahlen  $f'(x_0)$  und  $x - x_0$  gegeben. Identifizieren wir  $f'(x_0) \in \mathbb{R}$  mit der linearen stetigen Abbildung  $\mathbb{R} \ni y \mapsto f'(x_0) \cdot y \in \mathbb{R}$ , so schreiben wir auch  $f'(x_0)\langle y \rangle$ . Entsprechendes gilt auch für vektorwertige Funktionen.

**Proposition 4.1.9.** Sei  $\Omega \subset \mathbb{R}$  offen. Seien  $f, g: \Omega \rightarrow \mathbb{R}$  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

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

**Bemerkung 4.1.10.** ★ Proposition 4.1.9 (i) besagt, dass die Ableitung eine lineare Abbildung auf den in  $x_0$  differenzierbaren Funktionen nach  $L(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}$  ist.

*Beweis von Proposition 4.1.9.*

- (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  $x \rightarrow x_0$  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  $x \rightarrow x_0$  über. Die Behauptung folgt.  $\square$

**Korollar 4.1.11.** Sei  $n \in \mathbb{N}_{>0}$ . Dann ist  $f_n: \mathbb{R} \rightarrow \mathbb{R}$  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

$$f'_n(x) = (f_{n-1}f_1)'(x) = f'_{n-1}(x)f_1(x) + f_{n-1}(x)f'_1(x)$$

$$= (n-1)x^{n-2} \cdot x + x^{n-1} \cdot 1 = nx^{n-1}. \quad \square$$

**Bemerkung 4.1.12.** ★ Einen Beweis, in dem  $(x+y)^n - x^n$  mit Hilfe von Binomialkoeffizienten geschrieben wird, lassen wir als Übung.

**Theorem 4.1.13** (Kettenregel). *Sei  $\Omega, \tilde{\Omega} \subset \mathbb{R}$  offen. Seien  $g: \Omega \rightarrow \mathbb{R}$  in  $x_0$  und  $f: \tilde{\Omega} \rightarrow \mathbb{R}$  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 4.1.7 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

$$\begin{aligned} (f \circ g)(x) - (f \circ g)(x_0) &= f'(g(x_0)) \langle g'(x_0) \langle x - x_0 \rangle \rangle \\ &\quad + f'(g(x_0)) \langle o(|x - x_0|) \rangle + o(|g(x) - g(x_0)|). \end{aligned}$$

- (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  $\varepsilon(|x - x_0|) \rightarrow 0$  für  $x \rightarrow x_0$  schreiben können, genügt der Nachweis, dass

$$\lim_{x \rightarrow x_0} \varepsilon(|g(x) - g(x_0)|) \cdot \frac{|g(x) - g(x_0)|}{|x - x_0|} = 0$$

gilt. Dies folgt nun aus der Stetigkeit und Differenzierbarkeit von  $g$  in  $x_0$  und Grenzwertsätzen.  $\square$

**Bemerkung 4.1.14.**  $\star$  Die Ableitung bezeichnet man nach Gottfried Wilhelm Leibniz auch mit  $\frac{df}{dx}$ . Schreibt man  $f(g(x))$  und bezeichnet die Variable von  $f$  mit  $y$  für  $g(x)$ , so ergibt sich ohne Argumente die an das Kürzen von Brüchen erinnernde Merkregel

$$\frac{df}{dx} = \frac{df}{dy} \frac{dy}{dx}.$$

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

**Korollar 4.1.15.** *Seien  $\Omega, \tilde{\Omega} \subset \mathbb{R}$  offen. Sei  $f: \Omega \rightarrow \tilde{\Omega}$  invertierbar und sei  $g: \tilde{\Omega} \rightarrow \Omega$  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  $\text{id} = g \circ f$ . Wir leiten dies ab und erhalten mit Hilfe der Kettenregel

$$1 = \text{id}' = g'(f(x_0)) \cdot f'(x_0).$$

Somit ist insbesondere  $f'(x_0) \neq 0$  und wir erhalten die behauptete Identität.  $\square$

Wir definieren iterativ höhere Ableitungen und zugehörige Funktionenräume.

**Definition 4.1.16.** Sei  $F$  ein Banachraum und sei  $\Omega \subset \mathbb{R}$  offen.

- (i) Fassen wir die Ableitung einer Funktion  $f: \Omega \rightarrow F$  als Funktion  $Df: \Omega \rightarrow F$  auf, so können wir induktiv die Abbildungen  $D^0 f = f: \Omega \rightarrow F$ ,  $D^1 f = Df: \Omega \rightarrow F$ ,  $D^2 f = D(Df): \Omega \rightarrow F$  sowie  $D^{n+1} f = D(D^n f): \Omega \rightarrow F$  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  $f^{(n)}: \Omega \rightarrow F$  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:

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

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  $C^0(\Omega) \equiv C^0(\Omega, \mathbb{R})$ ,  $C^k(\Omega) \equiv C^k(\Omega, \mathbb{R})$ ,  $\dots$

**Bemerkung 4.1.17.**  $\star$  Ist  $\overline{\Omega}$  kompakt, also z. B.  $\Omega \subset \mathbb{R}^n$  offen und beschränkt, so ist jede stetige Fortsetzung einer Funktion nach  $\overline{\Omega}$  automatisch auch beschränkt.

**Definition 4.1.18.**

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

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

$\star$  Als Indices der Norm verwenden wir auch  $C^0$ ,  $C^0(\Omega, F)$ ,  $C^0(\overline{\Omega}, F)$ ,  $L^\infty$ ,  $L^\infty(\Omega, F)$ , gebräuchlich sind aber auch 0 und  $\infty$ . Ist  $F = \mathbb{R}$ , so lassen wir dieses Argument  $F$  auch weg.

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

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

$\star$  Gebräuchlich sind auch hier wieder alle Varianten wie bei der  $C^0$ -Norm oder  $\|\cdot\|_k$ .

Um nachweisen zu können, dass diese Räume Banachräume sind, benötigen wir, dass Stetigkeit unter gleichmäßiger Konvergenz erhalten ist.

**Theorem 4.1.19.** *Seien  $E, F$  metrische Räume. Sei  $(f_n)_n$  eine Folge von stetigen Funktionen  $f_n: E \rightarrow F$ , die gleichmäßig gegen eine Funktion  $f: E \rightarrow F$  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.  $\square$

*Alternativbeweis mit Doppelfolgen.*  $\star$  Sei  $x_0 \in E$  beliebig. Sei  $(x_m)_m \subset E$  eine Folge mit  $x_m \rightarrow x_0$  für  $m \rightarrow \infty$ . Wir wollen nachweisen, dass

$$f(x_0) = \lim_{m \rightarrow \infty} f(x_m)$$

gilt. Dies ist äquivalent zu

$$\lim_{n \rightarrow \infty} \underbrace{\lim_{m \rightarrow \infty} f_n(x_m)}_{=f_n(x_0)} = \lim_{m \rightarrow \infty} \underbrace{\lim_{n \rightarrow \infty} f_n(x_m)}_{=f(x_m)}$$

und folgt direkt aus unserem Resultat über die Vertauschbarkeit von Grenzwerten bei Doppelfolgen, wenn die Konvergenz bezüglich des einen Index gleichmäßig ist.  $\square$

Ein entsprechendes Resultat gilt auch für Reihen:

**Theorem 4.1.20.**  $\star$  Sei  $E$  ein metrischer Raum,  $F$  ein Banachraum und  $f_n: E \rightarrow F$  eine Folge stetiger Funktionen. Konvergiert die Reihe  $((f_n))_n$  für  $n \rightarrow \infty$  gleichmäßig gegen eine Funktion  $f: E \rightarrow F$ , so ist  $f$  stetig.

*Beweis.* Wir wenden Theorem 4.1.19 auf die Folge  $(s_n)_n$  der Partialsummen

$$s_n(x) = \sum_{k=0}^n f_k(x)$$

an.  $\square$

**Theorem 4.1.21.** Sei  $\Omega \subset \mathbb{R}$  offen und sei  $F$  ein Banachraum. Sei  $k \in \mathbb{N}$ . Dann ist  $C^k(\overline{\Omega}, F)$  mit der  $C^k$ -Norm ein Banachraum.

*Beweis.*  $\star$  Übung. Benutze, dass der Grenzwert gleichmäßig konvergenter stetiger Funktionen nach Theorem 4.1.19 wieder stetig ist und dass  $C^0(\overline{\Omega}, F)$  ein Banachraum ist.  $\square$

**Proposition 4.1.22.** Sei  $\Omega \subset \mathbb{R}$  offen, sei  $n \in \mathbb{N}_{>0}$  und sei  $f: \Omega \rightarrow \mathbb{R}^n$  mit  $f = (f^i)_{1 \leq i \leq n}$  eine Abbildung. Dann ist  $f$  genau dann in  $x_0$  differenzierbar, wenn die Komponenten  $f^i$ ,  $1 \leq i \leq n$ , in  $x_0$  differenzierbar sind und es gilt dann

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

*Beweis.* Übung.  $\square$

**Definition 4.1.23.** Seien  $E$  ein metrischer Raum und  $f: E \rightarrow \mathbb{R}$  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  $f(x_0) \leq f(x)$  für alle  $x \in U$  existiert.
- (ii) Dann besitzt  $f$  in  $x_0$  ein **(globales) Minimum**, falls  $f(x_0) \leq f(x)$  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 4.1.24.** *Seien  $a < b \in \mathbb{R}$ , also  $a, b \in \mathbb{R}$  und  $a < b$ , sei  $f: (a, b) \rightarrow \mathbb{R}$  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  $f'(x_0) = 0$ .  $\square$

**Definition 4.1.25.** Sei  $\Omega \subset \mathbb{R}$  offen. Sei  $f: \Omega \rightarrow \mathbb{R}$  eine in  $x_0 \in \Omega$  differenzierbare Funktion mit  $f'(x_0) = 0$ . Dann heißt  $x_0$  **kritischer Punkt** von  $f$ .

**Bemerkung 4.1.26.**

- (i) Proposition 4.1.24 zeigt, dass jedes lokale Extremum einer auf einer offenen Teilmenge von  $\mathbb{R}$  definierten Funktion ein kritischer Punkt ist.
- (ii)  $\mathbb{R} \ni x \mapsto \frac{1}{4}x^4 - \frac{1}{2}x^2 \in \mathbb{R}$  besitzt drei lokale Extrema, nämlich  $-1, 0, 1$ , 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 4.1.27** (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

$$m := \inf_{[a, b]} f \quad \text{und} \quad M := \sup_{[a, b]} f.$$

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 4.1.24 folgt also  $f'(\xi) = 0$ .  $\square$

**Theorem 4.1.28** (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  $g: [a, b] \rightarrow \mathbb{R}$  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.  $\square$

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

**Proposition 4.1.29.** Sei  $\Omega \subset \mathbb{R}$  offen,  $F$  ein Banachraum und sei  $f: \Omega \rightarrow F$  in  $x_0$  differenzierbar. Sei  $\varphi: F \rightarrow \mathbb{R}$  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  $\varphi$  auf die Gleichung (4.1) an und beachte, dass  $\varphi(o(|x - x_0|)) \subset o(|x - x_0|)$  gilt. Details: Übung.  $\square$

**Lemma 4.1.30.** 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.

„ $\geq$ “: 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  $|\langle u, \varphi \rangle| \leq \|\varphi\| \cdot \|u\|$ .

„ $\leq$ “: 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

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

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  $\|\varphi_0\| = \sup_{\|x\|=1} |\langle x, u \rangle| \leq \sup_{\|x\|=1} \|x\| \cdot \|u\|$ .

$\|u\| = 1$  aufgrund der Cauchy-Schwarzschen Ungleichung. Schließlich gelten noch  $\|\varphi_0\| = \sup_{\|x\|=1} |\langle x, u \rangle| \geq |\langle u, u \rangle| = \|u\|^2 = 1$  und  $\varphi_0(u) = \|u\|^2 = 1$ .  $\square$

**Beispiel 4.1.31.** ★ Man überlegt sich am Beispiel einer Spirale

$$\mathbb{R} \ni t \mapsto (\sin t, \cos t, t),$$

dass der Mittelwertsatz für vektorwertige Funktionen nicht genau in der Form des obigen Mittelwertsatzes gelten kann.

**Theorem 4.1.32** (Vektorwertiger Mittelwertsatz). Seien  $a < b \in \mathbb{R}$ . Sei  $F$  ein Banachraum, sei  $f: [a, b] \rightarrow F$  stetig und in  $(a, b)$  differenzierbar. Dann gilt

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

*Beweis.* Sei  $\varphi \in F^*$  mit  $\|\varphi\| = 1$  beliebig. Wir definieren eine Funktion  $g: [a, b] \rightarrow \mathbb{R}$  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 4.1.30

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

Da dies für beliebige  $\varphi \in F^*$  mit  $\|\varphi\| = 1$  gilt, folgt die Behauptung nun aus Lemma 4.1.30.  $\square$

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 4.1.33.** Sei  $\Omega \subset \mathbb{R}$  ein offenes Intervall, sei  $F$  ein Banachraum und sei  $f: \Omega \rightarrow F$  differenzierbar. Seien  $x, z_0 \in \Omega$ . Dann gilt

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

*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  $g'(x) = f'(x) - f'(z_0)$  folgt die Behauptung nun direkt aus dem (vektorwertigen) Mittelwertsatz.  $\square$

**Bemerkung 4.1.34.**  $\star$  In Korollar 4.1.33 dürfen wir überall, auch im Beweis,  $f'(z_0)$  durch ein anderes Element  $v \in F$  ersetzen.

**Korollar 4.1.35.** Sei  $\Omega \subset \mathbb{R}$  ein offenes Intervall,  $F$  ein Banachraum und  $f: \Omega \rightarrow F$  differenzierbar. Gilt  $f' \equiv 0$ , so ist  $f$  eine konstante Abbildung.

*Beweis.* Dies folgt direkt aus dem vektorwertigen Mittelwertsatz.  $\square$

**Definition 4.1.36.** Sei  $I \subset \mathbb{R}$  ein Intervall. Eine Funktion  $f: I \rightarrow \mathbb{R}$  heißt

- (i) **monoton wachsend**, falls für alle  $x_1, x_2 \in I$  aus  $x_1 \leq x_2$  die Ungleichung  $f(x_1) \leq f(x_2)$  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.
- (iii)  $\star$  **(strikt) monoton fallend**, falls für alle  $x_1, x_2 \in I$  aus  $x_1 < x_2$  die Ungleichung  $f(x_1) \geq f(x_2)$  bzw.  $f(x_1) > f(x_2)$  im strikten Fall folgt.

**Proposition 4.1.37.** Sei  $f: (a, b) \rightarrow \mathbb{R}$  differenzierbar.

- (i) Dann ist  $f$  genau dann monoton wachsend, wenn  $f'(x) \geq 0$  für alle  $x \in (a, b)$  gilt.
- (ii) Gilt  $f'(x) > 0$  für alle  $x \in (a, b)$ , so ist  $f$  strikt monoton wachsend.
- (iii)  $\star$  Für (strikt) monoton fallende Funktionen gelten analoge Aussagen.

*Beweis.*

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

**Bemerkung 4.1.38.**  $\star$  Das Beispiel  $\mathbb{R} \ni x \mapsto x^3$  zeigt, dass strikte Monotonie im Allgemeinen nicht  $f'(x) > 0$  impliziert.

**Theorem 4.1.39.** Sei  $I = (a, b) \subset \mathbb{R}$ . Sei  $f: I \rightarrow \mathbb{R}$  differenzierbar und gelte  $f' \neq 0$  in  $I$ . Dann ist  $J := f(I) \subset \mathbb{R}$  offen und  $f$  besitzt eine differenzierbare Inverse  $f^{-1}: J \rightarrow I$  mit

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

*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  $f'(x) \neq 0$  für  $x \in (a, b)$  ausgeschlossen ist.
- (iv) Dieselbe Argumentation liefert, dass  $f$  eine offene Abbildung ist. Somit ist  $f^{-1}: J \rightarrow I$  stetig,  $f: I \rightarrow J$  also ein Homöomorphismus.

(v) Da  $f$  ein Homöomorphismus ist, folgt

$$x \rightarrow x_0 \iff f(x) \rightarrow f(x_0).$$

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  $x \rightarrow x_0$  mit  $x \neq x_0$  oder, aufgrund der obigen Äquivalenz und der Injektivität äquivalent dazu,  $\xi \rightarrow \xi_0$  mit  $\xi \neq \xi_0$  und erhalten die Behauptung.  $\square$

**Beispiel 4.1.40.**  $\star$  Für  $f_n: \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ ,  $n \in \mathbb{N}_{>0}$ , mit  $x \mapsto x^n$  ist  $g(x) = x^{1/n}$  die Umkehrfunktion und es gilt

$$g'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{nx^{\frac{1}{n}(n-1)}}.$$

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

Als wichtiges Hilfsmittel benötigen wir den folgenden Satz.

**Theorem 4.2.1** (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*

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

**Bemerkung 4.2.2.**  $\star$  Nach dem Satz von Rolle gilt  $g(b) - g(a) \neq 0$ . Somit ist die linke Seite wohldefiniert. Der übliche Mittelwertsatz, auf Zähler und Nenner getrennt angewandt, liefert in der vorliegenden Situation die Existenz von  $\xi, \eta \in I$  mit

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

Der Unterschied zum verallgemeinerten Mittelwertsatz ist, dass in diesem zusätzlich  $\xi = \eta$  gilt.

*Beweis von Theorem 4.2.1.*

- (i) Aufgrund des Zwischenwertsatzes gilt  $g'(x) > 0$  für alle  $x \in (a, b)$  oder  $g'(x) < 0$  für alle  $x \in (a, b)$ . Nach Proposition 4.1.37 ist  $g$  daher strikt monoton. Also gilt insbesondere  $g(a) \neq g(b)$ .
- (ii) Daher können wir die Funktion  $h$  mit

$$h(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a))$$

definieren. Es gilt  $h(a) = f(a) = h(b)$ . Somit finden wir nach dem Satz von Rolle ein  $\xi \in I$  mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(\xi)$$

wie behauptet.  $\square$

Der folgende Beweis lässt sich gut auf den vektorwertigen Fall verallgemeinern.

*Alternativer Beweis von Theorem 4.2.1  $\star$ .*

- (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 4.1.39 besitzt  $g$  daher eine in  $J := (\alpha, \beta)$  differenzierbare Inverse  $\varphi$  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  $\xi = \varphi(\gamma)$ .  $\square$

Es gilt die folgende vektorwertige Verallgemeinerung

**Theorem 4.2.3.**  $\star$  Sei  $F$  ein Banachraum, sei  $I = (a, b) \subset \mathbb{R}$  ein Intervall und seien  $f \in C^0([a, b], F)$ ,  $g \in C^0([a, b])$  in  $I$  differenzierbar. Gilt  $g'(x) \neq 0$  für alle  $x \in I$ , so folgt

$$\|f(b) - f(a)\| \leq \sup_{\xi \in (a, b)} \left\| \frac{f'(\xi)}{g'(\xi)} \right\| \cdot |g(b) - g(a)|.$$

*Beweisidee.* Gehe wie beim Beweis des verallgemeinerten Mittelwertsatzes vor und benutze den vektorwertigen Mittelwertsatz statt der reellen Version.  $\square$

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

**Proposition 4.2.4.** 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

$$\lim_{\substack{U \ni x \rightarrow a \\ x \neq a}} \frac{f(x)}{g(x)}, \quad \text{falls} \quad \lim_{\substack{U \ni x \rightarrow a \\ x \neq a}} \frac{f'(x)}{g'(x)}$$

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  $x \rightarrow a$  und somit  $\xi(x) \rightarrow a$  und erhalten die Behauptung.  $\square$

**Bemerkung 4.2.5.**

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

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

- (ii)  $\star$  Weiterhin gilt unter Benutzung von Schulwissen

$$\lim_{y \rightarrow \infty} \frac{\log(y)}{y} = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\log(1/x)}{1/x} = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\frac{x}{1} \left(-\frac{1}{x^2}\right)}{-\frac{1}{x^2}} = \lim_{x \rightarrow 0} x = 0.$$

- (iii) ★ Existieren die Grenzwerte  $\lim_{x \rightarrow a} \frac{f(x)-f(a)}{x-a} = f'(a)$  und  $\lim_{x \rightarrow a} \frac{g(x)-g(a)}{x-a} = g'(a)$  sowie der Quotient  $\frac{f'(a)}{g'(a)}$ , so funktioniert folgender einfacherer Beweis, den man von rechts her liest:

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{g(x) - g(a)} = \lim_{x \rightarrow a} \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} = \frac{\lim_{x \rightarrow a} \frac{f(x)-f(a)}{x-a}}{\lim_{x \rightarrow a} \frac{g(x)-g(a)}{x-a}} = \frac{f'(a)}{g'(a)}$$

- (iv) ★ Gegebenenfalls muss man etwas umformen um einen Quotienten der Gestalt „ $\frac{0}{0}$ “ zu bekommen.
- (v) ★ Es ist möglich, von der Existenz des Grenzwertes von  $\frac{f''(x)}{g''(x)}$  auf die des Grenzwertes von  $\frac{f'(x)}{g'(x)}$  und dann weiter auf die des Grenzwertes von  $\frac{f(x)}{g(x)}$  zu schließen.
- (vi) ★ Falls der Grenzwert der Funktionswerte existiert, braucht der Grenzwert der Ableitungen nicht zu existieren, siehe [6].
- (vii) ★ Linksseitige Umgebungen definieren wir analog zu rechtsseitigen.

Dazu gibt es noch eine Variante mit Grenzwerten gleich unendlich.

**Proposition 4.2.6.** ★ Sei  $U$  eine einseitige Umgebung von  $a \in \mathbb{R}$  und seien  $f, g \in C^0(U)$  in  $U \setminus \{a\}$  differenzierbar mit  $g'(x) \neq 0$  für alle  $x \in U$ . Gelte

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \infty.$$

Dann existiert

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)}, \quad \text{falls} \quad \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

existiert, und die beiden Grenzwerte stimmen überein.

*Beweis.* Sei ohne Einschränkung  $a$  der linke Endpunkt von  $U$ , also  $I = [a, b] \subset U$  mit  $a < b$ . Für beliebige  $x, x_0$  mit  $a < x < x_0 < b$  gibt es aufgrund des verallgemeinerten Mittelwertsatzes ein  $\xi \in (x, x_0)$  mit

$$\frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi)}{g'(\xi)}.$$

Sei

$$\alpha = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.$$

Sei  $\varepsilon > 0$ . Dann gibt es ein  $x_0 \in (a, b)$  mit  $\frac{f'(\xi)}{g'(\xi)} \leq \alpha + \varepsilon$  für alle  $\xi \in (a, x_0)$ . Benutzen wir solche  $\xi$ , so erhalten wir aus der obigen Folgerung aus dem verallgemeinerten Mittelwertsatz

$$\overline{\lim}_{x \rightarrow a} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} \leq \alpha + \varepsilon.$$

Wegen  $\lim_{x \rightarrow a} g(x) = \infty$  erhalten wir (Die erste Gleichheit gilt auch ohne Limes und man sieht sie leicht durch Multiplikation mit  $g(x) - g(x_0)$ .)

$$\begin{aligned} \overline{\lim}_{x \rightarrow a} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} &= \overline{\lim}_{x \rightarrow a} \left\{ \frac{f(x)}{g(x)} \left( 1 + \frac{g(x_0)}{g(x) - g(x_0)} \right) - \frac{f(x_0)}{g(x) - g(x_0)} \right\} \\ &= \overline{\lim}_{x \rightarrow a} \frac{f(x)}{g(x)}. \end{aligned}$$

Da  $\varepsilon > 0$  beliebig war, erhalten wir  $\overline{\lim}_{x \rightarrow a} \frac{f(x)}{g(x)} \leq \alpha$ .

Entsprechend schließt man von

$$\underline{\lim}_{x \rightarrow a} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} \geq \alpha - \varepsilon \quad \text{auf} \quad \underline{\lim}_{x \rightarrow a} \frac{f(x)}{g(x)} \geq \alpha.$$

Es gilt also

$$\alpha \leq \liminf_{x \rightarrow a} \frac{f(x)}{g(x)} \leq \limsup_{x \rightarrow a} \frac{f(x)}{g(x)} \leq \alpha.$$

Die Behauptung folgt.  $\square$

**Proposition 4.2.7.**  $\star$  Sei  $U$  eine Umgebung von  $\infty \in \overline{\mathbb{R}}$ . Seien  $f, g: U \rightarrow \mathbb{R}$  auf  $U \cap \mathbb{R}$  stetig und differenzierbar mit  $g'(x) \neq 0$  für alle  $x \in U \cap \mathbb{R}$ . Gelte

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = 0 \quad \text{oder} \quad \lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = \infty.$$

Dann existiert

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)}, \quad \text{falls} \quad \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$$

existiert, und die beiden Grenzwerte stimmen überein.

*Beweis.* Definiere

$$\begin{aligned} \varphi(x) &:= f(x^{-1}), & \varphi(0) &:= \lim_{x \rightarrow \infty} f(x), \\ \psi(x) &:= g(x^{-1}) & \text{und} & \quad \psi(0) := \lim_{x \rightarrow \infty} g(x). \end{aligned}$$

Dann erfüllen  $\varphi$  bzw.  $\psi$  die Voraussetzungen (Stetigkeit, Differenzierbarkeit,  $\psi' \neq 0$ ) unserer bereits bewiesenen l'Hospitalischen Regeln und wir erhalten

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\varphi(x)}{\psi(x)} = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\varphi'(x)}{\psi'(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$$

wie behauptet.  $\square$

**4.3. Differentiation von Funktionenfolgen.** Wir haben in Theorem 4.1.19 bereits gesehen, dass gleichmäßige Konvergenz Stetigkeit erhält. Gleichmäßige Konvergenz sichert auch, dass wir den Grenzwert in der Folge und die Ableitung, letztlich auch ein Grenzwert, vertauschen dürfen, dass also

$$\lim_{n \rightarrow \infty} f'_n = \left( \lim_{n \rightarrow \infty} f_n \right)'$$

gilt.

**Theorem 4.3.1.** Sei  $\Omega \subset \mathbb{R}$  offen. Sei  $F$  ein Banachraum und  $f_n: \Omega \rightarrow F$  eine Folge stetig differenzierbarer Funktionen mit  $f_n \rightarrow f$  und  $f'_n \rightrightarrows g$  für eine Funktion  $g: \Omega \rightarrow F$ . 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 4.1.19, 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

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

$$\equiv I_1 + I_2.$$

Wir zeigen zunächst, dass der Term  $I_2$  klein wird. Nach Korollar 4.1.33 mit Bemerkung 4.1.34 und  $g(x_0)$  statt  $f'(z_0)$ , 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  $n$ . Zusammengekommen folgt gerade die Behauptung.  $\square$

Ein entsprechendes Resultat gilt auch für Reihen von Funktionen.

**Theorem 4.3.2.** *★ Sei  $\Omega \subset \mathbb{R}$  offen. Sei  $F$  ein Banachraum und sei  $f_n: \Omega \rightarrow F$  eine Folge stetig differenzierbarer Funktionen. Konvergiere  $((f_n(x)))_{n \in \mathbb{N}}$  für jedes  $x \in \Omega$ . Definiere die Funktion  $f: \Omega \rightarrow F$  durch*

$$f(x) := \sum_{n=0}^{\infty} f_n(x).$$

*Nehme an, dass die Reihe  $((f'_n(x)))_n$  bzw. die Folge  $(s_k)_k$  mit  $s_k(x) = \sum_{n=0}^k f'_n(x)$  gleichmäßig in  $x$  gegen eine Funktion  $g: \Omega \rightarrow F$  konvergiert. Dann ist  $f$  stetig differenzierbar und es gilt  $f' = g$ .*

*Beweis.* Wir wenden Theorem 4.3.1 auf die Folge  $(s_k)_k$  der Partialsummen an. Die Stetigkeit von  $g$  folgt aus Theorem 4.1.20.  $\square$

Damit können wir die Ableitung einer Potenzreihe im Inneren ihres Konvergenzradius durch gliedweises Differenzieren ermitteln.

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

**Theorem 4.3.3.** *Sei  $((a_n x^n))_{n \in \mathbb{N}}$  eine Potenzreihe in  $\mathbb{R}$  mit Konvergenzradius  $r > 0$ . Dann ist die Funktion*

$$\begin{aligned} f: B_r(0) &\rightarrow \mathbb{R}, \\ f(x) &:= \sum_{n=0}^{\infty} a_n x^n \end{aligned}$$

*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 Konvergenzradius 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  $((na_n x^{n-1}))_n$  konvergiert.

Nach Lemma 4.3.4 gilt  $\lim_{n \rightarrow \infty} n^{1/n} = 1$ . Damit folgt

$$\frac{1}{r} = \overline{\lim}_{n \rightarrow \infty} |a_n|^{1/n} = \overline{\lim}_{n \rightarrow \infty} (n|a_n|)^{1/n}$$

wie behauptet.

- (ii) Sei  $\rho < r$ . Dann konvergiert die Potenzreihe  $((na_n x^{n-1}))_n$  in  $B_\rho(0)$  gleichmäßig (Analysis I) und ist nach Theorem 4.3.2 (wir dürfen Differentiation und  $n \rightarrow \infty$  für die Partialsummen vertauschen) als gliedweise Ableitung der Potenzreihe  $((a_n x^n))_n$  in  $B_\rho(0)$  die Ableitung von  $f$ . Da  $\rho$  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 4.3.4.**  $\star$  Es gilt

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

*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.  $\square$

*Alternativbeweis mit Schulwissen.*  $\star$  Aufgrund der l'Hospitalschen Regel gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{1/n} &= \lim_{n \rightarrow \infty} \exp\left(\frac{\log n}{n}\right) = \exp\left(\lim_{n \rightarrow \infty} \frac{\log n}{n}\right) \\ &= \exp\left(\lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{1}\right) = \exp(0) = 1 \end{aligned}$$

wie behauptet.  $\square$

**Beispiel 4.3.5.** 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 4.4.1.** 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 [7].  $\square$

**Bemerkung 4.4.2.** Unsere Aussagen über Reihen und deren gliedweise Ableitungen erlauben es, zu zeigen, dass

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

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

*Beweis.* Übung. □

Unter Vorwegnahme des Integrals zeigen wir das Gronwallsche Lemma und wie man daraus die Eindeutigkeit der Lösung einer linearen gewöhnlichen Differentialgleichung erhält.

**Lemma 4.4.3** (Gronwallsches Lemma). *★ Seien  $a, b \geq 0$  und  $T \geq 0$ . Sei  $f: [0, T] \rightarrow \mathbb{R}_+$  stetig und gelte*

$$f(t) \leq a + b \int_0^t f(s) ds$$

für alle  $t \in [0, T]$ . Dann folgt

$$f(t) \leq ae^{bt}$$

für alle  $t \in [0, T]$ .

*Beweis.* Ohne Einschränkung dürfen wir annehmen, dass  $b > 0$  gilt, da die Aussage sonst trivial ist. Nach Voraussetzung erhalten wir

$$\begin{aligned} \frac{d}{dt} \left( e^{-bt} \int_0^t f(s) ds \right) &= -be^{-bt} \int_0^t f(s) ds + e^{-bt} f(t) \\ &\leq -be^{-bt} \int_0^t f(s) ds + e^{-bt} \left( a + b \int_0^t f(s) ds \right) \\ &\leq ae^{-bt}. \end{aligned}$$

Wir integrieren dies nun unter Benutzung des Hauptsatzes der Differential- und Integralrechnung, also  $\int_0^t \frac{d}{ds} f(s) ds = f(t) - f(0)$ , und erhalten

$$\begin{aligned} e^{-bt} \int_0^t f(s) ds &= e^{-bt} \int_0^t f(s) ds - 0 \\ &\leq \int_0^t ae^{-bs} ds = -\frac{a}{b} e^{-bs} \Big|_0^t = \frac{a}{b} (1 - e^{-bt}). \end{aligned}$$

Dies multiplizieren wir nun mit  $e^{bt}$  und verwenden nochmals die Integralungleichung aus der Voraussetzung und erhalten

$$f(t) \leq a + b \int_0^t f(s) ds \leq a + b \left[ \frac{a}{b} (e^{bt} - 1) \right] = ae^{bt}$$

wie behauptet. □

**Korollar 4.4.4.** *★ Die Lösung  $x$  aus Theorem 4.4.1 ist (wie dort behauptet) eindeutig bestimmt.*

*Beweis.* Seien  $x$  und  $y$  Lösungen der Differentialgleichungen  $\dot{x} = Ax$  bzw.  $\dot{y} = Ay$ . Wir definieren  $\varphi(t) := x(t) - y(t)$  und erhalten

$$\dot{\varphi}(t) = \dot{x}(t) - \dot{y}(t) = Ax(t) - Ay(t) = A(x(t) - y(t)) = A\varphi(t).$$

Wir integrieren dies und erhalten mit dem Hauptsatz der Differential- und Integralrechnung

$$\varphi(t) = \varphi(0) + \int_0^t A\varphi(s) ds.$$

Nun betrachten wir die Beträge beider Seiten, wenden die Dreiecksungleichung an, ziehen die Beträge in das Integral hinein und erhalten unter Verwendung der Operatornorm  $\|A\|$

$$\|\varphi(t)\| \leq \|\varphi(0)\| + \int_0^t \|A\| \cdot \|\varphi(s)\| ds,$$

also für  $f(t) := \|\varphi(t)\|$  die Ungleichung

$$f(t) \leq f(0) + \int_0^t \|A\| \cdot f(s) ds.$$

Ist  $f(0) = 0$ , so folgt aufgrund des Gronwallschen Lemmas  $f \equiv 0$  und daher die behauptete Eindeutigkeit.  $\square$

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

**Bemerkung 4.4.5** (Potenzfunktionen).  $\star$  Seien  $n, m \in \mathbb{N}_{>0}$ . Dann ist die Funktion  $f_n: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  mit  $f_n(x) = x^n$  unendlich oft differenzierbar mit  $f'_n(x) = nx^{n-1}$  und wegen  $f'_n(x) > 0$  für  $x > 0$  strikt monoton wachsend. Somit ist  $f_n: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  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  $(a^n)^{\frac{1}{m}} = (a^{\frac{1}{m}})^n$ , wie man durch Potenzieren mit  $m$  sieht. Wir definieren

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

Damit gelten die Potenzgesetze

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

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

**Lemma 4.4.6** (Exponentialfunktion). Die Exponentialfunktion  $\exp: \mathbb{R} \rightarrow \mathbb{R}_{>0}$  ist ein Homöomorphismus und besitzt eine  $C^\infty$ -Inverse:  $\log: \mathbb{R}_{>0} \rightarrow \mathbb{R}$  mit  $\log'(x) = \frac{1}{x}$ .

*Beweis.* Beispiel 4.3.5 liefert die Differenzierbarkeit und  $\exp' = \exp$ . Für  $x > 0$  gilt klarerweise  $\exp(x) > 0$ . Wegen  $1 = \exp(0) = \exp(x) \cdot \exp(-x)$  ist  $\exp(x)$  für alle  $x \in \mathbb{R}$  positiv. Weiterhin folgt wegen  $\exp(x) \rightarrow \infty$  für  $x \rightarrow \infty$  daraus auch  $\exp(x) \rightarrow 0$  für  $x \rightarrow -\infty$ . Dank der Stetigkeit ist  $\exp: \mathbb{R} \rightarrow \mathbb{R}_{>0}$  somit surjektiv. Nach Theorem 4.1.39 ist  $\exp$  ein Homöomorphismus (siehe Beweis), die Inverse  $\log$  ist differenzierbar und es gilt  $\log'(\exp(x)) = \frac{1}{\exp(x)}$ . Da sich jedes  $y > 0$  als  $\exp(x)$  darstellen läßt, erhalten wir den Ausdruck für  $\log'$ . Daher ist der Logarithmus beliebig oft differenzierbar.  $\square$

**Lemma 4.4.7.** 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.  $\square$

Motiviert durch diese Beziehung für rationale Exponenten definieren wir

**Definition 4.4.8.** Seien  $a > 0$  und  $x \in \mathbb{R}$ . Dann definieren wir

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

**Bemerkung 4.4.9.**  $\star$

- (i) Für  $x \in \mathbb{Q}$  stimmt diese Definition nach Lemma 4.4.7 mit der bereits bekannten Definition überein. Definieren wir  $a^x := \lim_{n \rightarrow \infty} a^{x_n}$  mit  $\mathbb{Q} \ni x_n \rightarrow x \in \mathbb{R}$ , so erhalten wir aufgrund der Stetigkeit Übereinstimmung mit Definition 4.4.8.
- (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  $x \rightarrow \infty$  und fällt schneller als jede Potenz für  $x \rightarrow -\infty$ .

**Proposition 4.4.10.** Es gelten für beliebiges  $n \in \mathbb{N}$

$$\lim_{x \rightarrow \infty} \frac{e^x}{x^n} = \infty \quad \text{und} \quad \lim_{x \rightarrow -\infty} e^x |x|^n = 0.$$

*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.  $\square$

**Proposition 4.4.11.** Der Logarithmus erfüllt für alle  $x, y, \varepsilon \in \mathbb{R}_{>0}$  und alle  $\lambda \in \mathbb{R}$

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

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.  $\square$

**Definition 4.4.12.** Sei  $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ . Sei  $x \in C^\infty(\mathbb{R}, \mathbb{R}^2)$  die eindeutig bestimmte Lösung der Differentialgleichung

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

Dann definieren wir die Funktionen **Sinus**  $\sin: \mathbb{R} \rightarrow \mathbb{R}$  und **Kosinus**  $\cos: \mathbb{R} \rightarrow \mathbb{R}$  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 4.4.13.**

- (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.
- (iii) ★ Es lässt sich zwar leicht überprüfen, dass die über Potenzreihen definierten Funktionen  $\sin$  und  $\cos$  die Differentialgleichung lösen, ohne weitere Theorie ist jedoch nicht klar, dass dies die einzigen Lösungen sind.

**Proposition 4.4.14.** *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.  $\square$

Sinus und Kosinus sind ungerade bzw. gerade Funktionen.

**Proposition 4.4.15.** *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.  $\square$

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  $\alpha$  und  $\beta$ . Dies benutzt jedoch implizit, dass die Abbildung, die  $\alpha$  auf die Drehmatrix um den Winkel  $\alpha$  abbildet, ein Gruppenhomomorphismus ist.

**Proposition 4.4.16** (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  $\Phi$  und  $\Psi$  also dieselbe lineare gewöhnliche Differentialgleichung erfüllen und  $\Phi(0) = \Psi(0)$  gilt, folgt  $\Phi = \Psi$ .  $\square$

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  $\sin = \text{im } \cos = [-1, 1]$  gilt. Dazu zeigen wir zunächst

**Lemma 4.4.17.** *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  $x \rightarrow \infty$  konvergiert die rechte Seite gegen minus Unendlich. Dies widerspricht  $\cos^2 x \leq 1$ .  $\square$

**Definition 4.4.18** (Die Kreiszahl  $\pi$ ). 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 4.4.19.** *Auf  $[0, \frac{\pi}{2}]$  ist der Sinus strikt monoton wachsend und es gilt*

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

*Beweis.* Nach Definition von  $\pi$  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  $\sin \frac{\pi}{2} = 1$ .  $\square$

**Korollar 4.4.20.** *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.  $\square$

Allgemeiner gilt

**Korollar 4.4.21.** *Für alle  $x \in \mathbb{R}$  gelten*

$$\begin{aligned} \sin\left(x + \frac{\pi}{2}\right) &= \cos x, & \sin(x + \pi) &= -\sin x, & \sin(x + 2\pi) &= \sin x, \\ \cos\left(x + \frac{\pi}{2}\right) &= -\sin x, & \cos(x + \pi) &= -\cos x, & \cos(x + 2\pi) &= \cos x. \end{aligned}$$

*Beweis.* Benutze die Additionstheoreme.  $\square$

**Korollar 4.4.22.** *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 4.4.21 sind der Sinus und der Kosinus mit Periode  $2\pi$  periodisch.

**Definition 4.4.23.** Sei  $E$  ein Vektorraum und  $F$  eine Menge. Sei  $0 \neq \xi \in E$ . Eine Abbildung  $f: E \rightarrow F$  heißt mit Periode  $\xi$  **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 4.4.24.** Die Abbildung

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

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

*Beweis.* Im Intervall  $[0, \frac{\pi}{2}]$  ist der Sinus strikt monoton wachsend von 0 bis 1. Somit ist  $\varphi$  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

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

$\Gamma_+$  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  $[\frac{\pi}{2}, \pi]$ ,  $[\pi, \frac{3\pi}{2}]$  sowie  $[\frac{3\pi}{2}, 2\pi]$  liefern nun die Behauptung über das Bild. Die Injektivität folgt aus der Injektivität auf jedem dieser Intervalle und Vorzeichenbetrachtungen.  $\square$

**Proposition 4.4.25** (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 Konvergenzradius 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.  $\square$

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

**Proposition 4.4.26.** Es gelten für alle  $x \in \mathbb{R}$

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

und

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

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 4.4.27.** *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  $z = i\pi + 2\pi ik$  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  $|e^z| = e^x \cdot |e^{iy}| = e^x$ . 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  $2\pi$ -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 4.4.28.** *Die Exponentialfunktion  $\mathbb{C} \ni z \mapsto e^z \in \mathbb{C}$  ist  $2\pi i$ -periodisch.*

**Definition 4.4.29.**

- (i) Für  $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$  definieren wir den **Tangens**,  $\tan x$ , durch  $\tan x = \frac{\sin x}{\cos x}$ .
- (ii)  $\star$  Für  $x \in (0, \pi)$  definieren wir den **Cotangens** bzw. **Kotangens**,  $\cot x$ , durch  $\cot x = \frac{\cos x}{\sin x}$ .

**Bemerkung 4.4.30.**  $\star$  Tangens und Cotangens sind unendlich oft differenzierbar und bilden ihren jeweiligen Definitionsbereich homöomorph auf  $\mathbb{R}$  ab. Es gelten

$$\tan' x = \frac{1}{\cos^2 x} = 1 + \tan^2 x \quad \text{und} \quad \cot' x = -\frac{1}{\sin^2 x} = -1 - \cot^2 x.$$

**Definition 4.4.31** (Arcusfunktionen).  $\star$  Wir definieren

$$\begin{aligned} \arcsin &:= \left( \sin \Big|_{(-\frac{\pi}{2}, \frac{\pi}{2})} \right)^{-1} : (-1, 1) \rightarrow \left( -\frac{\pi}{2}, \frac{\pi}{2} \right), \\ \arccos &:= \left( \cos \Big|_{(0, \pi)} \right)^{-1} : (-1, 1) \rightarrow (0, \pi), \\ \arctan &:= \left( \tan \Big|_{(-\frac{\pi}{2}, \frac{\pi}{2})} \right)^{-1} : \mathbb{R} \rightarrow \left( -\frac{\pi}{2}, \frac{\pi}{2} \right), \\ \operatorname{arccot} &:= \left( \cot \Big|_{(0, \pi)} \right)^{-1} : \mathbb{R} \rightarrow (0, \pi). \end{aligned}$$

**Bemerkung 4.4.32.**  $\star$

- (i) Für  $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$  gelten  $\sin' x = \cos x \neq 0$  und  $\tan' x = \frac{1}{\cos^2 x} \neq 0$  und für  $x \in (0, \pi)$  gelten  $\cos' x = -\sin x \neq 0$  sowie  $\cot' x = -\frac{1}{\sin^2 x} \neq 0$ . Daher sind die Arcusfunktionen wohldefiniert.
- (ii) Es gilt

$$\begin{aligned} \arcsin' x &= \frac{1}{\sqrt{1-x^2}}, & \arccos' x &= -\frac{1}{\sqrt{1-x^2}}, \\ \arctan' x &= \frac{1}{1+x^2}, & \operatorname{arccot}' x &= -\frac{1}{1+x^2}. \end{aligned}$$

*Beweis.* Übung. □

**Bemerkung 4.4.33.**

- (i) Aus der Eulerschen Formel erhalten wir

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

Somit definieren wir die komplexen trigonometrischen Funktionen durch

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

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

- (ii) ★ Mit rein komplexen Argumenten erhalten wir Funktionen mit neuen Namen:

$$\sinh x := -i \sin ix = \frac{1}{2} (e^x - e^{-x}), \quad \cosh x := \cos ix = \frac{1}{2} (e^x + e^{-x})$$

und

$$\tanh x := \frac{\sinh x}{\cosh x}.$$

- (iii) ★ Es gelten die folgenden Beziehungen:

$$\cosh^2 x - \sinh^2 x = 1,$$

$$\cosh(x+y) = \cosh x \cdot \cosh y + \sinh x \cdot \sinh y,$$

$$\sinh(x+y) = \sinh x \cdot \cosh y + \cosh x \cdot \sinh y,$$

$$\sinh' x = \cosh x,$$

$$\cosh' x = \sinh x,$$

$$\tanh' x = \frac{1}{\cosh^2 x}.$$

*Beweis.* Übung. □

**Bemerkung 4.4.34** (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  $r \geq 0$  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  $\mathbb{R}_{\leq 0} \subset \mathbb{C}$  unstetig.

**Lemma 4.4.35.**

- (i) Die Exponentialfunktion bildet den Streifen

$$S = \{z = x + iy : -\pi < y \leq \pi\}$$

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

- (ii) Sei  $\log: \mathbb{C} \setminus \{0\} \rightarrow S$  die Umkehrfunktion von  $\exp: S \rightarrow \mathbb{C} \setminus \{0\}$ , so gilt  $\log' z = \frac{1}{z}$  für alle Punkte in  $\mathbb{C} \setminus \mathbb{R}_{\leq 0}$ .

*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. Polynome ★.

**Bemerkung 4.5.1** (Wiederholung).

- (i) Sei  $E$  ein Vektorraum. Dann sind die Polynome  $P(\mathbb{K}, E)$  mit Koeffizienten in  $E$  aus der linearen Algebra bekannt. Wir schreiben  $P(\mathbb{K})$  für  $P(\mathbb{K}, \mathbb{K})$ . Der Grad eines Polynoms  $\deg f$ , hat Werte in  $\mathbb{N} \cup \{-\infty\}$ , der führende Koeffizient Werte in  $E$ . In  $P(\mathbb{K})$  gilt  $\deg(fg) = \deg f + \deg g$ .
- (ii) Jedes Polynom  $p(x) = \sum_{k=0}^n a_k x^k$  lässt sich in der Form  $\sum_{k=0}^n b_k (x - x_0)^k$  schreiben. Dabei gilt  $a_n = b_n$ . Grad und führender Koeffizient sind also vom **Entwicklungspunkt**  $x_0$  unabhängig.

*Beweis.* Schreibe

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

und expandiere den Ausdruck mit Hilfe der binomischen Formel.  $\square$ 

- (iii) Sei  $f$  ein Polynom mit  $\deg f \geq 1$ . Gelte  $f(x_0) = 0$ . Dann gibt es ein Polynom  $g$  mit  $\deg g = (\deg f) - 1$  und  $f(x) = (x - x_0)g(x)$ .

*Beweis.* Polynomdivision. Alternativ kann man  $f$  als  $f(x) = \sum_{k=0}^n a_k (x - x_0)^k$  schreiben und sieht, dass  $a_0 = 0$  gilt.  $\square$

- (iv) Ein Polynom vom Grad  $n \geq 1$ , das nicht das Nullpolynom ist, besitzt höchstens  $n$  Nullstellen, wobei wir Nullstellen mit ihren Vielfachheiten zählen.

*Beweis.* Dies folgt direkt durch sukzessives Abspalten von Nullstellen von  $f$  in der Form  $f(x) = (x - x_0)g(x)$ .  $\square$

- (v) Zwei Polynome vom Grad kleiner oder gleich  $n \geq 1$ , die an  $n + 1$  Stellen übereinstimmen, sind gleich.

*Beweis.* Spalte induktiv alle Nullstellen des Polynoms, das als Differenz der beiden Polynome definiert ist, ab.  $\square$

- (vi) Mit der Gradfunktion bilden die Polynome einen Euklidischen Ring, d. h. es gibt eine Division mit Rest.
- (vii) Komplexe nichtkonstante Polynome besitzen mindestens eine Nullstelle (Fundamentalsatz der Algebra, siehe z. B. Funktionentheorievorlesung) und zerfallen daher in Linearfaktoren (Induktionsbeweis).

**Proposition 4.5.2** (Partialbruchzerlegung).

Seien  $f, g$  komplexe Polynome mit  $\deg f < \deg g$ . Seien  $\zeta_i$ ,  $1 \leq i \leq m$ , die (paarweise verschiedenen) Nullstellen von  $g$  mit Vielfachheiten  $m_i$ , dann besitzt  $f/g$  eine eindeutige Darstellung

$$\frac{f}{g}(z) \equiv \frac{f(z)}{g(z)} = \sum_{i=1}^m \sum_{j=1}^{m_i} \frac{c_{ij}}{(z - \zeta_i)^j}$$

mit  $c_{ij} \in \mathbb{C}$ .**Bemerkung 4.5.3.**

- (i) Der Quotient zweier Polynome bzw. Polynomfunktionen heißt eine **rationale** Funktion. Eine rationale Funktion  $f/g$  mit  $\deg f < \deg g$  heißt **echt rationale** Funktion.

- (ii) Mittels Polynomdivision lässt sich jede rationale Funktion als Summe eines Polynoms und einer echt rationalen Funktion darstellen.

*Beweis von Proposition 4.5.2.*

- (i) Existenz: Wir zeigen zunächst per Induktion nach  $\deg g$ , dass es solch eine Zerlegung gibt.

Im Fall  $\deg g = 1$  sind wir fertig.

Sei also  $\deg g = n$  und die Behauptung bereits für  $\deg g < n$  bewiesen.

Nach dem Fundamentalsatz der Algebra besitzt  $g$  eine Nullstelle  $\zeta \in \mathbb{C}$ . Sei  $m$  die Vielfachheit dieser Nullstelle, gelte also  $g(z) = (z - \zeta)^m g_1(z)$  für ein komplexes Polynom  $g_1$  mit  $g_1(\zeta) \neq 0$ . Wir behaupten, dass es ein  $c \in \mathbb{C}$  gibt, so dass

$$\frac{f(z)}{g(z)} = \frac{c}{(z - \zeta)^m}$$

eine echt rationale Funktion ist, bei der das Polynom im Nenner einen Grad  $< n$  besitzt. Dazu bringen wir den Ausdruck auf den Hauptnenner

$$\frac{f(z)}{g(z)} - \frac{c}{(z - \zeta)^m} = \frac{f(z) - c g_1(z)}{(z - \zeta)^m g_1(z)}.$$

Nun wählen wir  $c$  so, dass  $f(\zeta) - c g_1(\zeta) = 0$  gilt, also  $c = \frac{f(\zeta)}{g_1(\zeta)}$ . Wegen  $g_1(\zeta) \neq 0$  ist dies wohldefiniert. Somit ist  $\zeta$  eine Nullstelle des Zählers und des Nenners auf der rechten Seite. Wir können also den gemeinsamen Faktor  $(z - \zeta)$  kürzen und erhalten

$$\frac{f(z)}{g(z)} - \frac{c}{(z - \zeta)^m} = \frac{f_1(z)}{(z - \zeta)^{m-1} \cdot g_1(z)}$$

für ein Polynom  $f_1(z)$  vom Grad  $\leq \deg f - 1$ . Somit können wir die Induktionsvoraussetzung anwenden und die Behauptung folgt per Induktion.

- (ii) Eindeutigkeit: Sei  $f/g$  ohne Einschränkung vollständig gekürzt. Sei

$$h(z) = \sum_{i=1}^n \sum_{j=1}^{n_i} \frac{d_{ij}}{(z - \xi_i)^j}$$

eine weitere solche Darstellung. Dann wird  $(f/g)(z)$  genau dann unbeschränkt, wenn  $z$  gegen eine Nullstelle von  $g$  konvergiert. Daher muss bis auf Umnummerierung  $\zeta_i = \xi_i$  gelten. Somit gilt auch  $n = m$ . Die Übereinstimmungen  $m_i = n_i$  und  $c_{ij} = d_{ij}$  erhalten wir schließlich, indem wir beide Ausdrücke mit  $(z - \zeta_i)^k$ ,  $k = \max\{n_i, m_i\}$ , multiplizieren und den Grenzwert  $z \rightarrow \zeta_i$  betrachten. Dann erhalten wir in beiden Fällen  $c_{ik}$  und  $d_{ik}$  bzw. 0, falls diese Ausdrücke nicht definiert sind. Wir subtrahieren nun den Term  $\frac{c_{ik}}{(z - \zeta_i)^k}$  und erhalten per Induktion die Behauptung.  $\square$

#### 4.6. Taylorsche Formeln.

**Bemerkung 4.6.1.** ★ 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 4.6.2.** Sei  $E$  ein reeller Banachraum,  $I = (a, b)$  ein Intervall mit  $0, 1 \in I$  und sei  $f \in C^n(I, E)$  mit  $n \geq 1$ . 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

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

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

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

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

**Theorem 4.6.3** (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

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

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  $0 \leq t \leq 1$ .  $\varphi$  ist auch noch etwas über  $[0, 1]$  hinaus von der Klasse  $C^n$ . Daher können wir Theorem 4.6.2 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.  $\square$

**Definition 4.6.4** (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  $\Omega$  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 4.6.5.** 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

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

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  $o(|x - x_0|^n)$  ist.

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

**Bemerkung 4.6.6.**

- (i)  $\star$  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  $n \rightarrow \infty$  punktweise gegen Null konvergieren, wenn also für alle  $x \in B_\varepsilon(x_0)$

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

gilt.

- (ii)  $\star$  Gibt es zu  $x_0 \in \Omega$  ein solches  $\varepsilon > 0$  mit  $\lim_{n \rightarrow \infty} T_n(f, x_0)(x) = f(x)$  bzw., äquivalent dazu,  $\lim_{n \rightarrow \infty} R_n(f, x_0)(x) = 0$  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

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

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

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

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

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

Sei dazu  $x_0 = 0$ . Betrachte nun die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit

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

Per Induktion zeigt man, dass die Ableitungen für  $x < 0$  von der Gestalt  $f^{(k)}(x) = \frac{p_k(x)}{x^{3k}} e^{-\frac{1}{x^2}}$  für Polynome  $p_k$  sind. Nun gilt

$$\lim_{y \rightarrow \infty} e^{-y} y^l = 0$$

für alle  $l > 0$ . Daraus erhalten wir erstens

$$\lim_{\substack{x \rightarrow 0 \\ x < 0}} f^{(k)}(x) = 0$$

für alle  $k \in \mathbb{N}$  und weiterhin (nutze lokale Beschränktheit der Ableitungen, deren Verschwinden im Ursprung und integriere mehrfach)

$$|f(x)| \leq |x|^{k+1}$$

für alle  $k \in \mathbb{N}$  und  $x \in (-\varepsilon_k, \varepsilon_k)$  für  $\varepsilon_k > 0$ . Somit gilt  $f^{(k)}(0) = 0$  für alle  $k \in \mathbb{N}$ . Wir erhalten also  $f \in C^\infty$ . Aufgrund der Formel für  $f^{(k)}(0)$  folgt somit  $T(f, x_0) \equiv 0$ , jedoch gilt in keiner Umgebung des Ursprunges  $f \equiv 0$ .

Wir leiten noch eine weitere Restgliedabschätzung her.

**Theorem 4.6.7.** Sei  $E$  ein reeller Banachraum, sei  $I \subset \mathbb{R}$  ein offenes Intervall mit  $0, 1 \in I$ , sei  $f: I \rightarrow E$  eine  $(n+1)$ -mal differenzierbare Funktion mit  $n \geq 0$ . Dann gilt

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

mit

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

*Beweis.* Wir definieren

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

$\varphi$  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 4.2.3 (oder Theorem 4.2.1), auf dem Intervall  $(0, 1)$  an und erhalten

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

wie behauptet.  $\square$

**Theorem 4.6.8** (Lagrangesche Restgliedabschätzung). *Sei  $\Omega \subset \mathbb{K}$  offen und konvex und seien  $x, x_0 \in \Omega$ . Sei  $E$  ein Banachraum über  $\mathbb{K}$  und  $f: \Omega \rightarrow E$  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

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

*Beweis.* Folge dem Beweis von Theorem 4.6.3 und benutze dabei Theorem 4.6.7 statt Theorem 4.6.2.  $\square$

Ohne Abzuschätzen erhält man

**Theorem 4.6.9** (Lagrangesches Restglied).  $\star$  *Sei  $\Omega \subset \mathbb{R}$  ein Intervall und seien  $x, x_0 \in \Omega$ . Sei  $f: \Omega \rightarrow \mathbb{R}$  eine  $(n+1)$ -mal differenzierbare Funktion mit  $n \geq 0$ . Dann gilt*

$$f(x) = T_n(f, x_0)(x) + \frac{1}{(n+1)!} f^{(n+1)}(\xi) \cdot (x - x_0)^{n+1}$$

für ein  $\xi \in [x, x_0]$ .

*Beweis.* Wir gehen wie im Beweis von Theorem 4.6.7 vor und erhalten in der dortigen Notation aber mit  $F$  statt  $f$

$$\begin{aligned} F(1) - \sum_{k=0}^n \frac{1}{k!} F^{(k)}(0) &= \varphi(1) - \varphi(0) \\ &= -\frac{\varphi(1) - \varphi(0)}{g(1) - g(0)} \\ &= -\frac{\varphi'(\xi)}{g'(\xi)} \end{aligned}$$

für ein  $\xi \in [x, x_0]$

$$= \frac{1}{(n+1)!} F^{(n+1)}(\xi).$$

Nun wenden wir dies für  $F(t) := f(tx + (1-t)x_0)$  an und erhalten aufgrund der Kettenregel die Behauptung.  $\square$

Weitere Restglieddarstellungen benutzen Integrale.

**Lemma 4.6.10.**  $\star$  Sei  $F$  ein Banachraum. Seien  $I = (a, b)$  und  $x, y \in I$ . Sei  $f \in C^n(I, F)$  mit  $n \in \mathbb{N}_{\geq 1}$ . Dann gilt

$$\begin{aligned} f(x) &= f(y) + \frac{x-y}{1!} f'(y) + \frac{(x-y)^2}{2!} f''(y) + \dots + \frac{(x-y)^{n-1}}{(n-1)!} f^{(n-1)}(y) \\ &\quad + \int_y^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt \\ &= \sum_{i=0}^{n-1} \frac{(x-y)^i}{i!} f^{(i)}(y) + \int_y^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt \end{aligned}$$

*Beweis.* Der Induktionsanfang „ $n = 1$ “ lautet

$$f(x) = f(y) + \int_y^x f'(t) dt$$

und ist gerade der Hauptsatz der Differential- und Integralrechnung.

Induktionsschritt: Gelte also die Behauptung bereits für ein festes  $n \in \mathbb{N}_{\geq 1}$ . Dann liefert die folgende Rechnung mit partieller Integration, dass sie auch für  $n+1$  gilt.

$$\begin{aligned} &\frac{(x-y)^n}{n!} f^{(n)}(y) + \int_y^x \underbrace{\frac{(x-t)^n}{n!}}_{=G(t)} \underbrace{f^{(n+1)}(t)}_{=F'(t)} dt \\ &= \frac{(x-y)^n}{n!} f^{(n)}(y) + \left\{ \frac{(x-t)^n}{n!} f^{(n)}(t) \right\} \Big|_{t=y}^{t=x} - \int_y^x -\frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt \\ &= \int_y^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt. \end{aligned}$$

Dies liefert gerade die Behauptung.  $\square$

Als direkte Verallgemeinerung daraus erhalten wir das folgende Korollar. Dies benutzt jedoch auf Banachräumen definierte Funktionen. Wir empfehlen daher, dies erst nach der mehrdimensionalen Differentialrechnung anzusehen.

**Korollar 4.6.11.**  $\star$  Seien  $E, F$  Banachräume. Sei  $\Omega \subset E$  offen. Seien  $x, x_0 \in \Omega$  und gelte  $[x, x_0] \subset \Omega$ . Sei  $f \in C^n(\Omega, F)$  mit  $n \in \mathbb{N}_{\geq 1}$ . Für  $v \in E$  und  $k \in \mathbb{N}$  schreiben wir  $\langle v \rangle^{(k)}$  für  $\langle v, \dots, v \rangle$ , wobei  $v$  in den spitzen Klammern genau  $k$ -mal

vorkommt. Dann gilt

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{1!} f'(x_0) \langle x - x_0 \rangle + \frac{1}{2!} f''(x_0) \langle x - x_0 \rangle^{(2)} + \dots \\ &\quad + \frac{1}{(n-1)!} f^{(n-1)}(x_0) \langle x - x_0 \rangle^{(n-1)} \\ &\quad + \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!} f^{(n)}(tx + (1-t)x_0) dt \langle x - x_0 \rangle^{(n)}. \end{aligned}$$

*Beweis.* Wende Lemma 4.6.10 auf die Funktion  $\varphi: (-\varepsilon, 1 + \varepsilon) \rightarrow F$  mit  $\varphi(t) := f(tx + (1-t)x_0)$  an.  $\square$

## 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  $f: [a, b] \rightarrow \mathbb{R}$  reell, nicht-negativ und beschränkt. Dann wollen wir mit dem Integral über die Funktion  $f$  den Flächeninhalt der Menge

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

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

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

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**.

★ Wir werden auch laxerweise  $P$  selbst als Partition bezeichnen.

- (ii) Sei  $P = \{x_0, x_1, \dots, x_n\}$  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 (unter Verwendung der laxen Bezeichnung)  $P \subset P'$  gilt.
- (iv) Seien  $P_1$  und  $P_2$  zwei Partitionen von  $I$ . Dann heißt die Partition  $P_1 \vee P_2 := P_1 \cup P_2$  (laxe Bezeichnung) die gemeinsame **Verfeinerung** von  $P_1$  und  $P_2$ .

**Lemma 5.1.2.** Seien  $P_1$  und  $P_2$  Partitionen von  $I = [a, b]$ . Dann gilt

$$P_1 \prec P_2 \implies \eta(P_2) \leq \eta(P_1).$$

*Beweis.* Klar.  $\square$

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  $Q$ :  $|Q|$ . Wir wollen  $|Q|$  definieren. Dass diese Rechtecke auf endlich vielen Linien überlappen trägt nichts zum Flächeninhalt bei. Wir fordern also

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

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  $\eta(P_k) \rightarrow 0$  für  $k \rightarrow \infty$  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 5.1.3.** Sei  $I = [a, b] \subset \mathbb{R}$  kompakt, sei  $E$  ein Banachraum und sei  $f: I \rightarrow E$  beschränkt. Sei  $P = \{x_0, x_1, \dots, x_n\}$  eine Partition von  $I$  und seien  $I_i$  sowie  $f_i$  wie oben.

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

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

(Für reellwertige Funktionen gilt  $\text{osc } f = \sup_{x \in I} f(x) - \inf_{y \in I} f(y)$ .)

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

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

(iii) Seien  $\xi_i \in I_i$  beliebig und  $Z := \{\xi_i: 0 \leq i \leq n-1\}$ . Die Punkte  $\xi_i$ ,  $0 \leq i \leq n-1$ , 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 5.1.4.** Sei  $f: I \rightarrow E$  beschränkt, seien  $P$  und  $P'$  zwei Partitionen von  $I$  mit  $P \prec P'$  und habe  $P'$  genau  $\lambda$  Teilpunkte mehr als  $P$ . Dann gelten

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

*Beweis.*

- (i) Seien  $I_i$ ,  $0 \leq i \leq n-1$  die Teilintervalle der Partition  $P$  und  $I'_j$ ,  $0 \leq j \leq m-1$  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  $f'_j = f|_{I'_j}$

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

- (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

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

**Lemma 5.1.5.** Seien  $I = [a, b]$  und  $E$  wie bisher. Sei  $f: I \rightarrow E$  beschränkt. Seien  $P$  und  $P'$  zwei Partitionen des Intervalles  $I$  mit Zwischenpunkten  $Z$  bzw.  $Z'$ .

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

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

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

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

*Beweis.*

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

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

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

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

- (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

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

□

**Lemma 5.1.6** (Riemannsches Integrabilitätskriterium). *Sei  $f: I = [a, b] \rightarrow E$  beschränkt. Dann sind die folgenden beiden Aussagen äquivalent:*

- (i)  $\forall \varepsilon > 0 \exists P \sigma(f, P) < \varepsilon$  und  
(ii)  $\forall \varepsilon > 0 \exists \delta > 0 \forall P \eta(P) < \delta \implies \sigma(f, P) < \varepsilon$ .

*Beweis.*

- „(i) $\implies$ (ii)“: Sei  $\varepsilon > 0$  beliebig und  $P$  eine Partition mit  $\sigma(f, P) < \varepsilon$ . Wir nehmen an, dass  $P$  gerade  $\lambda$  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  $\lambda$  zusätzliche Teilpunkte gegenüber  $P'$ . Es folgt

$$\begin{aligned} \sigma(f, P') & \leq \sigma(f, P^*) + |\sigma(f, P') - \sigma(f, P^*)| \\ & \leq \sigma(f, P) + 2\lambda\|f\|_{L^\infty} \cdot \eta(P') \quad (\text{nach Lemma 5.1.4}) \\ & \leq \varepsilon + 2\lambda\|f\|_{L^\infty} \cdot \delta = \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

- „(ii) $\implies$ (i)“: Ist klar. □

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

**Definition 5.1.7** (Riemannsches Integral). Sei  $I = [a, b] \subset \mathbb{R}$  kompakt und sei  $E$  ein Banachraum.

- (i) Eine beschränkte Funktion  $f: I \rightarrow E$  heißt **Riemann integabel** oder **integabel**, falls  $f$  eine der Bedingungen aus Lemma 5.1.6 erfüllt.

Wir sagen dann, dass  $f$  das **Riemannsche Integrabilitätskriterium** erfüllt.

- (ii) Sei  $f$  Riemann integabel und  $(P_k)_k$  eine Folge von Partitionen von  $I$  mit Zwischenpunkten  $(Z_k)_k$  und  $\eta(P_k) \rightarrow 0$  für  $k \rightarrow \infty$ . Dann definieren wir das **(bestimmte) Riemannsche Integral** von  $f$  über  $I$  durch

$$\int_a^b f \equiv \int_a^b f(x) dx \equiv \int_I f := \lim_{k \rightarrow \infty} S(P_k, Z_k).$$

(Der Zusatz „bestimmt“ deutet dabei an, dass wir das Integrationsintervall  $I$  festgelegt haben.)

**Theorem 5.1.8.** *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 5.1.5 folgt dann

$$(5.1) \quad \|S(P, Z) - S(P', Z')\| \leq \sigma(f, P) + \sigma(f, P') < 2\varepsilon.$$

- (ii) Sei  $(P_k)_k$  eine beliebige Folge von Partitionen von  $I$  mit  $\eta(P_k) \rightarrow 0$  für  $k \rightarrow \infty$  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  $\lim_{k \rightarrow \infty} S(P_k, Z_k)$  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

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

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.  $\square$

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

**Korollar 5.1.9.** *Für beliebiges  $\varepsilon > 0$  gibt es ein  $\delta > 0$ , so dass für alle Partitionen  $P$  mit  $\eta(P) < \delta$  und Zwischenpunkten  $Z$  die Ungleichung*

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

*gilt.*

## 5.2. Integrationsregeln.

**Proposition 5.2.1.** *Sei  $f: I = [a, b] \rightarrow E$  Riemann integabel und sei  $[\alpha, \beta] \subset [a, b]$ . Dann ist  $f|_{[\alpha, \beta]}: [\alpha, \beta] \rightarrow E$  ebenfalls Riemann integabel.*

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

**Proposition 5.2.2.** *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.*  $\star$

- (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 Integabilitätskriterium 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 5.2.3.** ★ 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 5.2.2 erhalten wir daraus

**Proposition 5.2.4.** ★ Sei  $f: I = [a, b] \rightarrow E$  integrierbar. Seien  $a_i \in I$ ,  $1 \leq i \leq n$  beliebig. Dann folgt

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

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

**Theorem 5.2.5** (Dreiecksungleichung). Sei  $f: [a, b] \rightarrow E$  (Riemann) integrierbar. Dann ist auch  $\|f\|: [a, b] \rightarrow \mathbb{R}$  (Riemann) integrierbar und es gilt

$$\left\| \int_a^b f \right\| \leq \int_a^b \|f\|.$$

*Beweis.*

- (i) Nach Dreiecksungleichung gilt für alle  $x, y \in [a, b]$

$$\|f(x) - f(y)\| \leq \|f(x) - f(y)\|$$

Somit erhalten wir  $\text{osc } \|f\| \leq \text{osc } f_i$  für alle Intervalle  $I_i \subset [a, b]$ . Daher gilt auch  $\sigma(\|f\|, P) \leq \sigma(f, P)$  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

$$\left\| \int_a^b f - S(f, P, Z) \right\| < \varepsilon \quad \text{und} \quad \left| \int_a^b \|f\| - S(\|f\|, P, Z) \right| < \varepsilon$$

gelten.

Wir verwenden nun die üblichen Bezeichnungen. Aus

$$S(\|f\|, P, Z) = |S(f, P, Z)|$$

folgt im Limes  $\left| \int_a^b \|f\| \right| = \int_a^b \|f\|$ . Somit erhalten wir

$$\begin{aligned} \left\| \int_a^b f \right\| &\leq \left\| \int_a^b f - S(f, P, Z) \right\| + |S(f, P, Z)| \\ &\leq \varepsilon + \left\| \sum_{i=0}^{n-1} f(\xi_i) \cdot |I_i| \right\| \\ &\leq \varepsilon + \sum_{i=0}^{n-1} \|f(\xi_i)\| \cdot |I_i| \\ &= \varepsilon + S(\|f\|, P, Z) = \varepsilon + |S(f, P, Z)| \end{aligned}$$

$$\begin{aligned}
&\leq \varepsilon + \left| S(\|f\|, P, Z) - \int_a^b \|f\| \right| + \left| \int_a^b \|f\| \right| \\
&\leq 2\varepsilon + \int_a^b \|f\|.
\end{aligned}$$

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

**Proposition 5.2.6** (Monotonie des Integrals). *Seien  $f, g: [a, b] \rightarrow \mathbb{R}$  Riemann integrabel. Gilt  $f \leq g$ , so folgt*

$$\int_a^b f \leq \int_a^b g.$$

*Insbesondere gilt*

$$-(b-a) \cdot \|f\|_{L^\infty} \leq \int_a^b f \leq (b-a) \cdot \|f\|_{L^\infty}.$$

*Beweis.*

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

Der erste Teil des folgenden Korollars wird auch als Mittelwertsatz bezeichnet. (Dabei verwenden wir schon das erst in Proposition 5.2.12 gezeigte Resultat, dass das Produkt von Riemann integrablen Funktionen wieder Riemann integrabel ist.)

**Korollar 5.2.7** (Mittelwertsatz). *Sei  $I = [a, b]$  und seien  $f, g: I \rightarrow \mathbb{R}$  Riemann integrabel. Gelten  $0 \leq g$  und  $m \leq f \leq M$ , so folgen*

$$m \int_a^b g \leq \int_a^b fg \leq M \int_a^b g \quad \text{und} \quad m(b-a) \leq \int_a^b f \leq M(b-a).$$

**Proposition 5.2.8.** *Sei  $f: I = [a, b] \rightarrow \mathbb{R}^n$  mit  $f = (f^i)_{1 \leq i \leq n}$  gegeben. Dann ist  $f$  genau dann Riemann integrabel, wenn dies für die Komponentenfunktionen zutrifft und es gilt*

$$\int_a^b f = \left( \int_a^b f^i \right)_{1 \leq i \leq n}.$$

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

*Beweis.* ★

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

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

für alle  $x, y \in I$  folgt

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

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  $S(f^i, P, Z)$  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 5.2.9.** Sei  $E$  ein Banachraum über dem Körper  $\mathbb{K}$ . Dann bilden die Riemann integrierbaren Funktionen  $f: I = [a, b] \rightarrow E$  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  $\lambda \in \mathbb{K}$

$$\int_a^b (f + g) = \int_a^b f + \int_a^b g \quad \text{und} \quad \int_a^b \lambda f = \lambda \int_a^b f.$$

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

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

$$\text{osc}(f|_J + g|_J) \leq \text{osc}(f|_J) + \text{osc}(g|_J)$$

erhalten wir

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

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

Der Rest des Beweises ist einfach: Übung.  $\square$

**Proposition 5.2.10.** Seien  $E, F$  Banachräume. Sei  $I = [a, b]$  und sei  $f \in R(I, E)$ . Sei  $\varphi: E \rightarrow F$  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  $\|\varphi(x) - \varphi(x_0)\| \leq L \cdot \|x - x_0\| \leq Lr$ , wobei  $L$  die Lipschitzkonstante von  $\varphi$  ist. 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  $\varphi$  auf  $R(f)$ . Dann folgt  $\|\varphi(x) - \varphi(y)\| \leq L \cdot \|x - y\|$  für alle  $x, y \in R(f)$ . Sei  $J \subset I$  beliebig. Wir erhalten dann  $\text{osc}(\varphi \circ f|_J) \leq L \cdot \text{osc}(f|_J)$  und folgern, dass  $\sigma(\varphi \circ f, P) \leq L \cdot \sigma(f, P)$  für alle Partitionen  $P$  von  $I$  gilt. Wir erhalten die Behauptung.  $\square$

**Korollar 5.2.11.** 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 5.2.12.** Sei  $I = [a, b]$  und seien  $f, g \in R(I, \mathbb{K})$ , so folgen

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

*Beweis.*

- (i) Die Beschränktheit 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

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

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

- (iii) Gehe analog vor. Übung.  $\square$

**Proposition 5.2.13.**  $\star$  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

$$\int_a^b f = \int_a^b g.$$

*Beweis.* Wähle in den approximierenden Riemannschen Summen die Zwischenpunkte aus der dichten Teilmenge, in der  $f$  und  $g$  übereinstimmen.  $\square$

*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 5.2.14.** 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  $f: I \rightarrow E$  konvergieren. Dann gilt auch  $f \in R(I, E)$  und für das Integral erhalten wir

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

*Beweis.*

- (i) Beschränktheit: Aus

$$\|f(x)\| \leq \|f_n(x)\| + \|f(x) - f_n(x)\| \leq \|f_n\|_{L^\infty} + \|f - f_n\|_{L^\infty}$$

folgt die Beschränktheit von  $f$ .

- (ii) Integrabilität:  $f_n \rightrightarrows f$  ist äquivalent zu  $\|f - f_n\|_{L^\infty} \rightarrow 0$  für  $n \rightarrow \infty$ . Sei  $A \subset I$  beliebig und seien  $x, y \in A$ . Sei  $n \in \mathbb{N}$  fest. Dann erhalten wir

$$\begin{aligned} \|f(x) - f(y)\| &\leq \|f(x) - f_n(x)\| + \|f_n(x) - f_n(y)\| + \|f_n(y) - f(y)\| \\ &\leq \|f - f_n\|_{L^\infty} + \text{osc}(f_n|_A) + \|f_n - f\|_{L^\infty} \end{aligned}$$

und somit

$$\text{osc}(f|_A) \leq \text{osc}(f_n|_A) + 2\|f - f_n\|_{L^\infty}$$

sowie für eine beliebige Partition  $P$  von  $I$

$$\sigma(f, P) \leq \sigma(f_n, P) + 2(b-a)\|f - f_n\|_{L^\infty}.$$

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 integrierbar.

- (iii) Wert des Integrals: Mit Dreiecksungleichung und der Monotonie des Integrals erhalten wir

$$\left\| \int_a^b f - \int_a^b f_n \right\|_E \leq \int_a^b \|(f - f_n)(x)\|_E dx \leq \int_a^b \|f - f_n\|_{L^\infty} \leq (b-a) \cdot \|f - f_n\|_{L^\infty}.$$

Dies liefert die Behauptung.  $\square$

**Bemerkung 5.2.15.** ★ Der punktweise (sogar monotone) Grenzwert von Riemann integrierbaren Funktionen ist i. a. nicht Riemann integrierbar: Sei dazu  $(x_i)_{i \in \mathbb{N}}$  eine Abzählung von  $\mathbb{Q} \cap [0, 1]$ . Wir definieren

$$f_n(x) := \begin{cases} 1, & \text{falls } x = x_i \text{ für ein } i \leq n, \\ 0, & \text{sonst.} \end{cases}$$

Es gilt  $f_n(x) \in R([0, 1], \mathbb{R})$  für alle  $n \in \mathbb{N}$  und  $f_n \rightarrow \chi_{\mathbb{Q} \cap [0, 1]}$  für  $n \rightarrow \infty$ , aber  $\chi_{\mathbb{Q} \cap [0, 1]} \notin R([0, 1], \mathbb{R})$ .

**Theorem 5.2.16** (Integration von Reihen). ★ Sei  $I = [a, b]$ , sei  $(f_n)_n$  eine Folge mit  $f_n \in R(I, E)$ ,  $n \in \mathbb{N}$ . Konvergiere die Reihe  $((f_n))_n$  in  $I$  gleichmäßig gegen  $f$ , gelte also insbesondere  $f(x) = \sum_{n=0}^{\infty} f_n(x)$ . Dann gelten  $f \in R(I, E)$  und

$$\int_a^b f = \sum_{n=0}^{\infty} \int_a^b f_n.$$

*Beweis.* Sei  $(s_n)_n$  die  $n$ -te Partialsumme der Folge. Da  $R(I, E)$  ein Vektorraum ist, folgt  $s_n \in R(I, E)$ . Die Behauptung erhalten wir nun direkt aus Theorem 5.2.14.  $\square$

**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 5.3.1.** Sei  $I = [a, b]$ . Sei  $f: I \rightarrow E$  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

$$|x - y| < \delta \implies \|f(x) - f(y)\| < \varepsilon$$

für alle  $x, y \in I$  gilt. Insbesondere gilt für alle Teilmengen  $J \subset I$  mit  $\text{diam } J = \sup_{x, y \in J} |x - y| < \delta$  die Abschätzung  $\text{osc}(f|_J) < \varepsilon$ .

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  $f \in R(I, E)$ .  $\square$

**Proposition 5.3.2.** Sei  $f: I = [a, b] \rightarrow \mathbb{R}$  monoton, so ist  $f$  Riemann integrierbar.

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

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

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  $0 \leq i \leq n-1$  charakterisiert ist. Für die Partition  $P_n$  bekommen wir eine Teleskopsumme und erhalten

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

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

**Definition 5.3.3.**

- (i) Sei  $A \subset X$  eine Menge. Dann ist  $\chi_A: X \rightarrow \mathbb{R}$ , die **charakteristische Funktion** der Menge  $A$  durch

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

definiert.

- (ii) Eine Funktion  $f: I = [a, b] \rightarrow E$  heißt **Treppenfunktion**, falls es eine Partition  $P = \{x_0, x_1, \dots, x_n\}$  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  $a_i \in E$ ,  $0 \leq i \leq n-1$ , (außer in  $x_n$ ). Wir sprechen auch von einer Treppenfunktion, wenn die Werte in den Teilpunkten  $x_i$  nicht zu dieser Definition passen.

- (iii) Eine Funktion  $f: I = [a, b] \rightarrow E$  heißt **stückweise stetig** bzw. **stückweise von der Klasse  $C^m$** ,  $m \geq 0$ , falls eine Partition  $P$  von  $I$  existiert, so dass  $f$  bzw. alle Ableitungen von  $f$  bis zur Ordnung  $m$  auf allen Intervallen  $\overset{\circ}{I}_i = (x_i, x_{i+1})$ ,  $0 \leq i \leq n-1$ , stetig sind und sich stetig auf jedes  $\overline{I}_i$  fortsetzen lassen.
- (iv)  $\star$  Stückweise affin lineare Funktionen, stückweise quadratische Funktionen sowie stückweise polynomiale Funktionen und stückweise konstante Funktionen (= Treppenfunktionen) sind analog definiert.

**Proposition 5.3.4.** *Stückweise stetige Funktionen und daher auch Treppenfunktionen sind Riemann integrierbar.*

*Beweis.*  $\star$  Benutze, dass die Einschränkungen auf die Intervalle  $I_i$  zu einer geeigneten Partition stetig und daher auch Riemann integrierbar sind. Nach Proposition 5.2.2 sind die Funktionen damit auch auf dem gesamten Definitionsbereich Riemann integrierbar.  $\square$

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

**Proposition 5.4.1** (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 5.2.7 erhalten wir  $(b-a) \inf f \leq \int f \leq (b-a) \sup f$ . Jede Zahl in diesem Bereich lässt sich aber aufgrund des Zwischenwertsatzes für stetige Funktionen in der Form  $f(\xi) \cdot (b-a)$  schreiben.  $\square$

Für vektorwertige Funktionen erhalten wir ebenfalls eine Variante des Mittelwertsatzes.

**Proposition 5.4.2** (Mittelwertsatz). *Sei  $I = [a, b]$  und sei  $f \in R(I, E)$ , so gilt*

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

*Beweis.* Nach Dreiecksungleichung gilt

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

wie behauptet.  $\square$

**Proposition 5.4.3.** Sei  $I = [a, b]$  und sei  $f \in R(I, E)$ . Dann definieren wir eine Abbildung

$$F: I \rightarrow E \quad \text{durch} \quad x \mapsto \int_a^x f.$$

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

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

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

wie behauptet.  $\square$

**Theorem 5.4.4** (Hauptsatz der Differential- und Integralrechnung). Sei  $I = [a, b]$  und sei  $f \in C^0(I, E)$ . Dann ist die Funktion  $F$  aus Proposition 5.4.3 in  $(a, b)$  (stetig) differenzierbar und es gilt dort  $F' = f$ .

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

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

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

Wenn wir nun durch  $|x - x_0|$  dividieren, konvergiert die rechte Seite für  $x \rightarrow x_0$  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  $f$ .  $\square$

**Korollar 5.4.5.** Sei  $\Omega \subset \mathbb{R}$  ein offenes Intervall. Sei  $f \in C^1(\Omega, E)$ . Dann gilt

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt$$

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  $\Omega$  überein. Somit ist die Differenz nach Korollar 4.1.35 eine Konstante in  $E$ . Für  $x = x_0$  gilt Gleichheit, somit ist die Konstante  $0 \in E$  und die Behauptung folgt.  $\square$

Aus Korollar 5.4.5 erhalten wir das folgende häufig verwendete Resultat.

**Korollar 5.4.6.** Sei  $\Omega \subset \mathbb{R}$  ein offenes Intervall. Sei  $f \in C^1(\Omega, E)$ . Dann gilt

$$f(x) - f(x_0) = \int_0^1 \frac{d}{dt} f(tx + (1-t)x_0) dt = \int_0^1 Df(tx + (1-t)x_0) \langle x - x_0 \rangle dt$$

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 5.4.5.  $\square$

**Definition 5.4.7.** Sei  $\Omega \subset \mathbb{R}$  ein offenes Intervall. Sei  $f: \Omega \rightarrow E$  stetig. Eine Funktion  $F: \Omega \rightarrow E$  mit  $F' = f$  heißt **Stammfunktion** von  $f$ .

**Bemerkung 5.4.8.**

- (i) Ist  $f$  stetig, so definiert  $F(x) = \int_{x_0}^x f$  eine Stammfunktion von  $f$ .
- (ii)  $\star$  Ist  $F$  eine Stammfunktion von  $f$  und  $c \in E$ , so ist auch  $x \mapsto F(x) + c$  eine Stammfunktion von  $f$ .
- (iii)  $\star$  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 4.1.35.  $\square$

- (iv) Sei  $F$  eine Stammfunktion von  $f$ . Dann folgt nach Korollar 5.4.5 für beliebige  $x_1, x_2 \in (a, b)$

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

- (v)  $\star$  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 5.4.9.** 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  $\text{im } \varphi \subset (a, b)$ . Dann ist die Funktion

$$G(x) := \int_a^{\varphi(x)} f(t) dt$$

in  $\Omega$  (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.  $\square$

### 5.5. Integralsätze und Transformationsregeln.

Die partielle Integration ist eine integrierte Version der Produktregel.

**Theorem 5.5.1** (Partielle Integration). Sei  $I = [a, b]$  und seien  $f, g \in C^1(I, \mathbb{K})$ . Dann gilt

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

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  $\int_{x_0}^x$  und benutzen  $\int_{x_0}^x (fg)' = (fg)|_{x_0}^x$ . Somit folgt die Behauptung in diesem Fall. Durch Grenzübergang erhalten wir aufgrund der Stetigkeit aller Summanden die Behauptung auch für beliebige  $x, x_0 \in I$ .  $\square$

**Beispiele 5.5.2.** ★

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

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

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

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

**Theorem 5.5.3** (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

$$\int_a^b f(x) \, dx = \int_\alpha^\beta f \circ \varphi(t) \varphi'(t) \, dt.$$

Die Transformationsformel ist letztlich eine integrierte Form der Kettenregel.

*Beweis.* Wir betrachten beide Seiten als Funktion der oberen Grenze  $\beta$ , also insbesondere auch  $b = \varphi(\beta)$ , benutzen  $\gamma$  statt  $\beta$  und  $c = \varphi(\gamma)$  statt  $b$ , da die obere Grenze nun variabel sein soll und erhalten mit Proposition 5.4.9

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

und andererseits

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

Daher stimmen die Ableitungen beider Seiten überein.

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

**Beispiel 5.5.4.** ★ 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

$$\int_0^a e^{-x^2} x \, dx = \frac{1}{2} \int_0^a f(\varphi(x)) \varphi'(x) \, dx = \frac{1}{2} \int_{\varphi(0)}^{\varphi(a)} e^{-t} \, dt$$

$$= \frac{1}{2} \int_0^{a^2} e^{-t} dt = -\frac{1}{2} e^{-t} \Big|_0^{a^2} = \frac{1}{2} (1 - e^{-a^2}).$$

**5.6. Integration rationaler Funktionen \*** Wir wollen in diesem Abschnitt lernen, wie man eine echt rationale reelle Funktion  $h = f/g$ , die als komplexe Partialbruchzerlegung durch

$$h(z) = \sum_{i=1}^m \sum_{j=1}^{m_i} \frac{c_{ij}}{(z - \zeta_i)^j}$$

gegeben ist, integriert.

Zur Vorbereitung zeigen wir

**Lemma 5.6.1.** *Sei  $f \in P(\mathbb{C})$  ein Polynom mit reellen Koeffizienten. Sei  $\zeta$  eine Nullstelle von  $f$ . Dann ist auch  $\bar{\zeta}$  eine Nullstelle von  $f$  und die Vielfachheiten der Nullstelle von  $\zeta$  und  $\bar{\zeta}$  stimmen überein.*

*Beweis.*

- (i) Da  $f$  reelle Koeffizienten hat, folgt  $\overline{f(z)} = f(\bar{z})$  für alle  $z \in \mathbb{C}$ . Insbesondere folgt aus  $f(\zeta) = 0$  also auch  $f(\bar{\zeta}) = 0$ .
- (ii) Sei  $\zeta$  eine nicht reelle Nullstelle von  $f$ . Wir können  $f$  mittels Polynomdivision als  $f(z) = (z - \zeta)(z - \bar{\zeta})g(z)$  mit einem Polynom  $g$  schreiben. Da  $(z - \zeta)(z - \bar{\zeta}) = z^2 - (\zeta + \bar{\zeta})z - \zeta\bar{\zeta}$  ein Polynom mit reellen Koeffizienten ist, ist auch  $g$  ein Polynom mit reellen Koeffizienten. Somit folgt die Aussage über die Vielfachheit der Nullstellen per Induktion.  $\square$

Da die Nenner in der Partialbruchzerlegung gerade von den Nullstellen von  $g$  in  $h = f/g$  herkommen, können wir die Partialbruchzerlegung auch in der Form

$$(5.2) \quad h(z) = \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{a_{ij}}{(z - \xi_i)^j} + \sum_{i=1}^{n'} \sum_{j=1}^{m'_i} \left( \frac{b_{ij}}{(z - \zeta_i)^j} + \frac{c_{ij}}{(z - \bar{\zeta}_i)^j} \right)$$

schreiben, wobei  $\xi_i$  die reellen Nullstellen von  $g$  und  $\zeta_i, \bar{\zeta}_i$  die komplexen, nicht reellen Nullstellen von  $g$  sind.

**Lemma 5.6.2.** *Sei  $h = f/g$  eine rationale Funktion mit reellen Koeffizienten. Dann gelten in der Partialbruchzerlegung in der Form (5.2) die Aussagen  $a_{ij} \in \mathbb{R}$  und  $\bar{b}_{ij} = c_{ij}$ .*

*Beweis.* Aus (5.2) und da die Koeffizienten reell sind, erhalten wir

$$h(\bar{z}) = \overline{h(z)} = \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{\bar{a}_{ij}}{(\bar{z} - \xi_i)^j} + \sum_{i=1}^{n'} \sum_{j=1}^{m'_i} \left( \frac{\bar{b}_{ij}}{(\bar{z} - \bar{\zeta}_i)^j} + \frac{\bar{c}_{ij}}{(\bar{z} - \zeta_i)^j} \right).$$

Dies gilt für beliebige  $z \in \mathbb{C}$ , also insbesondere auch für  $z$  statt  $\bar{z}$ . Wir vergleichen das Ergebnis mit (5.2) und erhalten durch Koeffizientenvergleich aufgrund der Eindeutigkeit der Partialbruchzerlegung die Behauptung.  $\square$

Für den rein reellen Anteil mit  $a_{ij}$  in (5.2) können wir eine reelle Stammfunktion mit  $\log|z - \xi_i|$  im Falle  $j = 1$  und als rationale Funktion für  $j > 1$  direkt angeben. Daher genügt es, im Falle

$$f(x) = \frac{b}{(x - \zeta)^m} + \frac{\bar{b}}{(x - \bar{\zeta})^m}, \quad m \geq 1,$$

eine reelle Stammfunktion anzugeben.

Diese kann man mit Fallunterscheidung direkt angeben:

**Bemerkung 5.6.3.**

- (i) Sei  $m \geq 2$ . Wir geben eine Stammfunktion  $F$  in zwei verschiedenen Formen an. An der ersten Form liest man direkt ab, dass es sich um eine Stammfunktion handelt, an der zweiten sieht man, dass die Funktion reell ist.

$$\begin{aligned} F(x) &= -\frac{1}{m-1} \left( \frac{b}{(x-\zeta)^{m-1}} + \frac{\bar{b}}{(x-\bar{\zeta})^{m-1}} \right) \\ &= -\frac{1}{m-1} \frac{b(x-\bar{\zeta})^{m-1} + \bar{b}(x-\zeta)^{m-1}}{(x^2 - x(\zeta + \bar{\zeta}) + |\zeta|^2)^{m-1}}. \end{aligned}$$

- (ii) Ist  $m = 1$ , so können wir die Funktion  $f$  in der Form

$$f(x) = \frac{b(x-\bar{\zeta}) + \bar{b}(x-\zeta)}{x^2 - x(\zeta + \bar{\zeta}) + |\zeta|^2} \equiv \frac{\alpha x + \beta}{x^2 + px + q}$$

mit  $\alpha, \beta, p, q \in \mathbb{R}$  darstellen. Der Nenner besitzt nach Voraussetzung keine reelle Nullstelle und ist daher überall strikt positiv. Wegen

$$x^2 + px + q = \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q \equiv \left(x + \frac{p}{2}\right)^2 - D$$

ist dies nur für  $D < 0$  möglich.

- (iii) Es gilt

$$\frac{d}{dx} \left( \frac{1}{\sqrt{-D}} \arctan \frac{x + \frac{p}{2}}{\sqrt{-D}} \right) = \frac{1}{x^2 + px + q}.$$

Dies rechnet man direkt unter Benutzung von  $\arctan' x = \frac{1}{1+x^2}$  nach.

- (iv) Es gilt

$$\frac{d}{dx} \frac{1}{2} \log(x^2 + px + q) = \frac{x + \frac{p}{2}}{x^2 + px + q}.$$

Auch dies rechnet man direkt nach.

- (v) Aus diesen Ausdrücken lässt sich nun stets eine reelle Stammfunktion zusammensetzen.

**5.7. Lebesguesches Integrabilitätskriterium \***

Wir wollen in diesem Abschnitt zeigen, dass eine beschränkte Funktion genau dann Riemann integrierbar ist, wenn sie außerhalb einer Nullmenge stetig ist.

**Definition 5.7.1** (Nullmenge).

- (i) Eine Menge  $M \subset \mathbb{R}$  heißt **(Lebesguesche) Nullmenge**, falls es zu jedem  $\varepsilon > 0$  eine höchstens abzählbare Überdeckung  $(I_i)_{i \in A}$  von  $M$  durch offene Intervalle mit

$$\sum_{i \in A} |I_i| < \varepsilon$$

gibt.

- (ii) Eine Aussage  $p = p(x)$  gilt **fast überall** (f. ü.) in  $M \subset \mathbb{R}$ , falls die Menge  $\{x \in M : p(x) \text{ gilt nicht}\}$  eine Nullmenge ist.

**Bemerkung 5.7.2.**

- (i) Wir haben nicht gefordert, dass die Intervalle  $I_i$  paarweise disjunkt sind.  
(ii) Jede höchstens abzählbare Teilmenge von  $\mathbb{R}$  ist eine Nullmenge.  
(iii) Jede Teilmenge einer Nullmenge ist selber wieder eine Nullmenge.  
(iv) Die höchstens abzählbare Vereinigung von Nullmengen ist wieder eine Nullmenge. (Benutze ein  $2^{-i}\varepsilon$ -Argument.)

- (v) Eine kompakte Menge  $K \subset \mathbb{R}$  ist genau dann eine Nullmenge, wenn es zu jedem  $\varepsilon > 0$  eine endliche Überdeckung  $(I_i)_{1 \leq i \leq n}$  mit offenen Intervallen und

$$\sum_{i=1}^n |I_i| < \varepsilon$$

gibt.

- (vi) Eine äquivalente Definition einer Nullmenge erhalten wir, wenn wir statt offenen Intervallen in der Überdeckung beliebige Intervalle zulassen.

*Beweis.* Einfache Übung. □

Für dieses Kapitel führen wir eine ad hoc Definition ein.

**Definition 5.7.3.** Seien  $E, F$  metrische Räume. Sei  $f: E \rightarrow F$  eine Abbildung.

- (i) Wir definieren die **Oszillation** von  $f$  durch

$$\operatorname{osc} f := \sup_{x, y \in E} d(f(x), f(y)).$$

- (ii) Ist  $A \subset E$ , so definieren wir die Oszillation von  $f$  in  $A$  durch

$$\operatorname{osc}(f, A) := \operatorname{osc}(f|_A).$$

- (iii) Ist  $x \in E$ , so definieren wir die Oszillation von  $f$  in  $x$  durch

$$\operatorname{osc}(f, x) := \lim_{r \downarrow 0} \operatorname{osc}(f, B_r(x)).$$

(Da  $r \mapsto \operatorname{osc}(f, B_r(x))$  monoton wachsend ist, existiert der Grenzwert, ist aber möglicherweise gleich  $+\infty$ .)

**Bemerkung 5.7.4.** Seien  $E, F$  metrische Räume und  $f: E \rightarrow F$  eine Abbildung.

- (i) Dann ist  $f$  genau dann in  $x \in E$  stetig, wenn  $\operatorname{osc}(f, x) = 0$  gilt.  
(ii) Mit  $\Sigma(f)$  bezeichnen wir die Menge der Unstetigkeitspunkte von  $f$ . Weiterhin definieren wir  $\Sigma_\varepsilon(f) := \{x \in E: \operatorname{osc}(f, x) \geq \varepsilon\}$ . Dann gilt

$$\Sigma(f) = \bigcup_{m=1}^{\infty} \Sigma_{\frac{1}{m}}(f).$$

**Bemerkung 5.7.5.** Sei  $f: E \rightarrow F$  eine Abbildung zwischen metrischen Räumen. Sei  $\varepsilon > 0$ . Dann ist die Menge  $\Sigma_\varepsilon(f)$  abgeschlossen.

*Beweis.* Gelte  $\Sigma_\varepsilon(f) \ni x_i \rightarrow x_0$  für  $i \rightarrow \infty$ . Sei  $r > 0$  beliebig. Es genügt der Nachweis, dass es  $a, b \in B_r(x_0)$  mit  $d(f(a), f(b)) \geq \varepsilon$  gibt. Sei dazu  $x_{i_0} \in B_{r/2}(x_0)$ . Dann gibt es nach Voraussetzung  $a, b \in B_{r/2}(x_{i_0})$  mit  $d(f(a), f(b)) \geq \varepsilon$ . Nach Dreiecksungleichung gilt  $a, b \in B_r(x_0)$  und die Behauptung folgt. □

Das Hauptergebnis dieses Kapitels ist

**Theorem 5.7.6.** Sei  $f: I = [a, b] \rightarrow E$  beschränkt. Dann ist  $f$  genau dann Riemann integrierbar, wenn  $f$  fast überall stetig ist.

*Beweis.*

„ $\Rightarrow$ “: Sei  $f \in R(I, E)$ . Es genügt der Nachweis, dass für jedes  $m \in \mathbb{N}_{>0}$  die Menge  $\Sigma_{\frac{1}{m}}(f)$  eine Nullmenge ist, denn dann ist auch  $\Sigma(f) = \bigcup_{i=1}^{\infty} \Sigma_{\frac{1}{m}}(f)$  eine Nullmenge.

Sei  $\varepsilon > 0$  beliebig und  $m \in \mathbb{N}_{>0}$  fest. Wir wollen zeigen, dass sich die Menge  $\Sigma_{\frac{1}{m}}(f)$  durch Intervalle mit aufsummierter Länge  $< \varepsilon$  überdecken lässt.

Da  $f$  Riemann integrierbar ist, gibt es eine Partition  $P = \{x_0, \dots, x_n\}$  von  $I$ , so dass

$$\sigma(f, P) < \frac{\varepsilon}{2m}$$

gilt. Wir definieren  $\Lambda \subset \{0, 1, \dots, n-1\}$  als die Menge der Indices  $i$ , so dass  $\Sigma_{\frac{1}{m}}(f) \cap \overset{\circ}{I}_i \neq \emptyset$  gilt. Dann erhalten wir

$$\Sigma_{\frac{1}{m}}(f) \subset \bigcup_{i \in \Lambda} I_i \cup \bigcup_{i=0}^n \{x_i\}.$$

Wir müssen nun insbesondere die Gesamtlänge der Intervalle in der ersten Vereinigung abschätzen. Sei dazu  $i \in \Lambda$  und  $\xi \in \Sigma_{\frac{1}{m}}(f) \cap \overset{\circ}{I}_i$ . Dann gibt es eine Kugel/ein Intervall  $B_\delta(\xi)$  mit  $B_\delta(\xi) \subset \overset{\circ}{I}_i$ . Es folgt nach Definition und Monotonie der Oszillation im Definitionsbereich

$$\frac{1}{m} \leq \text{osc}(f, \xi) \leq \text{osc}(f, B_\delta(\xi)) \leq \text{osc}(f, I_i) \equiv \text{osc}(f_i).$$

Wir summieren über alle  $i \in \Lambda$  und erhalten aus

$$\frac{1}{m} \sum_{i \in \Lambda} |I_i| \leq \sum_{i \in \Lambda} \text{osc}(f_i) \cdot |I_i| \leq \sigma(f, P) < \frac{\varepsilon}{2m}$$

die Abschätzung

$$\sum_{i \in \Lambda} |I_i| < \frac{\varepsilon}{2}.$$

Die endlich vielen Zwischenpunkte aus der Partition  $P$  überdecken wir mit Intervallen  $J_i$ ,  $0 \leq i \leq n-1$ , mit  $x_i \in J_i$  der aufsummierten Gesamtlänge  $< \frac{\varepsilon}{2}$ . Insgesamt erhalten wir also

$$\Sigma_{\frac{1}{m}}(f) \subset \bigcup_{i \in \Lambda} I_i \cup \bigcup_{i=0}^n J_i \quad \text{und} \quad \sum_{i \in \Lambda} |I_i| + \sum_{i=0}^n |J_i| < \varepsilon.$$

„ $\Leftarrow$ “: Sei umgekehrt  $\Sigma(f)$  eine Nullmenge. Sei  $\varepsilon > 0$  beliebig und sei  $(J_k)_{k \in \Lambda}$  eine höchstens abzählbare Überdeckung von  $\Sigma(f)$  durch offene Intervalle  $J_k$  mit  $\sum_{k \in \Lambda} |J_k| < \varepsilon$ .

Zu jedem Punkt  $\xi \in I \setminus \Sigma(f)$  gibt es ein  $\delta > 0$ , so dass

$$\text{osc}(f, \overline{B_\delta(\xi) \cap I}) \equiv \text{osc}(f, \overline{B(\xi)}) < \varepsilon$$

gilt. Die Mengen  $(J_k)_{k \in \Lambda}$  und  $(B(\xi))_{\xi \in I \setminus \Sigma(f)}$  bilden zusammen eine offene Überdeckung von  $I$ . Da  $I$  kompakt ist, besitzt diese Überdeckung eine endliche Teilüberdeckung,

$$I \subset \bigcup_{k \in \Lambda'} J_k \cup \bigcup_{j=1}^N B(\xi_j).$$

Wir betrachten nun die Partition  $P$ , deren Teilpunkte alle Endpunkte der obigen Intervalle sowie  $a$  und  $b$  sind. Dann liegt jedes Teilintervall  $I_i$ ,  $0 \leq i \leq n-1$ , in einem  $\bar{J}_k$  mit  $k \in \Lambda'$  oder in einem  $\overline{B(\xi_j) \cap I}$ ,  $1 \leq j \leq N$ . Wir definieren  $\Lambda_1$  und  $\Lambda_2$  als eine Zerlegung von  $\{0, \dots, n-1\}$ , so dass für die Intervalle  $I_i$  mit  $i \in \Lambda_1$  stets  $I_i \subset \bar{J}_k$  für ein  $k \in \Lambda'$  und für  $i \in \Lambda_2$  stets  $I_i \subset \overline{B(\xi_j) \cap I}$  für ein  $1 \leq j \leq N$  gelten. Wir erhalten also die Abschätzung

$$\sigma(f, P) = \sum_{i=0}^{n-1} \text{osc } f_i \cdot |I_i| = \sum_{i \in \Lambda_1} \text{osc } f_i \cdot |I_i| + \sum_{i \in \Lambda_2} \text{osc } f_i \cdot |I_i|$$

$$\leq 2 \cdot \|f\|_{L^\infty} \cdot \varepsilon + \varepsilon \cdot |I| = (2 \cdot \|f\|_{L^\infty} + (b-a)) \cdot \varepsilon.$$

Da  $\varepsilon > 0$  beliebig war, folgt die Riemann-Integrierbarkeit.  $\square$

**Lemma 5.7.7.** *Sei  $I \subset \mathbb{R}$  ein Intervall mit nichtleerem Inneren und sei  $A \subset I$  eine Lebesgue-Nullmenge. Dann liegt  $I \setminus A$  dicht in  $I$ .*

*Beweis.* Falls nicht, so gibt es  $x \in I$  und  $\varepsilon > 0$  mit  $B_\varepsilon(x) \subset I$  und  $B_\varepsilon(x) \cap (I \setminus A) = \emptyset$ . Also folgt  $B_\varepsilon(x) \subset A$ . Dies ist bei einer Nullmenge  $A$  natürlich nicht möglich. Formal führen wir dies wie folgt zu einem Widerspruch. Die Menge  $K := \overline{B_{\varepsilon/2}(x)}$  ist kompakt. In einer Überdeckung von  $A$  durch höchstens abzählbar viele offene Intervalle mit einer aufsummierten Gesamtlänge  $< \delta$  gibt es also endlich viele,  $I_i$ ,  $1 \leq i \leq n$ , die bereits  $K$  überdecken. Wir erhalten

$$\varepsilon = |K| \leq \sum_{i=1}^n |I_i| < \delta,$$

wobei man „ $\leq$ “ in der obigen Ungleichung durch sukzessives Verschieben der Intervalle, so dass sie direkt aneinander anstoßend liegen, zeigt. Da  $A$  eine Nullmenge ist, können wir  $\delta > 0$  beliebig vorgeben. Widerspruch.  $\square$

**Korollar 5.7.8.** *Sei  $f \in R(I, E)$ . Dann gilt*

$$\|f\|_{L^1} \equiv \int_I \|f\| = 0 \quad \Longleftrightarrow \quad f \equiv 0 \text{ fast überall.}$$

*Beweis.*

„ $\Rightarrow$ “: Gelte  $\|f\|_{L^1} = 0$ . Wir definieren  $\Lambda := I \setminus \Sigma(f)$  und behaupten, dass  $f|_\Lambda \equiv 0$  gilt:

Sei dazu  $\xi \in \Lambda$  mit  $f(\xi) \neq 0$ . Dann ist  $f$  in  $\xi$  stetig und es gibt eine (relativ) offene Kugel  $B_\delta(\xi)$  mit  $0 < \frac{1}{2}\|f(\xi)\| \leq \|f(x)\|$  für alle  $x \in B_\delta(\xi)$ . Wir dürfen ohne Einschränkung annehmen, dass zumindest die rechte oder linke Hälfte der Kugel/des Intervalles  $B_\delta(\xi)$  in  $I$  liegt. Sonst müssen wir in der folgenden Abschätzung  $\delta$  durch  $|I|$  ersetzen. Aufgrund der Monotonie des Integrals im Integrationsbereich bei nichtnegativen **Integranden** (:= der Funktion „unter“ dem Integral) und im Integranden erhalten wir

$$\int_a^b \|f\| \geq \int_{B_\delta(\xi) \cap I} \|f\| \geq \frac{1}{2} \|f(\xi)\| \cdot \delta > 0.$$

Widerspruch.

„ $\Leftarrow$ “: Gelte fast überall  $f = 0$ . Wir definieren  $\Lambda := \{x \in I : f(x) \neq 0\}$ . Dann ist  $\Lambda$  nach Voraussetzung eine Lebesgue-Nullmenge. Nach Lemma 5.7.7 liegt  $I \setminus \Lambda$  also dicht in  $I$ . Auf einer dichten Teilmenge von  $I$  gilt also  $f = 0$ . Nach Proposition 5.2.13 folgt daher  $\int_I f = \int_I 0 = 0$  wie behauptet.  $\square$

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

**Definition 5.8.1.** 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 5.8.2.**

- (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  $\text{dist}(\overline{A}, \partial V) > 0$ .

(iii) Wir erinnern daran, dass  $\nearrow$  bzw.  $\searrow$  monotone Konvergenz bezeichnen.

**Definition 5.8.3.** Sei  $(x_n)_n \subset \mathbb{R}$  eine Folge und  $a \in \mathbb{R}$ .

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

**Definition 5.8.4.** Sei  $E$  ein Banachraum und sei  $\emptyset \neq I \subset \mathbb{R}$  ein beliebiges nicht-leeres Intervall.

- (i) Eine Funktion  $f: I \rightarrow E$  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**  $\int_a^b f$  existiert oder konvergiert, falls

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

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

- (iii)  $\star$  Entsprechend definieren wir im Falle  $I = (a, b]$  und  $f \in R_{\text{loc}}(I, E)$ , falls der folgende Grenzwert existiert,

$$\int_a^b f := \lim_{\delta \downarrow 0} \int_{a+\delta}^b f.$$

- (iv)  $\star$  Ist  $I = (a, b)$  und  $f \in R_{\text{loc}}(I, E)$ , so definieren wir, sofern diese Grenzwerte existieren,

$$\lim_{\delta \downarrow 0} \int_c^{b-\delta} f + \lim_{\varepsilon \downarrow 0} \int_{a+\varepsilon}^c f$$

für ein beliebiges  $c \in (a, b)$ .

(Diese Definition hängt nicht von der speziellen Wahl von  $c \in (a, b)$  ab.)

**Beispiel 5.8.5.**  $\star$  Es gilt

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

Ganz analog definieren wir für unbeschränkte Integrationsbereiche

**Definition 5.8.6.**  $\star$

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

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

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

- (ii) Falls  $f \in R_{\text{loc}}((-\infty, a], E)$  ist, so definieren wir entsprechend

$$\int_{-\infty}^a f = \lim_{x \rightarrow -\infty} \int_x^a f,$$

falls dieser Grenzwert existiert.

- (iii) Ist  $f \in R_{\text{loc}}(\mathbb{R}, E)$ , so definieren wir

$$\int_{-\infty}^{\infty} f = \lim_{x \rightarrow -\infty} \int_x^a f + \lim_{y \rightarrow \infty} \int_a^y f$$

für ein beliebiges  $a \in \mathbb{R}$ , falls diese Grenzwerte beide existieren.

(Diese Definition ist wiederum von der Wahl von  $a$  unabhängig.)

**Beispiel 5.8.7.** Es gilt

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

**Bemerkung 5.8.8.** ★

- (i) Für  $f \in R([a, b], E)$  stimmt das uneigentliche Integral über  $(a, b)$  mit dem gewöhnlichen Integral über  $[a, b]$  überein, da das Integral für beschränkte Funktionen bezüglich seiner Grenzen stetig ist.
- (ii) Das uneigentliche Integral ist linear in  $f$ , additiv bezüglich seiner Grenzen und im Falle  $E = \mathbb{R}$  auch monoton.

**Lemma 5.8.9.** ★ Ist  $f \in R_{\text{loc}}((a, b), \mathbb{R})$  mit  $f \geq 0$  in  $(a, b)$ ,  $a, b \in \mathbb{R}$  bzw.  $(a, b) = (-\infty, \infty)$ , so gelten

$$\int_a^b f = \lim_{\delta \downarrow 0} \int_{a+\delta}^{b-\delta} f \quad \text{bzw.} \quad \int_{-\infty}^{\infty} f = \lim_{x \rightarrow \infty} \int_{-x}^x f,$$

wobei wir auch den Wert  $+\infty$  zulassen.

*Beweis.* Übung. □

**Bemerkung 5.8.10.** ★ Ist

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

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

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

**Definition 5.8.11.**

- (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 Cauchyriterien erfüllt sind. Im Fall  $\int_a^\infty f$  geben wir es hier an, in den anderen Fällen lauten sie analog.

**Proposition 5.8.12** (Cauchyriterium). *Das uneigentliche Integral  $\int_a^\infty f$  existiert genau dann, wenn*

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

*gilt.*

*Beweis.* Dies folgt direkt nach Definition. □

Damit erhalten wir direkt

**Korollar 5.8.13.** *Ein absolut konvergentes Integral ist konvergent.*

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

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

□

**Proposition 5.8.14** (Majorantenriterium).

*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  $\int_a^\infty g$ , 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 5.8.15.** Ist das Integrationsintervall nichtkompakt, so können unerwartete Effekte auftreten. Auch bei gleichmäßiger Konvergenz gilt dann im Allgemeinen nicht mehr

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

Dies sehen wir an dem folgenden Beispiel. Definiere für  $n \in \mathbb{N}$  Funktionen  $f_n: \mathbb{R}_+ \rightarrow \mathbb{R}$  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  $\int_0^\infty f_n = 1 \not\rightarrow 0$  für  $n \rightarrow \infty$ .

Ist der Grenzwert  $\lim_{x \rightarrow \infty} \int_a^x f_n$  in der Definition des uneigentlichen Integrales jedoch gleichmäßig in  $n$ , so erhalten wir

**Theorem 5.8.16.** Sei  $I = [a, \infty)$  und sei  $f_n \in R_{\text{loc}}(I, E)$  eine Folge von Funktionen, die lokal gleichmäßig (und daher auch kompakt, d. h. gleichmäßig auf kompakten Teilmengen) gegen eine Funktion  $f: I \rightarrow E$  konvergieren. Konvergieren die uneigentlichen Integrale  $\int_a^\infty f_n$  gleichmäßig, d. h. die Cauchybedingung bezüglich der oberen Grenze kann gleichmäßig für alle  $n$  erfüllt werden, so ist  $f$  lokal Riemann integrierbar,  $\int_a^\infty f$  existiert und es gilt

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

*Beweis.*

- (i) Die Behauptung  $f \in R_{\text{loc}}(I, E)$  folgt direkt aus Theorem 5.2.14.
- (ii) Wir schreiben die gleichmäßige Konvergenz der Integrale explizit als gleichmäßige Cauchybedingung auf. Es gilt

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

Mit Theorem 5.2.14 können wir zum Grenzwert  $n \rightarrow \infty$  übergehen und erhalten für das  $\varepsilon$  und das  $R$  von oben

$$\left\| \int_x^y f \right\| \leq \varepsilon \quad \text{für alle } x, y \geq R.$$

Diese Cauchybedingung besagt aber gerade, dass  $\int_a^\infty f$  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

$$\left\| \int_R^\infty f_n \right\| + \left\| \int_R^\infty f \right\| \leq 2\varepsilon$$

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

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

$$\left\| \int_a^R (f - f_n) \right\| \leq \varepsilon$$

gilt. Zusammengenommen erhalten wir somit

$$\begin{aligned} \left\| \int_a^\infty (f - f_n) \right\| &\leq \left\| \int_a^R (f - f_n) \right\| + \left\| \int_R^\infty (f - f_n) \right\| \\ &\leq \varepsilon + \left\| \int_R^\infty f \right\| + \left\| \int_R^\infty f_n \right\| \leq 3\varepsilon. \end{aligned}$$

Die Behauptung folgt. □

**Bemerkung 5.8.17.** Eine Folge von uneigentlichen Integralen  $\int_a^\infty f_n$  ist gleichmäßig konvergent, falls es eine Funktion  $g \in R_{\text{loc}}([a, \infty), \mathbb{R})$  mit  $\int_a^\infty g < \infty$  und  $\|f_n(x)\| \leq g(x)$  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. □

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

**Proposition 5.9.1.** Sei  $\Omega$  ein metrischer Raum und  $f \in C^0(I \times \Omega, E)$ . Dann ist

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

in  $\Omega$  stetig.

*Beweis.*

- (i) Sei  $x_n \rightarrow x_0$  eine konvergente Folge in  $\Omega$ . 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  $f_n := f(\cdot, x_n)$  und  $g := f(\cdot, x_0)$ . Aufgrund der gleichmäßigen Stetigkeit folgt  $f_n \rightrightarrows g$  für  $n \rightarrow \infty$ . Da wir den Limes bei gleichmäßiger Konvergenz mit dem Integral vertauschen dürfen, erhalten wir

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

wie behauptet. □

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

**Definition 5.9.2** (Partielle Ableitungen).

- (i) Seien  $\Omega_i \subset \mathbb{R}$ ,  $1 \leq i \leq n$ , offene nichtleere Intervalle. Sei  $\Omega := \prod_{i=1}^n \Omega_i \subset \mathbb{R}^n$  und sei  $f: \Omega \rightarrow E$  eine Funktion. Dann heißt die Funktion  $f$  im Punkt  $x_0 = (x_0^i)$  nach  $x^k$  oder in Richtung  $k$  **partiell differenzierbar**, falls die Funktion  $g: \Omega_k \rightarrow E$  mit  $g(t) := f(x_0^1, \dots, x_0^{k-1}, t, x_0^{k+1}, \dots, x_0^n) \equiv f(\hat{x}_0^k, t)$  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) ★ Eine entsprechende Definition gilt auch im Falle, dass eine offene Menge  $\Omega \subset \mathbb{R}^n$  nicht in Produktgestalt ist. Wir schränken dabei  $f$  auf eine Umgebung in Produktgestalt ein.
- (iii) Ist  $f$  in ganz  $\Omega$  partiell nach  $x^k$  differenzierbar, so heißt  $f$  in  $\Omega$  nach  $x^k$  differenzierbar. Wir erhalten in diesem Fall eine Abbildung  $D_k f: \Omega \rightarrow E$ .
- (iv) Ist die **partielle Ableitung**  $D_k f$  stetig, so heißt  $f$  **stetig partiell** nach  $x^k$  **differenzierbar**.
- (v)  $f$  heißt **partiell differenzierbar**, falls alle Ableitungen  $D_k f$ ,  $1 \leq k \leq n$ , existieren und **stetig partiell differenzierbar**, falls diese zusätzlich stetig sind.

- (vi) 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.

**Theorem 5.9.3.** 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

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

stetig differenzierbar und es gilt

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

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

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

Wir schätzen dies mit der Dreiecksungleichung für Integrale und dem Mittelwertsatz in der Form von Korollar 4.1.33 ab und erhalten

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

wobei das  $\xi$ , in dem das Supremum bzw. Maximum angenommen wird, von  $t$  abhängen darf. Aufgrund der gleichmäßigen Stetigkeit der Einschränkung von  $\frac{\partial f}{\partial x}$  auf die kompakte Menge  $[a, b] \times \overline{B_r(x_0)}$  für  $r > 0$  mit  $B_{2r}(x_0) \subset \Omega$  konvergiert die rechte Seite für  $x \rightarrow x_0$  gegen Null und wir erhalten aus  $0 = \lim_{x \rightarrow x_0} I$  die Behauptung.

Die Stetigkeit der Ableitung folgt nun aus Proposition 5.9.1.  $\square$

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

**Korollar 5.9.4.** Sei  $\Omega \subset \mathbb{R}$  nichtleer und offen, seien  $f$ ,  $\frac{\partial f}{\partial x} \in C^0(I \times \Omega, E)$ . Seien  $\varphi, \psi: \Omega \rightarrow \overset{\circ}{I}$  in  $x_0 \in \Omega$  differenzierbar, so ist

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

in  $x_0$  differenzierbar und es gilt

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

*Beweis.*

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

$$F(x) = \int_{\psi(x)}^{\varphi(x)} f(t, x) dt = \int_0^{\varphi(x)} f(t, x) dt - \int_0^{\psi(x)} f(t, x) dt$$

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

(ii) Es gilt

$$\begin{aligned} & F(x) - F(x_0) \\ &= \int_0^{\varphi(x)} f(t, x) dt - \int_0^{\varphi(x_0)} f(t, x_0) dt \\ &= \int_0^{\varphi(x)} f(t, x) - f(t, x_0) dt + \int_0^{\varphi(x)} f(t, x_0) dt - \int_0^{\varphi(x_0)} f(t, x_0) dt \\ &= \int_0^{\varphi(x_0)} f(t, x) - f(t, x_0) dt + \int_{\varphi(x_0)}^{\varphi(x)} f(t, x) - f(t, x_0) dt + \int_{\varphi(x_0)}^{\varphi(x)} f(t, x_0) dt. \end{aligned}$$

Wir dividieren nun durch  $x - x_0$  und lassen  $x \rightarrow x_0$ . Dann konvergiert der erste Term auf der rechten Seite wegen Theorem 5.9.3 gegen den Integralterm aus der Behauptung und der dritte Term konvergiert nach Proposition 5.4.9 gegen den ersten Term auf der rechten Seite in der Behauptung.

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

$$0 \stackrel{!}{=} \lim_{x \rightarrow x_0} \left\| \frac{1}{x - x_0} \int_{\varphi(x_0)}^{\varphi(x)} f(t, x) - f(t, x_0) dt \right\| \equiv \lim_{x \rightarrow x_0} I.$$

Es gilt aufgrund des vektorwertigen Mittelwertsatzes

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

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

**Theorem 5.9.5** (Doppelintegrale). *Seien  $I = [a, b]$  und  $J = [\alpha, \beta]$  kompakte Intervalle und  $f \in C^0(I \times J, E)$ , so gilt*

$$\int_{\alpha}^{\beta} \left( \int_a^b f(t, x) dt \right) dx = \int_a^b \left( \int_{\alpha}^{\beta} f(t, x) dx \right) dt,$$

d. h. es kommt nicht auf die Integrationsreihenfolge an und wir schreiben auch  $\int_{I \times J} f$ .

*Beweis.* Nach Proposition 5.9.1 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

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

Es gilt  $\varphi(\alpha) = 0 = \psi(\alpha)$ . Weiterhin sind  $\varphi$  und  $\psi$  in  $(\alpha, \beta)$  nach Theorem 5.9.3 ( $\psi$ ) (die dafür nötigen Voraussetzungen wollen wir nachfolgend prüfen), Proposition 5.9.1 ( $\varphi$  und  $\psi$ ) und dem Hauptsatz der Differential- und Integralrechnung ( $\varphi$  und  $\psi$ ) differenzierbar und es gilt dort

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

Somit stimmen  $\varphi$  und  $\psi$  überein.

Für die Anwendbarkeit von Theorem 5.9.3 müssen wir noch die Stetigkeit von  $g: I \times J \rightarrow E$  mit

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

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 5.9.1 stetig ist, gibt es ein  $\delta > 0$ , so dass  $\|g(t, y_0) - g(t_0, y_0)\| < \varepsilon$  gilt. Wir erhalten

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

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

Per Induktion erhält man daraus die folgende Verallgemeinerung.

**Theorem 5.9.6** (Mehrfachintegrale).  $\star$  Seien  $I_i = [a_i, b_i]$ ,  $i = 1, \dots, n$ , kompakte Intervalle,  $E$  ein Banachraum und  $f \in C^0\left(\prod_{i=1}^n I_i, E\right)$ , so ist

$$\int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x^1, \dots, x^n) dx^1 \cdots dx^n$$

wohldefiniert und ergibt unabhängig von der Integrationsreihenfolge stets denselben Wert.

*Parameterabhängige uneigentliche Integrale*  $\star$ . Wir behandeln hier nur den Fall  $I = [a, \infty)$ .

**Proposition 5.9.7.** Sei  $I = [a, \infty)$ , sei  $\Omega$  ein metrischer Raum und sei  $f \in C^0(I \times \Omega, E)$ . Angenommen, die Integrale

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

konvergieren gleichmäßig in  $x \in \Omega$ , d. h. der Limes  $\lim_{T \rightarrow \infty} \int_a^T f(t, x) dt$  ist gleichmäßig in  $x$ , dann ist  $F$  in  $\Omega$  stetig.

*Beweis.* Wir gehen ähnlich wie beim Beweis von Proposition 5.9.1 vor.

Sei  $x_n \rightarrow x_0$  eine konvergente Folge in  $\Omega$ . Dann können wir  $f$  auf  $I \times K$  mit der kompakten Menge  $K = \{x_n : n \in \mathbb{N}\}$  einschränken und dürfen daher ohne Einschränkung annehmen, dass  $f$  lokal gleichmäßig stetig ist.

Somit konvergiert  $f_n := f(\cdot, x_n)$  lokal gleichmäßig gegen  $f := f(\cdot, x_0)$ . Nach Theorem 5.8.16 folgt  $\int f_n \rightarrow \int f$  und somit die Behauptung.  $\square$

**Proposition 5.9.8.** Sei  $I = [a, \infty)$ , sei  $\Omega \subset \mathbb{K}$  offen, seien  $f, D_2 f \equiv \frac{\partial f}{\partial x} \in C^0(I \times \Omega, E)$ . Angenommen,

$$\int_a^\infty \frac{\partial f}{\partial x}(t, x) dt$$

konvergiert gleichmäßig in  $x \in \Omega$  und

$$\int_a^\infty f(t, x) dt$$

konvergiert für alle  $x \in \Omega$ .

Dann konvergieren die Integrale

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

lokal gleichmäßig in  $\Omega$ ,  $F$  ist stetig differenzierbar und es gilt

$$F'(x) = \int_a^\infty \frac{\partial f}{\partial x}(t, x) dt.$$

*Beweis.*

(i) Sei  $a \leq t_n \rightarrow \infty$  eine beliebige Folge. Wir definieren

$$F_n(x) := \int_a^{t_n} f(t, x) dt.$$

Nach Definition von  $F$  gilt  $F_n(x) \rightarrow F(x)$  für alle  $x \in \Omega$ .

Da das Integrationsgebiet  $[a, t_n]$  kompakt ist, erhalten wir aus Theorem 5.9.3 die Differenzierbarkeit von  $F_n$  und eine Formel für die Ableitung. Da die Ausdrücke für die Ableitung  $F'_n$  nach Voraussetzung gleichmäßig in  $x$  konvergieren, erhalten wir

$$F'_n(x) = \int_a^{t_n} \frac{\partial f}{\partial x}(t, x) dt \Rightarrow \int_a^\infty \frac{\partial f}{\partial x}(t, x) dt.$$

Die Funktionenfolge  $F_n$  konvergiert also punktweise und die Ableitungen  $F'_n$  konvergieren gleichmäßig. Daher dürfen wir die beiden Grenzwerte ( $n \rightarrow \infty$  und Differentiation) vertauschen und erhalten

$$F'(x) = \lim_{n \rightarrow \infty} F'_n(x) = \int_a^\infty \frac{\partial f}{\partial x}(t, x) dt.$$

- (ii) Zur lokal gleichmäßigen Konvergenz: Sei dazu  $B = B_\delta \subset \Omega$  ein Ball. Wir wollen  $F_n \rightrightarrows F$  in  $B$  zeigen. Aufgrund des Mittelwertsatzes gilt für  $x, x_0 \in B$

$$\|(F_n(x) - F(x)) - (F_n(x_0) - F(x_0))\| \leq \sup_{\xi \in [x_0, x]} \|F'_n(\xi) - F'(\xi)\| \cdot |x - x_0|.$$

Aufgrund der Dreiecksungleichung erhalten wir in  $B$

$$\|F_n(x) - F(x)\| \leq \|F_n(x_0) - F(x_0)\| + \sup_{\xi \in B} \|F'_n(\xi) - F'(\xi)\| \cdot \delta.$$

Auf der rechten Seite konvergiert nun für  $n \rightarrow \infty$  der erste Term aufgrund der punktweisen Konvergenz  $F_n \rightarrow F$  in  $\Omega$  gegen Null und wegen  $F'_n \rightrightarrows F'$  konvergiert der zweite Term gegen Null.  $\square$

Für Doppelintegrale mit einem unbeschränkten Integrationsbereich erhalten wir

**Proposition 5.9.9.** *Seien  $I = [a, \infty)$ ,  $J = [\alpha, \beta]$ ,  $f \in C^0(I \times J, E)$  und sei*

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

*gleichmäßig konvergent. Dann gilt*

$$\int_\alpha^\beta F(x) dx = \int_\alpha^\beta \int_a^\infty f(t, x) dt dx = \int_a^\infty \int_\alpha^\beta f(t, x) dx dt.$$

*Beweis.* Nach Proposition 5.9.7 ist  $F$  stetig. Somit ist das Integral auf der linken Seite wohldefiniert.

Sei nun  $a \leq t_n \rightarrow \infty$  eine beliebige Folge. Nach Theorem 5.9.5 erhalten wir

$$\underbrace{\int_\alpha^\beta \int_a^{t_n} f(t, x) dt dx}_{=: F_n(x)} = \int_a^{t_n} \int_\alpha^\beta f(t, x) dx dt.$$

Nach Voraussetzung gilt  $F_n \rightrightarrows F$  und das Integrationsintervall  $[\alpha, \beta]$  kompakt ist, dürfen wir den Grenzwert  $\lim_{n \rightarrow \infty}$  und  $\int_\alpha^\beta$  vertauschen und erhalten

$$\begin{aligned} \int_\alpha^\beta \int_a^\infty f(t, x) dt dx &= \int_\alpha^\beta \lim_{n \rightarrow \infty} \int_a^{t_n} f(t, x) dt dx = \lim_{n \rightarrow \infty} \int_\alpha^\beta \int_a^{t_n} f(t, x) dt dx \\ &= \lim_{n \rightarrow \infty} \int_a^{t_n} \int_\alpha^\beta f(t, x) dx dt = \int_a^\infty \int_\alpha^\beta f(t, x) dx dt \end{aligned}$$

wie behauptet.  $\square$

Lassen wir die gleichmäßige Konvergenz in der Voraussetzung weg, so gibt es Gegenbeispiele:

**Beispiel 5.9.10.** Sei  $f \in C^1([-\varepsilon, \infty))$  und gelte  $f(0) = \lim_{t \rightarrow \infty} f(t)$ . Wir definieren  $F$  nachfolgend und erhalten

$$F(x) := \int_0^\infty f'(xt) dt = \frac{1}{x} f(xt) \Big|_0^\infty = 0$$

für alle  $x > 0$ . Somit gilt  $\int_0^\beta F(x) dx = 0$  für alle  $\beta > 0$ . Andererseits gilt für beliebige  $t > 0$

$$\int_0^\beta f'(xt) dx = \frac{1}{t}(f(\beta t) - f(0)).$$

Somit folgt

$$\int_0^\infty \int_0^\beta f'(xt) dx dt = \int_0^\infty \frac{1}{t}(f(\beta t) - f(0)) dt.$$

Im Falle  $f(t) = t^2 e^{-t}$  und  $\beta = 1$  erhalten wir für die rechte Seite

$$\int_0^\infty t \beta^2 e^{-\beta t} dt > 0.$$

Somit dürfen wir in diesem Falle die Integrationsreihenfolge nicht umdrehen.

*Doppelintegrale mit unbeschränktem Integrationsbereich  $\star$ .*

**Theorem 5.9.11.** Seien  $I = [a, \infty)$  und  $J = [\alpha, \infty)$ . Sei  $f \in C^0(I \times J, E)$ . Angenommen, die uneigentlichen Integrale

$$\int_a^\infty \|f(t, x)\| dt \quad \text{und} \quad \int_\alpha^\infty \|f(t, x)\| dx$$

konvergieren gleichmäßig. Existiert dann eines der beiden Doppelintegrale

$$\int_a^\infty \int_\alpha^\infty \|f(t, x)\| dx dt \quad \text{oder} \quad \int_\alpha^\infty \int_a^\infty \|f(t, x)\| dt dx,$$

so existieren beide und ihre Werte stimmen überein.

*Beweis.* Angenommen,  $\int_a^\infty \int_\alpha^\infty \|f(t, x)\| dx dt$  existiert. Sei  $\beta > \alpha$ . Nach Proposition 5.9.9 erhalten wir dann

$$\int_\alpha^\beta \int_a^\infty \|f(t, x)\| dx dt = \int_a^\beta \int_\alpha^\beta \|f(t, x)\| dx dt \leq \int_a^\beta \int_\alpha^\infty \|f(t, x)\| dx dt.$$

Die Existenz dieser in  $\beta$  gleichmäßigen oberen Schranke besagt gerade, dass auch das zweite Doppelintegral konvergiert. Im Grenzwert  $\beta \rightarrow \infty$  erhalten wir daraus

$$\int_\alpha^\infty \int_a^\infty \|f(t, x)\| dx dt \leq \int_a^\infty \int_\alpha^\infty \|f(t, x)\| dx dt.$$

Mit vertauschten Rollen der beiden Integrale erhalten wir die umgekehrte Abschätzung und somit insgesamt die Behauptung.  $\square$

Für die Integrale ohne Normen erhalten wir

**Korollar 5.9.12.** Sei die Situation aus Theorem 5.9.11 gegeben.

Nochmals explizit: Seien  $I = [a, \infty)$ ,  $J = [\alpha, \infty)$ . Sei  $f \in C^0(I \times J, E)$ . Angenommen, die uneigentlichen Integrale

$$\int_a^\infty \|f(t, x)\| dt \quad \text{und} \quad \int_\alpha^\infty \|f(t, x)\| dx$$

konvergieren gleichmäßig und eines der Doppelintegrale

$$\int_a^\infty \int_\alpha^\infty \|f(t, x)\| dx dt \quad \text{oder} \quad \int_\alpha^\infty \int_a^\infty \|f(t, x)\| dt dx$$

existiert (und damit nach Theorem 5.9.11 beide).

Dann existieren die Doppelintegrale

$$\int_a^\infty \int_\alpha^\infty f(t, x) dx dt \quad \text{und} \quad \int_\alpha^\infty \int_a^\infty f(t, x) dt dx$$

und stimmen überein.

*Beweis.*

- (i) In einem Banachraum impliziert absolute Konvergenz einfache Konvergenz. Somit konvergieren die Integrale

$$F(x) := \int_a^\infty f(t, x) dt \quad \text{und} \quad G(t) := \int_\alpha^\infty f(t, x) dx$$

gleichmäßig und sind nach Proposition 5.9.7 (wie im Beweis von Proposition 5.9.9) stetige Funktionen.

- (ii) Existenz: Nach Theorem 5.9.11 konvergieren die Doppelintegrale absolut. Daher konvergieren auch die uneigentlichen Integrale über  $F$  bzw.  $G$  absolut. Daraus folgt die Existenz der beiden angegebenen Doppelintegrale.
- (iii) Gleichheit: Da die Integrale  $F(x)$  gleichmäßig in  $x$  konvergieren, können wir Proposition 5.9.9 anwenden und erhalten für alle  $n > \alpha$

$$(5.3) \quad \int_a^\infty \int_\alpha^n f(t, x) dx dt = \int_\alpha^n \int_a^\infty f(t, x) dt dx.$$

Definiere nun

$$G_n(t) := \int_\alpha^n f(t, x) dx.$$

Dann konvergieren die Funktionen  $G_n$  für  $n \rightarrow \infty$  nach Voraussetzung gleichmäßig nach  $G$ . Andererseits gilt

$$\|G_n(t)\| \leq \int_\alpha^n \|f(t, x)\| dx \leq \int_\alpha^\infty \|f(t, x)\| dx \equiv \Gamma(t).$$

$\Gamma$  ist eine von  $n$  unabhängige obere über das Intervall  $[a, \infty)$  integrierbare Schranke. Somit folgt nach Bemerkung 5.8.17, dass die Integrale  $\int_a^\infty G_n(t) dt$  gleichmäßig konvergent sind. Nun wenden wir Theorem 5.8.16 auf  $G_n$  und  $G$  an und erhalten unter Benutzung von (5.3)

$$\begin{aligned} \int_a^\infty \int_\alpha^\infty f(t, x) dx dt &= \int_a^\infty G(t) dt = \lim_{n \rightarrow \infty} \int_a^\infty G_n(t) dt \\ &= \lim_{n \rightarrow \infty} \int_a^\infty \int_\alpha^n f(t, x) dx dt = \lim_{n \rightarrow \infty} \int_\alpha^n \int_a^\infty f(t, x) dt dx \end{aligned}$$

$$= \int_{\alpha}^{\infty} \int_a^{\infty} f(t, x) dt dx$$

wie behauptet.  $\square$

## 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 6.1.1.

- (i) Seien  $E, F$  Banachräume. Sei  $\Omega \subset E$  offen. Dann heißt eine Abbildung  $f: \Omega \rightarrow F$  in  $x_0 \in \Omega$  **differenzierbar**, falls es eine stetige lineare Abbildung  $A \in L(E, F)$  mit

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

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  $\Omega$  differenzierbar, so heißt  $f$  in  $\Omega$  **differenzierbar**. Wir nennen dann die Abbildung

$$\begin{aligned} Df: \Omega &\rightarrow L(E, F), \\ x &\mapsto Df(x) \end{aligned}$$

**Ableitung** oder **Differential** von  $f$ .

- (iv) Ist  $Df: \Omega \rightarrow L(E, F)$  stetig, so heißt  $f$  in  $\Omega$  **stetig differenzierbar**.

**Bemerkung 6.1.2.**  $\star$  Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow F$  in  $x_0 \in \Omega$  differenzierbar.

- (i) Ist  $\dim E < \infty$ , so ist jede lineare Abbildung  $A: E \rightarrow F$  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.  
 (v) Ist  $E = \mathbb{R}^0 = \{0\}$ , so ist  $Df(0) = 0$  und  $L(\mathbb{R}^0, F) = \{0\}$ .

**Bemerkung 6.1.3.**  $\star$  Seien  $E, F$  Banachräume und sei  $\Omega \subset E$  offen. Seien die Funktionen  $f, g: \Omega \rightarrow F$  in  $x_0 \in \Omega$  differenzierbar. Dann sind auch  $f + g$  und  $\lambda f$  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  $\Omega \rightarrow F$ .

*Beweis.* Klar.  $\square$

#### Beispiele 6.1.4.

- (i) Konkretes Beispiel: Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 \\ x^2 y \\ y^3 \end{pmatrix}.$$

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

$$\left| \frac{u^2}{\|(u, v)\|} \right| = \left| \frac{u^2}{\sqrt{u^2 + v^2}} \right| = \left| \frac{u}{\sqrt{u^2 + v^2}} \right| \cdot |u| \leq |u|$$

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  $3 \times 2$ -Matrix gerade die partiellen Ableitungen der Komponenten stehen. Wir werden dies später allgemeiner zeigen.

- (ii) Sei  $f: \Omega \rightarrow F$  konstant. Dann ist  $f$  in jedem Punkt aus  $\Omega$  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  $E_i = \mathbb{R}$ . Wir werden multilineare Abbildungen später für die Behandlung höherer Ableitungen benutzen.

**Definition 6.1.5.** Seien  $E_1, E_2, \dots, E_n, F$  Banachräume. Sei

$$A: E_1 \times E_2 \times \dots \times E_n \equiv \prod_{j=1}^n E_j \rightarrow F$$

eine Abbildung.

- (i) Sind die Abbildungen  $E_i \ni z \mapsto A(x^1, \dots, x^{i-1}, z, x^{i+1}, \dots, x^n)$  für alle  $(x^1, \dots, \widehat{x^i}, \dots, x^n) \in \prod_{\substack{j=1 \\ j \neq i}}^n E_j$  und alle  $1 \leq i \leq n$  linear, so heißt  $A$  **multilinear**. Wir verwenden auch in diesem Fall spitze Klammern,  $A(\dots)$ .
- (ii) Wir bezeichnen den Banachraum aller stetigen multilinearen Abbildungen

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

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

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

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

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

**Bemerkung 6.1.6.**

- (i) Sei  $A: E_1 \times \dots \times E_n \rightarrow F$  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  $c \geq 0$ , so dass die Abschätzung

$$\|A \langle x^1, \dots, x^n \rangle\| \leq c \cdot \|x^1\| \cdot \dots \cdot \|x^n\|$$

für alle  $(x^1, \dots, x^n) \in \prod_{i=1}^n E_i$  gilt.

- (ii)  $L(E_1, \dots, E_n; F)$  ist mit der angegebenen Norm ein Banachraum.
- (iii) Ist  $A: E_1 \times E_2 \rightarrow F$  multilinear, so gilt  $A(\lambda x, \lambda y) = \lambda^2 A(x, y)$ .  
Ist  $A: E_1 \times E_2 \rightarrow F$  linear, so gilt  $A(\langle \lambda x, \lambda y \rangle) = \lambda A(\langle x, y \rangle)$ .
- (iv) Es gibt einen normerhaltenden Isomorphismus

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

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

- (v) ★ Analoge Bezeichnungen verwendet man in der Linearen Algebra auch für beliebige Vektorräume  $E_1, \dots, E_n, F$  und lineare nicht notwendigerweise stetige Abbildungen.

*Beweis.* Wichtige Übung. □

**Beispiel 6.1.7.** 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

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

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  $\|u\| + \|v\| \leq 1$ . Dann folgt

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

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

★ 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 6.1.8** (Gradient). Seien  $H$  ein Hilbertraum,  $\Omega \subset H$  offen und  $f: \Omega \rightarrow \mathbb{K}$  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  $\text{grad } f(x_0)$  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 6.1.9.

(i) ★ Konkretes Beispiel zum Selbststudium:

Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  die Abbildung  $(x, y) \mapsto x^2 \cdot y$ . Dann gelten

$$Df(x_0, y_0)\langle u, v \rangle = 2x_0y_0u + x_0^2v$$

und

$$\nabla f(x_0, y_0) = \begin{pmatrix} 2x_0y_0 \\ x_0^2 \end{pmatrix}.$$

*Beweis.* Wir benutzen die Definition der Ableitung

$$f((x_0, y_0) + (u, v)) = f((x_0, y_0)) + Df((x_0, y_0))\langle(u, v)\rangle + o(\|(u, v)\|)$$

und erhalten

$$\begin{aligned} (x_0 + u)^2(y_0 + v) - x_0^2y_0 &= x_0^2y_0 - x_0^2y_0 + 2x_0uy_0 + u^2y_0 + x_0^2v + 2x_0uv + u^2v \\ &= 2x_0y_0u + x_0^2v + o(\|(u, v)\|) \\ &= \begin{pmatrix} 2x_0y_0 & x_0^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + o(\|(u, v)\|) \\ &= \left\langle \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} 2x_0y_0 \\ x_0^2 \end{pmatrix} \right\rangle + o(\|(u, v)\|). \end{aligned} \quad \square$$

(ii) Sei  $H$  ein reeller Hilbertraum und  $f: H \rightarrow \mathbb{R}$  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.  $\square$

**Theorem 6.1.10** (Kettenregel). *Seien  $E, F, G$  Banachräume und seien  $\Omega \subset E$  und  $V \subset F$  offen. Seien  $g: \Omega \rightarrow F$  in  $x_0$  und  $f: V \rightarrow G$  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

$$\begin{aligned} f \circ g(x) - f \circ g(x_0) &= f'(g(x_0))\langle g'(x_0)\langle x - x_0 \rangle \rangle + f'(g(x_0))\langle o(\|x - x_0\|) \rangle \\ &\quad + o(\|g(x) - g(x_0)\|). \end{aligned}$$

- (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  $\varepsilon$ , die  $\varepsilon \rightarrow 0$  für  $g(x) \rightarrow g(x_0)$  erfüllt. Wir erhalten nach Anwendung der Norm und der Dreiecksungleichung auf die Definition von Differenzierbarkeit für  $g$

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

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

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

und die Behauptung folgt.  $\square$

**Korollar 6.1.11.** *Seien  $E, E'$  Banachräume, seien  $\Omega \subset E$  und  $\Omega' \subset E'$  offene Teilmengen und sei  $f: \Omega \rightarrow \Omega'$  invertierbar und  $g: \Omega' \rightarrow \Omega$  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  $\text{id}_\Omega = g \circ f$  und  $\text{id}_{\Omega'} = f \circ g$ . Setze  $y_0 := f(x_0)$ . Wir erhalten aufgrund der Kettenregel

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

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

**Theorem 6.1.12.** *Seien  $E, E'$  Banachräume. Seien  $\Omega \subset E$  und  $\Omega' \subset E'$  offene Teilmengen. Sei  $f: \Omega \rightarrow \Omega'$  ein Homöomorphismus mit Inverse  $g: \Omega' \rightarrow \Omega$ . 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  $\|A^{-1}y\| \leq c^{-1} \cdot \|y\|$  für alle  $y$  gilt. Wir wählen speziell  $y = Ax$  und erhalten  $c \cdot \|x\| \leq \|Ax\|$ . Somit erhalten wir in unserer Situation ein  $c > 0$  mit

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

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

$$\begin{aligned} \|f(x) - f(x_0)\| &\geq \|f'(x_0)\langle x - x_0 \rangle\| - o(\|x - x_0\|) \\ &\geq c\|x - x_0\| - \varepsilon(\|x - x_0\|) \cdot \|x - x_0\| \\ &\geq \frac{c}{2}\|x - x_0\| \end{aligned}$$

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

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

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

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

Wegen  $x \rightarrow x_0 \iff y \rightarrow y_0$  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 6.1.13.** *Seien  $E$  ein Banachraum,  $\Omega \subset E$  offen und  $F = \prod_{i=1}^n F_i$  ein Produktraum von Banachräumen. Sei  $f: \Omega \rightarrow F$  mit  $f = (f^i)_{1 \leq i \leq n}$  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*

$$Df(x_0) = (Df^1(x_0), \dots, Df^n(x_0)),$$

wobei wir  $L(E, F)$  und  $\prod_{i=1}^n L(E, F_i)$  (mit Maximumsnorm) wie nachfolgend angeben 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  $A: x \mapsto (A_1x, \dots, A_nx)$  zu.
- (c) Nach Definition der Maximumsnorm auf dem Produktraum gilt  $\|Ax\| = \max_{1 \leq i \leq n} \|A_ix\|$  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  $\pi_i: F \rightarrow F_i$  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

$$\begin{aligned} f(x) &= f(x_0) + (Df^1(x_0), \dots, Df^n(x_0)) \langle x - x_0 \rangle \\ &\quad + \underbrace{(o(\|x - x_0\|), \dots, o(\|x - x_0\|))}_{\in o(\|x - x_0\|)}, \end{aligned}$$

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

**Proposition 6.1.14** (Produkt- und Quotientenregel). *Seien  $E$  ein Banachraum,  $\Omega \subset E$  offen und  $f, g: \Omega \rightarrow \mathbb{K}$  in  $x_0 \in \Omega$  differenzierbar.*

- (i) *Dann ist auch  $h := fg$  in  $x_0$  differenzierbar und es gilt für  $Dh(x_0) \in L(E; \mathbb{K})$  sowie  $v \in E$*

$$Dh(x_0)\langle v \rangle = Df(x_0)\langle v \rangle \cdot g(x_0) + f(x_0) \cdot Dg(x_0)\langle v \rangle \in \mathbb{K}.$$

- (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  $\varphi: \Omega \rightarrow \mathbb{K}^2$  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.  $\square$

**Theorem 6.1.15.** *Seien  $E$  ein Banachraum,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow \mathbb{R}$  eine in  $x_0 \in \Omega$  differenzierbare Funktion, die in  $x_0$  ein lokales Maximum annimmt. Dann gilt  $Df(x_0) = 0$ .*

Analog zum Fall eines eindimensionalen Definitionsgebietes definieren wir kritische Punkte.

**Definition 6.1.16.** Sei  $E$  ein Banachraum,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow \mathbb{R}$  in  $x_0 \in \Omega$  differenzierbar. Dann heißt  $x_0$  **kritischer Punkt**, falls  $Df(x_0) = 0$  gilt.

*Beweis von Theorem 6.1.15.* 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  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  mit  $\varphi(t) := f(x_0 + te)$  in  $t = 0$  differenzierbar und nimmt in  $t = 0$  ein lokales Maximum an. Somit gilt nach Proposition 4.1.24

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

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

Die Funktionenräume der einmal stetig differenzierbaren Funktionen definieren wir formal genauso wie für  $\Omega \subset \mathbb{R}$ . Auch hier ist  $C^1(\overline{\Omega}, F)$  mit der  $C^1$ -Norm ein Banachraum.

**Definition 6.1.17.** ★ Seien  $E, F$  Banachräume und  $\Omega \subset E$  offen. Dann definieren wir die folgenden Funktionenräume:

- (i)  $C^1(\Omega, F) := \{f: \Omega \rightarrow F : f \text{ ist in } \Omega \text{ stetig differenzierbar}\}.$
- (ii)  $C^1(\overline{\Omega}, F) := \{f \in C^1(\Omega, F) : f \text{ und } Df \text{ lassen sich stetig und beschränkt auf } \overline{\Omega} \text{ fortsetzen}\}.$
- (iii) Wir definieren die  $C^1$ -Norm auf  $C^1(\overline{\Omega}, F)$  durch

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

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

**Theorem 6.1.18.** Seien  $E, F$  Banachräume und sei  $\Omega \subset E$  offen. Dann ist der Vektorraum  $C^1(\overline{\Omega}, F)$  mit der  $C^1$ -Norm ein Banachraum.

*Beweis.* Übung.  $\square$

*Eulersche Homogenitätsrelation.*

**Definition 6.1.19.**

- (i) Sei  $E$  ein  $\mathbb{R}$ -Vektorraum. Dann heißt  $\Gamma \subset E$  ein **Kegel**, falls

$$x \in \Gamma \implies tx \in \Gamma$$

für alle  $t > 0$  gilt.

- (ii) Seien  $E, F$  zwei  $\mathbb{R}$ -Vektorräume und  $\Gamma \subset E$  ein Kegel. Dann heißt eine Funktion  $\varphi: \Gamma \rightarrow F$  **positiv homogen** vom Grade  $\alpha$  mit  $\alpha \in \mathbb{R}$ , falls

$$\varphi(tx) = t^\alpha \varphi(x)$$

für alle  $t > 0$  und alle  $x \in \Gamma$  gilt.

**Proposition 6.1.20** (Eulersche Homogenitätsrelation). Seien  $E, F$  Banachräume,  $\Omega \subset E$  ein offener Kegel und sei  $\varphi: \Omega \rightarrow F$  differenzierbar. Dann ist  $\varphi$  genau dann positiv homogen vom Grade  $\alpha$ , wenn die **Eulersche Homogenitätsrelation**

$$D\varphi(x)\langle x \rangle = \alpha \varphi(x)$$

für alle  $x \in \Omega$  gilt.

*Beweis.*

„ $\Rightarrow$ “: Sei  $\varphi$  positiv homogen vom Grade  $\alpha$ . 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.

„ $\Leftarrow$ “:  $\star$  Gelte umgekehrt die Eulersche Homogenitätsrelation. Definiere wieder  $f(t) := \varphi(tx)$  für festes  $x \in \Omega$  und  $t > 0$ . Wir differenzieren und erhalten mit Hilfe der Eulerschen Homogenitätsrelation

$$f'(t) = D\varphi(tx)\langle x \rangle = \frac{1}{t}D\varphi(tx)\langle tx \rangle = \frac{1}{t}\alpha\varphi(tx) = \frac{1}{t}\alpha f(t).$$

Dies ist eine Differentialgleichung für  $f$ . Andererseits löst auch  $h(t) := t^\alpha f(1)$ ,  $t > 0$ , dieselbe Differentialgleichung, wie man direkt nachrechnet. Nun gilt für  $t > 1$  aufgrund des Hauptsatzes der Differential- und Integralrechnung

$$\begin{aligned} \|f(t) - h(t)\| &\leq \underbrace{\|f(1) - h(1)\|}_{=0} + \left\| \int_1^t f'(s) ds - \int_1^t h'(s) ds \right\| \\ &\leq \int_1^t \|f'(s) - h'(s)\| ds = \int_1^t \frac{1}{s} |\alpha| \cdot \|f(s) - h(s)\| ds \end{aligned}$$

und das Gronwallsche Lemma, angewandt auf  $\psi(t) = \|f(t) - h(t)\|$ , liefert aufgrund der Beschränktheit von  $\frac{1}{s}|\alpha|$  für  $s \in [1, t]$  die Eindeutigkeit. (Solch eine Argumentation wird uns im dritten Semester noch in allgemeineren Situationen begegnen. Der Fall  $\alpha = 0$  funktioniert auch direkt ohne Gronwallsches Lemma.) Eine analoge Argumentation funktioniert im Falle  $t < 1$ . Somit folgt  $f = h$  und wir erhalten die Behauptung.  $\square$

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

**Theorem 6.2.1** (Mittelwertsatz (MWS)). *Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und sei  $f: \Omega \rightarrow F$  differenzierbar. Seien  $x, y \in \Omega$  mit  $[x, y] \subset \Omega$ . Dann folgt*

$$\|f(x) - f(y)\| \leq \sup_{0 \leq t \leq 1} \|Df(tx + (1-t)y)\| \cdot \|x - y\| = \sup_{z \in [x, y]} \|Df(z)\| \cdot \|x - y\|.$$

*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  $\Omega$  offen ist, ist  $\varphi$  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

$$\begin{aligned} \|f(x) - f(y)\| &= \|\varphi(1) - \varphi(0)\| \leq \sup_{0 \leq t \leq 1} \|\varphi'(t)\| \cdot |1 - 0| \\ &= \sup_{0 \leq t \leq 1} \|Df(tx + (1-t)y)\langle x - y \rangle\| \\ &\leq \sup_{0 \leq t \leq 1} \|Df(tx + (1-t)y)\| \cdot \|x - y\| \end{aligned}$$

wie behauptet.  $\square$

Analog zum Eindimensionalen erhalten wir auch hier das folgende

**Korollar 6.2.2.** *Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow F$  differenzierbar. Seien  $x, y, x_0 \in \Omega$  beliebig mit  $[x, y] \subset \Omega$ . Dann gilt*

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

*Beweis.* Wende den Mittelwertsatz auf die Funktion

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

an.  $\square$

**Bemerkung 6.2.3.** ★ Auch hier können wir wie im Eindimensionalen  $Df(x_0)$  durch ein beliebiges Element  $A \in L(E, F)$  ersetzen.

**Korollar 6.2.4.** Seien  $E, F$  Banachräume,  $\Omega \subset E$  ein **Gebiet**, d. h. eine offene und zusammenhängende Menge, und  $f: \Omega \rightarrow F$  differenzierbar. Gilt  $Df \equiv 0$  in  $\Omega$ , so ist  $f$  konstant.

*Beweis.* Sei  $x_0 \in \Omega$ . Aus Theorem 6.2.1 folgt, dass  $f$  auf Geradenstücken in  $\Omega$ , d. h. Mengen der Form  $[x, y] \subset \Omega$ , konstant ist. Somit folgt, dass  $\Lambda := \{x \in \Omega: f(x) = f(x_0)\}$  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  $[y, z] \subset \Omega$  ist  $f$  auf  $[y, z]$  konstant. Somit folgt insbesondere  $z \in \Lambda$  und  $\Lambda$  ist offen.
- abgeschlossen: Aufgrund der Stetigkeit von  $f$  ist  $\Lambda$  (relativ) abgeschlossen. (Wie bei der Offenheit könnte man hier auch vom Grenzwert einer in  $\Omega$  konvergenten Folge her argumentieren.)
- nichtleer: Es gilt  $x_0 \in \Lambda$ .

Da  $\Omega$  zusammenhängend ist, folgt die Behauptung.  $\square$

**6.3. Differentiation von Funktionenfolgen.** Wir erhalten ähnliche Resultate wie im Eindimensionalen, z. B. wie in Theorem 4.3.1.

**Theorem 6.3.1.** Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen. Sei  $(f_n)_n \subset C^1(\Omega, F)$  eine Funktionenfolge mit  $f_n \rightarrow f$  und  $Df_n \rightrightarrows g$  für Funktionen  $f: \Omega \rightarrow F$  und  $g: \Omega \rightarrow L(E, F)$ . Dann ist  $f$  ebenfalls stetig differenzierbar und es gilt  $Df = g$ . Ist  $\Omega$  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  $\Omega$  ein Ball ist. Seien  $x, x_0 \in \Omega$ . Der Mittelwertsatz, Korollar 6.2.2, impliziert

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

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

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

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  $n \rightarrow \infty$  gegen Null. Somit folgt

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

Da  $g$  stetig ist, folgt  $\sup_{y \in [x, x_0]} \|g(y) - g(x_0)\| \rightarrow 0$  für  $x \rightarrow x_0$ . 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  $d := \text{diam } \Omega$ . Sei  $\varepsilon > 0$ . Sei  $x \in \Omega$  beliebig und sei  $y \in \Omega$  fest gewählt. Dann impliziert der Mittelwertsatz, angewandt auf die Funktion  $f_n - f_m$ ,

$$\|(f_n(x) - f_m(x)) - (f_n(y) - f_m(y))\| \leq \sup_{z \in \Omega} \|Df_n(z) - Df_m(z)\| \cdot d.$$

Aufgrund der gleichmäßigen Konvergenz ist die rechte Seite für große  $n, m$  kleiner als  $\varepsilon$  und aufgrund der punktweisen Konvergenz gilt  $\|f_n(y) - f_m(y)\| \leq \varepsilon$  für große  $n, m$ . Somit folgt für solche  $n, m$  aufgrund der Dreiecksungleichung für beliebige  $x \in \Omega$

$$\begin{aligned} & \|f_n(x) - f_m(x)\| \\ & \leq \|(f_n(x) - f_m(x)) - (f_n(y) - f_m(y))\| + \|f_n(y) - f_m(y)\| \leq 2\varepsilon. \end{aligned}$$

Die Behauptung folgt.  $\square$

Aufgrund der Überlegungen des letzten Abschnittes können wir statt  $f_n \rightarrow f$  auch nur  $f_n(x_0) \rightarrow f(x_0)$  für ein  $x_0 \in \Omega$  fordern und erhalten

**Theorem 6.3.2.** *★ Seien  $E, F$  Banachräume und  $\Omega \subset E$  ein Gebiet. Sei  $(f_n)_n \in C^1(\Omega, F)$  eine Funktionenfolge mit  $f_n(x_0) \rightarrow f(x_0)$  für ein  $x_0 \in \Omega$  (wobei wir  $f(x_0)$  als diesen Grenzwert definieren). Nehme an, dass  $Df_n$  lokal gleichmäßig gegen  $g$  konvergiert. Dann definiert  $f(x) := \lim_{n \rightarrow \infty} f_n(x)$  für  $x \in \Omega$  eine Funktion  $f: \Omega \rightarrow F$ . Weiterhin ist  $f$  stetig differenzierbar, es gilt  $Df = g$  und die Funktionen  $f_n$  konvergieren lokal gleichmäßig gegen  $f$ .*

*Beweis.*

- (i) Wir zeigen zunächst, dass  $f_n \rightarrow f$  gilt.
- (ii)  $g$  ist als gleichmäßiger Limes stetiger Funktionen in geeigneten Bällen dort und deshalb überall selbst ebenfalls stetig.
- (iii) Definiere

$$\Lambda := \{x \in \Omega: (f_n(x))_n \text{ konvergiert}\}.$$

Es gilt  $x_0 \in \Lambda$  und somit  $\Lambda \neq \emptyset$ .

- (iv) In Theorem 6.3.1 (iii) haben wir gesehen, dass  $f_n \rightrightarrows f$  in einem Ball  $B_\delta$ , in dem  $Df_n \rightrightarrows g$  gilt, gilt, falls  $f_n(y) \rightarrow f(y)$  für ein  $y \in B_\delta$  gilt, selbst wenn die Wortwahl nicht besonders abwechslungsreich ist.

Genauer haben wir dort nur die entsprechende Cauchyfolgenbedingung gezeigt. Mit  $m \rightarrow \infty$  fort dort aber sofort die Existenz von  $f$  und die Behauptung  $f_n \rightrightarrows f$  in  $B_\delta$ .

- (v)  $\Lambda$  ist offen: Sei  $y \in \Lambda$ . Nach (iv) erhalten wir aufgrund der Konvergenz in  $y$ , dass  $f_n \rightrightarrows f$  in einem Ball  $B_\delta(y) \subset \Omega$  gilt, in dem  $Df_n$  gleichmäßig gegen  $g$  konvergiert.
- (vi)  $\Lambda$  ist abgeschlossen: Sei  $x_k \rightarrow \hat{x} \in \Omega$  eine Folge in  $\Omega$  mit  $x_k \in \Lambda$  für alle  $k$ . Sei  $B_\delta(\hat{x}) \subset \Omega$  ein Ball, in dem  $Df_n$  gleichmäßig gegen  $g$  konvergiert. Für große  $k$  gilt  $x_k \in B_\delta(\hat{x})$ . Wegen  $f_n(x_k) \rightarrow f(x_k)$  für solche  $k$  erhalten wir nach (iv)  $f_n \rightrightarrows f$  in  $B_\delta(\hat{x})$  und daher insbesondere Konvergenz in  $\hat{x}$ . Somit ist  $\Lambda$  auch abgeschlossen.
- (vii) Da  $\Omega$  zusammenhängend ist, folgt  $\Lambda = \Omega$  und somit  $f_n \rightarrow f$  in  $\Omega$ .
- (viii) Nun folgt der Rest der Behauptung nach Theorem 6.3.1.  $\square$

*Differentiation von Reihen.*

**Theorem 6.3.3.** *★ Seien  $E, F$  Banachräume. Sei  $\Omega \subset E$  ein Gebiet und  $(f_n)_n \in C^1(\Omega, F)$ . Angenommen, die Reihe  $((f_n(x_0)))_n$  konvergiert für ein  $x_0 \in \Omega$  und zu jedem  $y \in \Omega$  gibt es eine Kugel  $B = B_\delta(y) \subset \Omega$ , in der  $((Df_n))_n$  gleichmäßig*

konvergiert. Dann konvergiert die Reihe  $((f_n))_n$  lokal gleichmäßig in  $\Omega$ . Definiere für  $x \in \Omega$

$$f(x) := \sum_{n \in \mathbb{N}} f_n(x).$$

Dann ist  $f \in C^1(\Omega, F)$  und es gilt

$$Df = \sum_{n \in \mathbb{N}} Df_n.$$

*Beweis.* Wende Theorem 6.3.2 auf die Folge der Partialsummen an.  $\square$

**Beispiel 6.3.4.** Sei  $E$  ein Banachraum. Definiere

$$\exp: L(E) \rightarrow L(E),$$

$$A \mapsto \exp A \equiv e^A := \sum_{n \in \mathbb{N}} \frac{A^n}{n!}.$$

Wir möchten nun die Ableitung der Exponentialfunktion ausrechnen. Dies wird im dritten Semester für lineare gewöhnliche Differentialgleichungen wichtig.

- (i) Definiere  $f_n: L(E) \rightarrow L(E)$  durch  $A \mapsto A^n$ .
- (ii) Wir definieren weiterhin die lineare Abbildung  $\varphi: L(E) \rightarrow L(E)^n$  durch  $A \mapsto (A, \dots, A)$  und erhalten  $D\varphi(A)\langle B \rangle = (B, \dots, B)$ .
- (iii) Definiere schließlich die multilineare Abbildung  $a \in L_n(L(E); L(E))$  durch

$$a(A_1, \dots, A_n) = A_1 \circ \dots \circ A_n.$$

Als multilineare Abbildung ist  $a$  differenzierbar und es gilt

$$Da(A_1, \dots, A_n)\langle B, \dots, B \rangle = \sum_{i=1}^n A_1 \circ \dots \circ A_{i-1} \circ B \circ A_{i+1} \circ \dots \circ A_n.$$

- (iv) Es gilt  $f_n = a \circ \varphi$  und daher nach Kettenregel

$$Df_n(A)\langle B \rangle = Da(\varphi(A))\langle D\varphi(A)\langle B \rangle \rangle = \sum_{i=1}^n \underbrace{A \circ \dots \circ A}_{i-1 \text{ Stück}} \circ B \circ \underbrace{A \circ \dots \circ A}_{n-i \text{ Stück}}.$$

- (v) Es gilt  $\|Df_n(A)\langle B \rangle\| \leq n\|A\|^{n-1}\|B\|$  und somit folgt  $\|Df_n(A)\| \leq n\|A\|^{n-1}$ . Also konvergiert die Reihe  $((\frac{1}{n!}Df_n(\cdot)))_n$  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

$$D \exp(A)\langle B \rangle = \sum_{n \in \mathbb{N}} \frac{1}{n!} Df_n(A)\langle B \rangle.$$

- (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

$$D \exp(A)\langle B \rangle = e^A B = B e^A.$$

**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 und als Anhang in der `tex`-Datei vorhanden.

Wir wiederholen Definition 5.9.2.

**Definition 6.4.1.** ★ Sei  $\Omega \subset \mathbb{R}^n$  offen. Sei  $F$  ein Banachraum und sei  $f: \Omega \rightarrow F$  eine Funktion. Sei  $x_0 = (x_0^i)_{1 \leq i \leq n} \in \Omega$  und sei  $\varepsilon > 0$ , so dass für alle  $z = (z^i)_{1 \leq i \leq n} \in \mathbb{R}^n$  mit  $|x_0^i - z^i| < \varepsilon$  auch  $z \in \Omega$  gilt.

- (i) Dann heißt  $f$  in  $x_0$  in Richtung  $i$ ,  $1 \leq i \leq n$ , **partiell differenzierbar**, falls die Abbildung

$$(x_0^i - \varepsilon, x_0^i + \varepsilon) \ni x^i \mapsto f(x_0^1, \dots, x_0^{i-1}, x^i, x_0^{i+1}, \dots, x_0^n) \equiv f(\hat{x}_0^i, x^i)$$

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  $i$ ,  $1 \leq i \leq n$ , partiell differenzierbar ist.
- (iii) Ist  $f$  in allen Punkten  $x \in \Omega$  partiell differenzierbar, so ist  $D_i f$  eine Abbildung  $D_i f: \Omega \rightarrow F$  und  $f$  heißt in  $\Omega$  **partiell differenzierbar**. Sind alle diese Abbildungen  $D_i f$ ,  $1 \leq i \leq n$ , stetig, so heißt  $f$  **stetig partiell differenzierbar**.

**Bemerkung 6.4.2.** ★

- (i) Die obige Bedingung an  $z$  ist äquivalent zu

$$\prod_{i=1}^n (x_0^i - \varepsilon, x_0^i + \varepsilon) \subset \Omega.$$

- (ii) Die Definition von partieller Differenzierbarkeit ist unabhängig von der Wahl von  $\varepsilon > 0$ . Auch genügt es, dass die Verbindungsstrecke

$$[(\hat{x}_0^i, x_0^i - \varepsilon), (\hat{x}_0^i, x_0^i + \varepsilon)] \subset \Omega$$

erfüllt.

*Beweis.* Übung. □

- (iii) Die Existenz eines solchen  $\varepsilon > 0$  folgt aus der Offenheit von  $\Omega$ , unabhängig von der Wahl einer (äquivalenten) Norm auf  $\mathbb{R}^n$ .
- (iv) Wir haben oben  $L(\mathbb{R}, F)$  und  $F$  vermöge der Abbildung

$$L(\mathbb{R}, F) \ni \varphi \mapsto \varphi(1) \in F$$

identifiziert und wollen dies später ohne weitere Erwähnung ebenfalls machen. Diese Abbildung ist ein Vektorraumisomorphismus und erhält die Norm:  $\|\varphi\|_{L(\mathbb{R}, F)} = \|\varphi(1)\|_F$ .

Man bezeichnet eine solche Abbildung auch als isometrischen Banachraumisomorphismus. (Bei einem Banachraumisomorphismus  $\Phi: E \rightarrow F$  wäre auch  $\frac{1}{c}\|x\|_E \leq \|\Phi(x)\|_F \leq c\|x\|_E$  für ein  $c > 0$  und beliebige  $x \in E$  möglich.)

- (v) Zur Abgrenzung gegenüber partiellen Ableitungen bezeichnen wir die übliche Ableitung gelegentlich als **totale** Ableitung.
- (vi) Wir werden zunächst eher die Bezeichnung  $D_i f$  verwenden, da ein Ausdruck wie  $\frac{\partial f}{\partial x^i}(x^2, x^1)$  oder  $f_1(x^2, x^1)$  leichter zu Missverständnissen führt. Später werden wir jedoch diese Schreibweisen bevorzugen.

**Proposition 6.4.3.** Sei  $\Omega \subset \mathbb{R}^n$  offen. Sei  $F$  ein Banachraum und sei  $f: \Omega \rightarrow F$  in  $x_0 \in \Omega$  differenzierbar. Dann ist  $f$  in  $x_0$  auch partiell differenzierbar.

Ist  $\chi_i: \mathbb{R} \hookrightarrow \mathbb{R}^n$  mit  $x \mapsto (0, \dots, 0, x, 0, \dots, 0)$  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  $(\hat{x}_0^i, x^i) \equiv (x_0^1, \dots, x_0^{i-1}, x^i, x_0^{i+1}, \dots, x_0^n) \in \Omega$  für alle  $x^i \in (x_0^i - \varepsilon, x_0^i + \varepsilon)$  gilt. Wir definieren die Abbildung

$$\varphi: (x_0^i - \varepsilon, x_0^i + \varepsilon) \rightarrow \Omega,$$

$$x^i \mapsto (\hat{x}_0^i, x^i).$$

Da  $\varphi$  affin linear ist, folgt

$$D\varphi(x^i)\langle t \rangle = (0, \dots, 0, t, 0, \dots, 0) = \chi_i \langle t \rangle$$

für alle  $x^i \in (x_0^i - \varepsilon, x_0^i + \varepsilon)$  und alle  $t \in \mathbb{R}$ , also  $D\varphi(x^i) = \chi_i$ . Nun gilt

$$f(\hat{x}_0^i, x^i) = f \circ \varphi(x^i).$$

Auf der linken Seite steht hier die Funktion aus der Definition der partiellen Ableitung. Wenden wir die Kettenregel auf die rechte Seite an, so erhalten wir

$$D_i f(x_0) = Df(\varphi(x_0^i)) \langle D\varphi(x_0^i) \rangle = Df(x_0) \circ \chi_i$$

und daher gerade die Behauptung.  $\square$

Die Ableitung einer Funktion ergibt sich aus den partiellen Ableitungen.

**Korollar 6.4.4.** Sei  $\Omega \subset \mathbb{R}^n$  offen und sei  $F$  ein Banachraum. Sei  $f: \Omega \rightarrow F$  in  $x_0 \in \Omega$  differenzierbar. Dann folgt

$$Df(x_0)\langle u \rangle = \sum_{i=1}^n D_i f(x_0) \langle u^i \rangle \quad \text{für } u = (u^i)_{1 \leq i \leq n} \in \mathbb{R}^n,$$

d. h. es gilt

$$Df(x_0) = \sum_{i=1}^n D_i f(x_0) \circ \pi_i,$$

wobei  $\pi_i: \mathbb{R}^n \rightarrow \mathbb{R}$  die kanonische Projektion auf die  $i$ -te Komponente bezeichnet.

*Beweis.* Es gilt nach Proposition 6.4.3 in der dortigen Notation und aufgrund der Linearität

$$\sum_{i=1}^n D_i f(x_0) \langle u^i \rangle = \sum_{i=1}^n Df(x_0) \langle \chi_i \langle u^i \rangle \rangle = Df(x_0) \left\langle \sum_{i=1}^n \chi_i \langle u^i \rangle \right\rangle = Df(x_0) \langle u \rangle$$

wie behauptet.  $\square$

**Bemerkung 6.4.5.** Sei  $\Omega \subset \mathbb{R}^n$  offen und  $f: \Omega \rightarrow \mathbb{R}$  in  $x_0 \in \Omega$  differenzierbar, so haben wir  $Df(x_0)$  mit einem Vektor in  $\mathbb{R}^n$  identifiziert. Nun gilt

$$\nabla f(x_0) = (f_1, \dots, f_n)^T.$$

Umgekehrt braucht eine partiell differenzierbare Funktion nicht differenzierbar zu sein.

**Beispiel 6.4.6.** Definiere  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  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 6.4.7.** Sei  $\Omega \subset \mathbb{R}^n$  offen und sei  $F$  ein Banachraum. Dann ist  $f: \Omega \rightarrow F$  genau dann in  $\Omega$  stetig differenzierbar, wenn  $f$  in  $\Omega$  stetig partiell differenzierbar ist.

*Beweis.* Wegen Proposition 6.4.3 genügt es zu zeigen, dass  $f$  in  $\Omega$  stetig differenzierbar ist, falls  $f$  in  $\Omega$  stetig partiell differenzierbar ist, dass also „ $\Leftarrow$ “ richtig ist.

Ist  $f$  differenzierbar, so können wir nach Korollar 6.4.4 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  $\Omega$  nachweisen. Sei dazu  $x_0 \in \Omega$  beliebig und  $\Omega$  ohne Einschränkung konvex. Wir wollen zeigen, dass

$$f(x) = f(x_0) + \sum_{i=1}^n D_i f(x_0) \langle x^i - x_0^i \rangle + o(\|x - x_0\|)$$

gilt.

Zur besseren Übersicht führen wir den Rest des Beweises nur im Falle  $n = 2$  durch. Es gilt

$$\begin{aligned} f(x^1, x^2) - f(x_0^1, x_0^2) &= f(x^1, x^2) - f(x_0^1, x^2) + f(x_0^1, x^2) - f(x_0^1, x_0^2) \\ (6.1) \quad &= D_1 f(x_0^1, x^2) \langle x^1 - x_0^1 \rangle + o_{x^2}(\|x^1 - x_0^1\|) \\ &\quad + D_2 f(x_0^1, x_0^2) \langle x^2 - x_0^2 \rangle + o(\|x^2 - x_0^2\|). \end{aligned}$$

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

$$\begin{aligned} \|o_{x^2}(\|x^1 - x_0^1\|)\| &= \|f(x^1, x^2) - f(x_0^1, x^2) - D_1 f(x_0^1, x^2) \langle x^1 - x_0^1 \rangle\| \\ &\leq \sup_{\xi \in [x^1, x_0^1]} \|D_1 f(\xi, x^2) - D_1 f(x_0^1, x^2)\| \cdot \underbrace{\|x^1 - x_0^1\|}_{\leq \|x - x_0\|}. \end{aligned}$$

Aufgrund der Stetigkeit von  $D_1 f$  sehen wir, dass der Term  $o_{x^2}(\|x^1 - x_0^1\|)$  ein  $o(\|x - x_0\|)$ -Term ist. Insbesondere wird also der erste Faktor in einer kleinen Umgebung von  $(x_0^1, x_0^2)$  unabhängig von  $x^2$  klein.

Nun zum ersten Term der rechten Seite in (6.1). Es gilt

$$\begin{aligned} D_1 f(x_0^1, x^2) \langle x^1 - x_0^1 \rangle &= D_1 f(x_0^1, x_0^2) \langle x^1 - x_0^1 \rangle \\ &\quad + (D_1 f(x_0^1, x^2) - D_1 f(x_0^1, x_0^2)) \langle x^1 - x_0^1 \rangle. \end{aligned}$$

Da  $D_1 f$  stetig ist, ist auch diese Differenz in  $o(\|x - x_0\|)$ .  $\square$

In Kombination mit der Kettenregel erhalten wir

**Theorem 6.4.8.** Seien  $\Omega \subset \mathbb{R}^n$  und  $\Omega' \subset \mathbb{R}^m$  offen. Sei  $F$  ein Banachraum. Seien  $f: \Omega \rightarrow F$  und  $g = (g^i)_i: \Omega' \rightarrow \Omega \subset \mathbb{R}^n$  differenzierbar:

$$\begin{array}{ccc} \Omega' & \xrightarrow{h} & \Omega \xrightarrow{f} F \\ \cap & \searrow g & \cap \\ \mathbb{R}^m & & \mathbb{R}^n \end{array}$$

Dann ist auch  $h = f \circ g$  differenzierbar und es gilt

$$Dh = \sum_{i=1}^n D_i f(g) \langle Dg^i \rangle,$$

d. h. mit Argumenten  $x' \in \Omega'$  und  $u \in \mathbb{R}^m$  erhalten wir

$$Dh(x') \langle u \rangle = \sum_{i=1}^n D_i f(g(x')) \langle Dg^i(x') \langle u \rangle \rangle$$

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

$$\begin{aligned} Dh(x)\langle u \rangle &= Df(g(x))\langle Dg(x)\langle u \rangle \rangle = Df(g(x)) \langle (Dg^1(x), \dots, Dg^n(x)) \langle u \rangle \rangle \\ &= \sum_{i=1}^n D_i f(g(x)) \langle \pi_i(Dg^1(x)\langle u \rangle, \dots, Dg^n(x)\langle u \rangle) \rangle \\ &= \sum_{i=1}^n D_i f(g(x)) \langle Dg^i(x)\langle u \rangle \rangle \end{aligned}$$

wie behauptet.  $\square$

**Bemerkung 6.4.9** (Jacobimatrix). Seien  $\Omega \subset \mathbb{R}^n$  offen und  $f: \Omega \rightarrow \mathbb{R}^m$  differenzierbar. Nach Korollar 6.4.4 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  $\sum_{i=1}^n \frac{\partial f^j}{\partial x^i}(x_0) u^i$ .

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  $\left( \frac{\partial f^j}{\partial x^i}(x_0) \right)_{\substack{1 \leq j \leq m \\ 1 \leq i \leq n}}$  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  $u = (u^i)_i \in \mathbb{R}^n$ .

Ist  $n = m$ , so heißt  $\det J_f(x_0) \equiv \det \frac{\partial f^j}{\partial x^i} \equiv \det \frac{df}{dx}$  **Funktionaldeterminante**.

Aus Theorem 6.4.8 erhalten wir für die Jacobimatrizen

$$J_h(x) = J_f(g(x)) \cdot J_g(x).$$

**Bemerkung 6.4.10** (Einsteinsche Summenkonvention).  $\star$  Physiker summieren gerne über Indices, die in einer Formel doppelt einmal unten und einmal oben auftauchen. Unter Benutzung dieser **Einsteinschen Summenkonvention** erhalten wir

$$Df(x_0)\langle u \rangle = \frac{\partial f}{\partial x^i}(x_0) u^i \equiv f_i(x_0) u^i.$$

Daher bezeichnen wir die  $j$ -te Komponente von  $f$  mit  $f^j$  und die  $i$ -te partielle Ableitung von  $f$  mit  $f_i$ . Wir werden die Einsteinsche Summenkonvention auch ab Vorlesungen über Differentialgeometrie oder Partielle Differentialgleichungen (meinen Spezialgebieten) verwenden.

Oben und unten stehende Indices erlauben es auch besser, zwischen Elementen in Vektorräumen und Dualräumen zu unterscheiden. Dementsprechend ist es sinnvoll, die Einträge von Matrizen mit  $a_j^i$  zu bezeichnen. Details dazu finden sich in [4].

Wir zeigen einen eindimensionalen Hebbarkeitssatz.

**Lemma 6.4.11.** *Sei  $\Omega \subset \mathbb{R}$  offen mit  $0 \in \Omega$ . Seien  $f, g: \Omega \rightarrow \mathbb{R}$  stetig und  $f$  in  $\Omega \setminus \{0\}$  stetig differenzierbar mit  $f' = g$  in  $\Omega \setminus \{0\}$ . Dann ist  $f$  in ganz  $\Omega$  stetig differenzierbar und es gilt dort  $f' = g$ .*

*Beweis.* Wir dürfen ohne Einschränkung annehmen, dass  $\Omega$  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

$$|f(x) - f(0) - g(0)x| \leq \varepsilon \cdot |x|$$

für alle  $|x| < \delta$  gibt.

Definiere dazu  $\varphi: [0, 1] \rightarrow \mathbb{R}$  durch

$$\varphi(t) := f(x) - f(tx) - g(0)(1-t)x.$$

Dann ist  $\varphi$  auf  $[0, 1]$  stetig und in  $(0, 1)$  differenzierbar. Der Mittelwertsatz ist also anwendbar. Es gelten

$$\begin{aligned} \varphi(1) &= 0, \\ |\varphi(0)| &= |f(x) - f(0) - g(0)x| \stackrel{!}{\leq} \varepsilon \cdot |x|, \\ |\varphi(0)| &= |\varphi(0) - \varphi(1)| \\ &\leq \sup_{t \in (0,1)} |\varphi'(t)| \cdot 1 \\ &= \sup_{t \in (0,1)} |-f'(tx)x - g(0)(-1)x| \\ &= \sup_{t \in (0,1)} |f'(tx) - g(0)| \cdot |x| \\ &\leq \sup_{0 < |y| < \delta} |f'(y) - g(0)| \cdot |x| \quad \text{für } |x| \leq \delta, \\ &= \sup_{0 < |y| < \delta} |g(y) - g(0)| \cdot |x|. \end{aligned}$$

Die Behauptung folgt nun aus der Stetigkeit von  $g$ . □

**Lemma 6.4.12.** *Sei  $\Omega \subset \mathbb{R}^n$  offen. Seien  $f: \Omega \rightarrow \mathbb{R}$  und  $g: \Omega \rightarrow \mathbb{R}^n$  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  $\Omega$  stetig differenzierbar und es gilt dort  $Df = g$ .*

*Beweis.* Nach Theorem 6.4.7 genügt es zu zeigen, dass  $f$  in  $\Omega$  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 6.4.11. □

Dieses Resultat überträgt sich wie folgt auf die erst im nächsten Kapitel definierten höheren Ableitungen.

**Korollar 6.4.13.** *★ Sei  $\Omega \subset \mathbb{R}^n$  offen. Sei  $H \subset \mathbb{R}^n$  eine Hyperebene. Sei  $f: \Omega \rightarrow \mathbb{R}$  in  $\Omega \setminus H$  eine  $k$ -mal stetig differenzierbare Funktion und seien sämtliche Ableitungen bis zur Ordnung  $k$  stetig auf ganz  $\Omega$  fortsetzbar. Dann ist auch  $f$  in ganz  $\Omega$   $k$ -mal stetig differenzierbar.*

*Beweis.* Induktion nach der Ordnung der Ableitung. □

**Bemerkung 6.4.14.** *★ Lemma 6.4.11 lässt sich nicht problemlos auf Funktionen übertragen, die auf dem Komplement einer abgeschlossenen Menge  $A$  mit  $\text{int } A = \emptyset$  definiert sind. Die Cantorsche Treppenfunktion ist ein Gegenbeispiel dafür.*

*Richtungsableitung.*

**Definition 6.4.15.** Seien  $E, F$  Banachräume und  $\Omega \subset E$  offen. Sei  $e \in E$  mit  $\|e\| = 1$ . Sei  $f: \Omega \rightarrow F$  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 6.4.16.**

- (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 6.4.6 ist im Ursprung in jede Richtung differenzierbar, jedoch nicht differenzierbar.
- (v) ★ Einen Vektor der Länge 1 nennen wir **Richtung** oder **Einheitsvektor**.

**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  $L(E, F)$ ,  $L(E, L(E, F))$ , ... mit der Operatornorm wieder Banachräume sind.

**Definition 6.5.1.** Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow F$  eine Abbildung. Sei  $x_0 \in \Omega$  beliebig.

- (i) Wir definieren  $D^0 f := f: \Omega \rightarrow F$  als die 0-te Ableitung von  $f$ . Ist  $f$  in  $U \subset \Omega$  differenzierbar, so definieren wir die **erste Ableitung** von  $f$  durch  $D^1 f := Df: U \rightarrow L(E, F)$ .
- (ii) Ist  $Df$  in einer Umgebung  $U$  von  $x_0$  definiert und die Abbildung  $Df: U \rightarrow L(E, F)$  in  $x_0$  differenzierbar, so heißt  $D(Df)(x_0) \equiv D^2 f(x_0): E \rightarrow L(E, F)$ , also  $D^2 f(x_0) \in L(E, L(E, F))$  **zweite Ableitung** von  $f$  in  $x_0$ .
- (iii) ★ Ist  $D^2 f: U \rightarrow L(E, L(E, F))$  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  $D^k f: U \rightarrow L(E, \dots, L(E, F) \dots)$  in  $x_0$  differenzierbar, so heißt die Ableitung dieser Funktion,  $D(D^k f)(x_0) \equiv D^{k+1} f(x_0)$ , 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  $\Omega$   **$k$ -mal differenzierbar**.
- (vi) ★ Existiert die  $k$ -te Ableitung von  $f$  in allen  $x \in \Omega$  und ist sie in  $\Omega$  stetig, so heißt  $f$  in  $\Omega$   **$k$ -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  $\bar{\Omega}$  stetig und beschränkt fortsetzbar, so ist  $f$  von der Klasse  $C^k(\bar{\Omega}, F)$ ,  $f \in C^k(\bar{\Omega}, F)$ .
- (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 6.5.2.**

- (i) ★ Da  $L(E, F)$  selbst wieder ein Banachraum ist, können wir die Definition von Differenzierbarkeit auf die Abbildung  $Df: U \rightarrow L(E, F)$  anwenden.

- (ii) Aus der linearen Algebra, siehe auch Bemerkung 6.1.6, kennen wir den (norm-treuen) 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 6.5.3.** *Sei  $\Omega \subset E$  im Banachraum  $E$  offen. Sei  $f: \Omega \rightarrow F$  in  $x_0$  zweimal differenzierbar. Dann ist, vermöge der Identifikation*

$$L(E, L(E, F)) \cong L_2(E; F),$$

$D^2 f(x_0) \in L_2(E; F)$  und es gilt

$$D^2 f(x_0)(u, v) \equiv D^2 f(x_0)\langle u, v \rangle = (D^2 f(x_0)\langle u \rangle) \langle v \rangle.$$

**Bemerkung 6.5.4.** ★

- (i) Hier stehen links die bilinearen Abbildungen, bei denen wir auch wieder  $\langle \cdot, \cdot \rangle$  benutzen werden um die Linearität in den angegebenen Argumenten zu verdeutlichen.
- (ii) Wir wollen künftig stets  $D^2 f(x_0)$  für das Element in  $L(E, L(E, F))$  oder  $L_2(E; F)$  schreiben.
- (iii) Entsprechend schreiben wir stets  $D^k f(x_0)$  für das entsprechende Element in

$$L(E, \dots, L(E, F) \dots)$$

oder in  $L_k(E; F)$ . Dabei benutzen wir induktiv, dass

$$L(E, L_k(E; F)) \cong L_{k+1}(E; F)$$

gilt.

*Beweis.* Übung. □

- (iv) Ist  $f$  in  $\Omega$  zunächst  $m$ -mal differenzierbar und  $D^m f$  in  $\Omega$  noch  $n$ -mal differenzierbar, so ist  $f$  in  $\Omega$  mindestens  $(m+n)$ -mal differenzierbar.

*Beweis.* Induktion. Übung □

**Definition 6.5.5** ( $C^k$ -Normen). Seien  $E, F$  Banachräume und  $\Omega \subset E$  offen. Dann definieren wir auf  $C^k(\overline{\Omega}, F)$  eine Norm, die  $C^k$ -Norm, durch

$$\|f\|_{C^k(\Omega, F)} := \sum_{i=0}^k \|D^i f\|_{C^0(\Omega, L_i(E; F))}$$

mit  $L_0(E, F) \equiv F$ .

Genauso wie im eindimensionalen Fall sehen wir, dass die Räume  $C^k(\overline{\Omega}, F)$  Banachräume sind.

**Theorem 6.5.6.** *Seien  $E, F$  Banachräume und sei  $\Omega \subset E$  offen. Sei  $k \in \mathbb{N}$ . Dann ist  $C^k(\overline{\Omega}, F)$  mit der  $C^k$ -Norm ein Banachraum.*

*Beweis.* Übung. □

**Lemma 6.5.7.** Seien  $E, F$  Banachräume und  $\Omega \subset E$  offen. Sei  $f: \Omega \rightarrow F$  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  $\varphi: U \rightarrow F$  durch  $\varphi(x) = Df(x)\langle v \rangle$ . Dann ist  $\varphi$  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  $\varphi$  als Komposition der linearen Abbildung

$$\Phi: L(E, F) \rightarrow F \quad \text{mit} \quad A \mapsto Av$$

und der Abbildung  $Df: U \rightarrow L(E, F): \varphi = \Phi \circ Df$ . Die Abbildung  $\Phi$  ist wegen

$$\|Av\|_F \leq \|A\|_{L(E, F)} \cdot \|v\|_E,$$

wobei nun anders als sonst  $\|v\|_E$  die Konstante liefert, stetig. 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  $\varphi$  somit in  $x_0$  differenzierbar und es gilt  $D\varphi = D\Phi(Df) \circ D^2f$  und in expliziter Form

$$\begin{aligned} D\varphi(x_0)\langle u \rangle &= D\Phi(Df(x_0))\langle D(Df(\cdot)\langle \cdot \rangle)(x_0)\langle u \rangle \rangle \\ &= D\Phi(Df(x_0))\langle D^2f(x_0)\langle \cdot \rangle\langle u \rangle \rangle = D^2f(x_0)\langle v, u \rangle, \end{aligned}$$

wobei  $Df(\cdot)\langle \cdot \rangle$  eine Funktion  $x \mapsto L(E, F)$  bezeichnet.  $\square$

Das folgende Resultat ist eine Variante des noch folgenden Satzes von Schwarz. Beachte, dass hier die Existenz der zweiten Ableitung in  $x_0$  und nicht nur die partiellen Ableitungen gefordert wird.

**Theorem 6.5.8.** Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und sei  $f: \Omega \rightarrow F$  in  $x_0 \in \Omega$  zweimal differenzierbar. Dann ist die bilineare Abbildung  $D^2f(x_0)$  symmetrisch, d. h. es gilt für alle  $u, v \in E$

$$D^2f(x_0)\langle u, v \rangle = D^2f(x_0)\langle v, u \rangle.$$

*Beweis.* Wir behaupten, dass

$$(6.2) \quad f(x_0 + u + v) - f(x_0 + u) - f(x_0 + v) + f(x_0) = D^2f(x_0)\langle u, v \rangle + o((\|u\| + \|v\|)^2)$$

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

$$\|D^2f(x_0)\langle u, v \rangle - D^2f(x_0)\langle v, u \rangle\| = o((\|u\| + \|v\|)^2)$$

und für alle Einheitsvektoren  $\bar{u}, \bar{v} \in E$  und alle  $|t| < r$  erhalten wir daraus

$$t^2 \|D^2f(x_0)\langle \bar{u}, \bar{v} \rangle - D^2f(x_0)\langle \bar{v}, \bar{u} \rangle\| = o(t^2).$$

Wir dividieren nun durch  $t^2 \neq 0$  und erhalten im Grenzwert  $t \rightarrow 0$  die Behauptung.

Somit ist noch (6.2) zu zeigen. Definiere für  $0 \leq t \leq 1$  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

$$\|\varphi(1) - \varphi(0) - \varphi'(0)\| \leq \sup_{0 \leq t \leq 1} \|\varphi'(t) - \varphi'(0)\|.$$

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

$$\begin{aligned} \varphi'(t) &= \{Df(x_0 + u + tv) - Df(x_0) - [Df(x_0 + tv) - Df(x_0)]\} \langle v \rangle \\ &= \{D^2f(x_0)\langle u + tv \rangle + o(\|u + tv\|) - D^2f(x_0)\langle tv \rangle - o(\|tv\|)\} \langle v \rangle \\ &= D^2f(x_0)\langle v \rangle \langle u \rangle + o(\|u + tv\|)\langle v \rangle - o(\|tv\|)\langle v \rangle \end{aligned}$$

und somit

$$\|\varphi'(t) - D^2f(x_0)\langle v \rangle \langle u \rangle\| = o(\|u + tv\|)\|v\| + o(\|tv\|)\|v\|.$$

(Zur hier angegebenen Reihenfolge der Argumente: Bei der ersten Ableitung gibt es für das lineare Argument noch eine Leerstelle, in die  $v$  nachher eingesetzt wird. Daher kommt das zunächst ausgeklammerte  $\langle v \rangle$ , obwohl es nun ganz hinten steht, an die erste Stelle. Die zweite Ableitung ist die Ableitung bevor  $v$  eingesetzt wird. Daher wird  $v$  auch bei der Anwendung der charakterisierenden Gleichung für die zweite Ableitung nicht mitberücksichtigt, sondern „wartet“ außen darauf, später als Argument eingesetzt zu werden.)

(Alternativ könnte man auch mit Lemma 6.5.7 argumentieren. Dabei wird allerdings nicht klar, warum  $o(\dots)$  nicht auch noch von  $v$ , was ja nun Bestandteil der betrachteten Funktion ist, abhängt.)

Wie man sich beispielsweise mit Hilfe der  $\varepsilon$ -Notation leicht überzeugt, sind beide Fehlerterme in  $o(\|u\| + \|v\|)^2$ . Hieraus folgt einerseits

$$\sup_{0 \leq t \leq 1} \|\varphi'(t) - \varphi'(0)\| \leq o(\|u\| + \|v\|)^2,$$

andererseits können wir den Ausdruck für  $\varphi'(t)$ , ausgewertet an der Stelle  $t = 0$ , auch in der Folgerung aus dem Mittelwertsatz einsetzen und erhalten

$$\varphi(1) - \varphi(0) = D^2f(x_0)\langle v, u \rangle + o(\|u\| + \|v\|)^2.$$

Nach Definition von  $\varphi$  ist dies aber gerade die Behauptung.  $\square$

Auch für zweite Ableitungen bekommen wir einen Zusammenhang zwischen  $D^2f$  und zweiten partiellen Ableitungen.

**Proposition 6.5.9.** *Sei  $\Omega \subset \mathbb{R}^n$  offen und sei  $F$  ein Banachraum. Sei  $f: \Omega \rightarrow F$  in  $x_0 \in \Omega$  zweimal differenzierbar und seien  $1 \leq i, j \leq n$ . Dann gelten*

(i)

$$D_i D_j f(x_0) = D^2f(x_0)\langle e_j, e_i \rangle$$

(ii) und

$$D_i D_j f(x_0) = D_j D_i f(x_0),$$

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 6.4.3 bereits  $D_j f(\cdot) = Df(\cdot) \circ \chi_j$ ,  $1 \leq j \leq n$ , mit der Standardeinbettung  $\mathbb{R} \rightarrow \mathbb{R}^n$  in die  $j$ -te Komponente gezeigt. Daraus erhalten wir  $D_j f(\cdot) = Df(\cdot)\langle e_j \rangle$

durch Anwenden auf 1. Wir differenzieren dies nochmals und benutzen dabei nochmals Proposition 6.4.3 und Lemma 6.5.7

$$D_i D_j f(x_0) = D(Df(\cdot)\langle e_j \rangle)(x_0)\langle e_i \rangle = D^2 f(x_0)\langle e_j, e_i \rangle.$$

(ii) Dies folgt nun unmittelbar aus Theorem 6.5.8. Es gilt nämlich

$$D_i D_j f(x_0) = D^2 f(x_0)\langle e_j, e_i \rangle = D^2 f(x_0)\langle e_i, e_j \rangle = D_j D_i f(x_0). \quad \square$$

Existieren die zweiten partiellen Ableitungen und sind stetig, so existieren auch die zweiten Ableitungen. Als Vorbereitung dafür zeigen wir

**Lemma 6.5.10.** *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  $\chi_i: \mathbb{R} \hookrightarrow \mathbb{R}^n$  die kanonische Einbettung in die  $i$ -te Komponente und  $\pi_i: \mathbb{R}^n \rightarrow \mathbb{R}$  die Projektion auf den  $i$ -ten Faktor, so ist ein möglicher solcher Banachraumisomorphismus durch  $\Phi: A \mapsto (A_i)_{1 \leq i \leq n}$  mit  $A_i = A \circ \chi_i$  und die Inverse dazu ist durch  $A = \sum_{i=1}^n A_i \circ \pi_i$  gegeben.*

*Beweis.* Direktes Nachrechnen. Beachte, dass wir nicht behaupten, dass der Banachraumisomorphismus normerhaltend ist.  $\square$

**Bemerkung 6.5.11.**  $\star$  Lemma 6.5.10 lässt sich sehr anschaulich mit Vektoren darstellen. Eine Abbildung  $A \in L(\mathbb{R}^n, F)$  entspricht einem Zeilenvektor, in dem die Elemente  $A\langle e_i \rangle \in F$  stehen. Dann operiert  $A$  als Vektormultiplikation auf  $x = (x^i)_{1 \leq i \leq n} \in \mathbb{R}^n$ :

$$A\langle x \rangle = (A\langle e_1 \rangle \quad \dots \quad A\langle e_n \rangle) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

Nun ist  $\Phi(A) \in F^n$  gerade dieser Zeilenvektor.

**Theorem 6.5.12.** *Sei  $\Omega \subset \mathbb{R}^n$  offen und sei  $F$  ein Banachraum. Sei  $f: \Omega \rightarrow F$  eine Abbildung. Dann ist  $f$  genau dann zweimal stetig in  $\Omega$  differenzierbar, wenn  $f$  in  $\Omega$  zweimal stetig partiell differenzierbar ist.*

*Beweis.* Nach Theorem 6.4.7 (stetige und stetige partielle Differenzierbarkeit sind äquivalent) ist die Behauptung für die ersten Ableitungen bekannt.

„ $\Rightarrow$ “: Sei  $f$  zunächst zweimal stetig differenzierbar. Nach Proposition 6.5.9 gilt  $D_i D_j f(x_0)\langle \cdot, \cdot \rangle = D^2 f(x_0)\langle e_j, e_i \rangle$  und wir erhalten die Behauptung für die zweiten Ableitungen.

„ $\Leftarrow$ “: Sei nun  $f$  umgekehrt zweimal stetig partiell differenzierbar,  $D_i f$  und  $D_j D_i f$ ,  $1 \leq i, j \leq n$  existieren also und sind stetig. Nach Theorem 6.4.7 sind somit  $f$  und  $D_i f$  in  $\Omega$  stetig differenzierbar.

Wir verwenden nun  $\Phi$  aus Lemma 6.5.10 und erhalten eine Abbildung  $\Phi \circ Df: \Omega \rightarrow F^n$ . Nach Proposition 6.4.3 sind die Komponenten dieser Abbildung gerade die partiellen Ableitungen  $D_i f$ .

Nach Voraussetzung sind die Funktionen  $D_i f$ , die ja gerade die Komponenten der Abbildung  $\Phi \circ Df$  sind, in  $\Omega$  differenzierbar. Somit ist nach Proposition 6.1.13 auch die Abbildung  $\Phi \circ Df$  in  $\Omega$  differenzierbar. Wir verknüpfen diese Funktion noch mit der linearen Abbildung  $\Phi^{-1}$  und erhalten, dass  $Df = \Phi^{-1} \circ \Phi \circ Df$  in  $\Omega$  differenzierbar ist.

Zur Stetigkeit der Abbildung  $x \mapsto D^2 f(x)$ : Seien  $u = (u^i)_{1 \leq i \leq n}, v = (v^j)_{1 \leq j \leq n} \in \mathbb{R}^n$ . Es gilt

$$\begin{aligned} D^2 f(x) \langle u, v \rangle &= D^2 f(x) \left\langle \sum_{i=1}^n u^i e_i, \sum_{j=1}^n v^j e_j \right\rangle \\ &= \sum_{i,j=1}^n u^i \cdot v^j \cdot D^2 f(x) \langle e_i, e_j \rangle = \sum_{i,j=1}^n u^i \cdot v^j \cdot D_j D_i f(x) \end{aligned}$$

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.  $\square$

**Korollar 6.5.13** (Satz von Schwarz). *Sei  $\Omega \subset \mathbb{R}^n$  offen und sei  $F$  ein Banachraum. Sei  $f: \Omega \rightarrow F$  zweimal stetig partiell differenzierbar, so sind die zweiten partiellen Ableitungen symmetrisch, d. h. für alle  $1 \leq i, j \leq n$  und alle  $x \in \Omega$  gilt*

$$D_i D_j f(x) = D_j D_i f(x).$$

*Beweis.* Nach Theorem 6.5.12 ist  $f$  in  $\Omega$  zweimal differenzierbar. Nun können wir Proposition 6.5.9 anwenden und erhalten die behauptete Vertauschbarkeit der partiellen Ableitungen.  $\square$

**Bemerkung 6.5.14.**

- (i) Sei  $\Omega \subset \mathbb{R}^n$  offen und sei  $f: \Omega \rightarrow \mathbb{R}$  zweimal differenzierbar. Dann erhalten wir für alle  $x \in \Omega$  und alle  $u, v \in \mathbb{R}^n$  nach Proposition 6.5.9 und der Linearität

$$D^2 f(x) \langle u, v \rangle = \sum_{i,j=1}^n D_i D_j f(x) u^j v^i.$$

- (ii) Wie in der linearen Algebra können wir auch hier die symmetrische Bilinearform  $D^2 f(x)$  durch eine Matrix darstellen, nämlich durch  $(D_i D_j f(x))_{1 \leq i, j \leq n}$ , die **Hessematrix** oder **Hessesche** von  $f$ .

Viele Eigenschaften, die wir für zweite Ableitungen gezeigt haben, übertragen sich per Induktion auf höhere Ableitungen. Wir erwähnen hier nur die entsprechenden Resultate.

**Bemerkung 6.5.15.** ★ Seien  $n, m \in \mathbb{N}_{>0}$ ,  $F$  ein Produktraum mit  $F = \prod_{j=1}^m F_j$  für Banachräume  $F_j$ ,  $\Omega \subset \mathbb{R}^n \equiv E$  offen und  $f: \Omega \rightarrow F$  eine Abbildung. Häufig machen wir von der Produktgestalt keinen Gebrauch; dies entspricht den Fällen  $n = 1$  oder  $m = 1$ .

- (i) Ist  $f$  eine  $k$ -mal differenzierbare Abbildung, so können wir  $D^k f(x_0)$  als Element in  $L_k(E; F)$  auffassen.  
(ii) Ist  $f$  eine  $k$ -mal differenzierbare Abbildung und  $D^k f$  eine  $l$ -mal differenzierbare Abbildung, so ist  $f$  eine  $k + l$ -mal differenzierbare Abbildung.  
(iii) Seien  $f$  eine  $k$ -mal differenzierbare Abbildung und  $u_1, \dots, u_{k-1} \in E$ . Dann ist die Abbildung  $\varphi: \Omega \rightarrow F$  mit

$$\varphi(x) = D^{k-1} f(x) \langle u_1, \dots, u_{k-1} \rangle$$

in  $\Omega$  differenzierbar und es gilt

$$D\varphi(x) \langle \cdot \rangle = D^k f(x) \langle u_1, \dots, u_{k-1}, \cdot \rangle.$$

Ist  $f$  sogar  $k+1$ -mal differenzierbar, so gilt eine entsprechende Aussage auch für die zweite Ableitung von  $\varphi, \dots$

- (iv) Sei  $f$  eine im Punkt  $x_0 \in \Omega$  mindestens  $k$ -mal differenzierbare Abbildung. So ist

$$E^k \ni (u_1, \dots, u_k) \mapsto D^k f(x_0) \langle u_1, \dots, u_k \rangle$$

eine symmetrische Abbildung, d. h. es gilt

$$D^k f(x_0) \langle u_1, \dots, u_k \rangle = D^k f(x_0) \langle u_{\sigma(1)}, \dots, u_{\sigma(k)} \rangle$$

für alle  $u_1, \dots, u_k \in E$  und alle  $\sigma \in S(k)$ , der symmetrischen Gruppe, die die ersten  $k$  natürlichen Zahlen beliebig permutiert.

- (v) Sei  $f$  in  $x_0 \in \Omega$  eine  $k$ -mal differenzierbare Abbildung. Dann ist  $f$  in  $x_0$  auch  $k$ -mal partiell differenzierbar und es gilt für alle  $i_j$  mit  $1 \leq i_j \leq n$ ,  $1 \leq j \leq k$ ,

$$D_{i_k} \dots D_{i_1} f(x_0) = D^k f(x_0) \langle \chi_{i_1}(1), \dots, \chi_{i_k}(1) \rangle.$$

Somit sind die  $k$ -ten partiellen Ableitungen symmetrisch; es gilt

$$D_{i_k} \dots D_{i_1} f(x_0) = D_{i_{\sigma(k)}} \dots D_{i_{\sigma(1)}} f(x_0)$$

für alle  $\sigma \in S(k)$ .

- (vi) Die Funktion  $f = (f^i)_{1 \leq i \leq m}$  ist genau dann in einem Punkt  $k$ -mal differenzierbar, wenn dies für alle Komponentenfunktionen  $f^i$ ,  $1 \leq i \leq m$ , gilt und es folgt in diesem Falle

$$D^k f(x) = (D^k f^1(x), \dots, D^k f^m(x)),$$

wobei wir  $L_k(E, F)$  mit  $\prod_{i=1}^m L_k(E, F_i)$  identifizieren.

- (vii) Auch für höhere Ableitungen  $k$ -mal differenzierbarer Funktionen gilt eine Kettenregel, allerdings sehen die Darstellungen höherer Ableitungen von Verknüpfungen komplizierter aus.  
 (viii) Ist  $f \in C^k(\Omega)$ , so gilt

$$D_{i_k} \dots D_{i_1} f(x) = D^k f(x) \langle e_{i_1}, \dots, e_{i_k} \rangle$$

für alle  $x \in \Omega$  und alle  $i_l$  mit  $1 \leq i_l \leq n$ ,  $1 \leq l \leq k$ .

Für Vektoren  $u_l = (u_l^1, \dots, u_l^n)^T$  folgt

$$D^k f(x) \langle u_1, \dots, u_k \rangle = \sum_{i_1, \dots, i_k=1}^n D_{i_k} \dots D_{i_1} f(x) u_1^{i_1} \dots u_k^{i_k}.$$

### Beispiele 6.5.16.

- (i) Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $(x, y) \mapsto x^2 y$  gegeben. Dann ist

$$f_{12}((x, y)) = D_2(2xy) = 2x = D_1(x^2) = f_{21}((x, y)).$$

- (ii) ★ Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0), \end{cases}$$

ist im Ursprung zweimal partiell differenzierbar, jedoch sind die zweiten partiellen Ableitungen dort weder stetig noch symmetrisch.

*Beweis.* Übung. □

- (iii) Sei  $B: \mathbb{R}^n \times \mathbb{R}^m \equiv \mathbb{R}^{n+m} \rightarrow \mathbb{R}^k$  bilinear. Somit ist  $B$  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  $DB(\cdot, \cdot)\langle(a, b)\rangle: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^k$  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  $D^3B \equiv 0$ .  $\square$

- (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

$$\Phi: GL(E) \ni A \mapsto A^{-1} \in GL(E) \subset L(E)$$

in  $GL(E)$  differenzierbar und es gilt  $D\Phi(A)\langle B \rangle = -A^{-1}BA^{-1}$ . (Vergleiche dies mit  $(g^{-1})' = -\frac{g'}{g^2}$ .) 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

$$\begin{aligned} o(\|B\|) &\stackrel{!}{=} A(B + A)^{-1} - AA^{-1} + AA^{-1}BA^{-1} \\ &= ((B + A)A^{-1})^{-1} - \mathbf{1} + BA^{-1} \\ &= (BA^{-1} + \mathbf{1})^{-1} - \mathbf{1} + BA^{-1}. \end{aligned}$$

Die Behauptung folgt nun mit Hilfe der Neumannschen Reihe, denn es gilt

$$(\mathbf{1} + BA^{-1})^{-1} = \mathbf{1} - BA^{-1} + (BA^{-1})^2 - (BA^{-1})^3 \pm \dots$$

für  $\|BA^{-1}\| < 1$ .  $\square$

## 6.6. Taylorsche Formel.

**Definition 6.6.1.** Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow F$  eine  $k$ -mal differenzierbare Abbildung. Seien  $u \in E$  und  $x \in \Omega$ . Dann setzen wir

$$D^k f(x)\langle u \rangle^k := D^k f(x) \underbrace{\langle u, \dots, u \rangle}_{k \text{ Stück}}$$

für  $k \in \mathbb{N}_{>0}$  und  $D^0 f(x)\langle u \rangle^0 = f(x)$ .

**Lemma 6.6.2.** Seien  $E, F$  Banachräume,  $\Omega \subset E$  offen und  $f: \Omega \rightarrow F$  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

$$\varphi^{(k)}(t) = D^k f(x + tu)\langle u \rangle^k.$$

*Beweis.* Kettenregel und Induktion.  $\square$

**Theorem 6.6.3** (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

$$f(x + u) = \sum_{k=0}^p \frac{1}{k!} D^k f(x)\langle u \rangle^k + o(\|u\|^p).$$

*Beweis.* Da  $\Omega$  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  $\varphi$  mit  $[0, 1] \ni t \mapsto \varphi(t) \equiv f(x + tu) \in F$  wie in Lemma 6.6.2. Dann ist  $\varphi$  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 4.6.2, an und erhalten

$$\varphi(1) = \sum_{k=0}^p \frac{1}{k!} \varphi^{(k)}(0) + R_p(\varphi, 0)(1)$$

sowie die Restgliedabschätzung

$$\|R_p(\varphi, 0)(1)\| \leq \frac{1}{(p-1)!} \sup_{0 < t < 1} \|\varphi^{(p)}(t) - \varphi^{(p)}(0)\|.$$

Die Behauptung folgt nun direkt mit Hilfe der Formel für die Ableitungen von  $\varphi$  aus Lemma 6.6.2.  $\square$

**Bemerkung 6.6.4** (Lagrangesche Restgliedabschätzung). Ist  $f \in C^{p+1}(\Omega, F)$ , so erhalten wir mit einem analogen Vorgehen, diesmal unter Verwendung von Theorem 4.6.7, die Abschätzung

$$f(x + u) = \sum_{k=0}^p \frac{1}{k!} D^k f(x) \langle u \rangle^k + O(\|u\|^{p+1}).$$

Stellen wir totale Ableitungen mit Hilfe der partiellen Ableitungen dar, so erhalten wir aus der Taylorschen Formel

**Korollar 6.6.5.** Sei  $\Omega \subset \mathbb{R}^n$  offen,  $F$  ein Banachraum und  $f \in C^p(\Omega, F)$  mit  $p \geq 1$ . Dann gilt für alle  $x, x + u \in \Omega$

$$f(x + u) = \sum_{k=0}^p \frac{1}{k!} \sum_{i_1, \dots, i_k=1}^n D_{i_1} \dots D_{i_k} f(x) u^{i_1} \dots u^{i_k} + o(\|u\|^p).$$

**Bemerkung 6.6.6** (Multiindexschreibweise).  $\star$  Ist  $\alpha \in \mathbb{N}^n$ , so schreiben wir  $D^\alpha f$  für die Ableitung  $D_1^{\alpha_1} \dots D_n^{\alpha_n} f$  und nennen  $|\alpha| = \sum_{i=1}^n \alpha_i$  die Ordnung der Ableitung.

Weiterhin schreiben wir  $\prod_{i=1}^n (u^i)^{\alpha_i} \equiv u^\alpha$ . In dieser Schreibweise erhalten wir die folgende Form der Taylorschen Formel:

$$f(x + u) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha f(x) u^\alpha + o(\|u\|^p)$$

mit  $\alpha! := \alpha_1! \cdot \dots \cdot \alpha_n!$ .

*Beweis.* Dies folgt aus rein kombinatorischen Überlegungen aus der obigen Taylorschen Formel.  $\square$

**6.7. Lokale Extrema.** Hinreichende und notwendige Bedingungen für lokale Extrema benutzen die aus der linearen Algebra bekannte Definitheit.

**Definition 6.7.1.** Sei  $E$  ein Banachraum und  $a \in L_2(E; \mathbb{R})$  eine symmetrische Bilinearform. Dann heißt  $a$

- (i)  $\star$  **positiv semidefinit**,  $a \succcurlyeq 0$ , falls  $a(x, x) \geq 0$  für alle  $x \in E$  gilt.
- (ii)  $\star$  **positiv definit**,  $a \succ 0$ , falls  $a(x, x) > 0$  für alle  $x \in E \setminus \{0\}$  gilt.
- (iii) **gleichmäßig positiv definit**, falls es ein  $c > 0$  mit  $a(x, x) \geq c$  für alle  $x \in E$  mit  $\|x\| = 1$  gibt.
- (iv)  $\star$  **negativ semidefinit**,  $a \preccurlyeq 0$ , falls  $-a$  positiv semidefinit ist;  
**negativ definit**,  $a \prec 0$ , falls  $-a$  positiv definit ist;  $\dots$

- (v) ★ **(gleichmäßig) definit**, falls  $a$  (gleichmäßig) positiv oder negativ definit ist.
- (vi) ★ **indefinit**, falls es Vektoren  $x, y \in E$  mit  $a(x, x) > 0$  und  $a(y, y) < 0$  gibt.
- (vii) 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  $A \succcurlyeq 0, \dots$
- (viii) ★ Im komplexen Fall benutzen wir die entsprechenden Definitionen für hermitesche (gleiche Linearitätseigenschaften wie beim komplexen Skalarprodukt) Sesquilinearformen  $a \in L_2(E; \mathbb{C})$ .
- (ix) ★ Häufig schreibt man auch einfach  $a \geq 0, a > 0, \dots$

**Bemerkung 6.7.2.** ★ Ist  $\dim E < \infty$ , so ist jede positiv definite Bilinearform auch gleichmäßig positiv definit.

Bei symmetrischen  $\mathbb{R}^{n \times n}$ -Matrizen kann man mit Hilfe der Eigenwerte Definitheit überprüfen, siehe [5]:

**Lemma 6.7.3.** ★ Sei  $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n}$  eine symmetrische Bilinearform darstellende Matrix und seien  $(\lambda_i)_{1 \leq i \leq n}$  deren Eigenwerte mit Vielfachheiten. Dann gelten

- (i)  $A \succcurlyeq 0 \iff \lambda_i \geq 0 \quad \forall i,$
- (ii)  $A \succ 0 \iff \lambda_i > 0 \quad \forall i,$
- (iii)  $A \preccurlyeq 0 \iff \lambda_i \leq 0 \quad \forall i,$
- (iv)  $A \prec 0 \iff \lambda_i < 0 \quad \forall i$  und
- (v)  $A$  ist indefinit  $\iff \exists i, j: \lambda_i < 0 < \lambda_j.$

*Beweis.* Diagonalisiere  $A$ . □

Bei symmetrischen  $\mathbb{R}^{n \times n}$ -Matrizen kann man die strikte Definitheit mit Hilfe der Vorzeichen von Unterdeterminanten untersuchen, siehe [5]:

**Lemma 6.7.4.** ★ Sei  $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n}$  eine symmetrische Bilinearform darstellende Matrix. Dann gelten

- (i)  $A \succ 0 \iff \det(a_{ij})_{1 \leq i, j \leq k} > 0 \quad \forall 1 \leq k \leq n,$
- (ii)  $A \prec 0 \iff (-1)^k \det(a_{ij})_{1 \leq i, j \leq k} > 0 \quad \forall 1 \leq k \leq n.$

*Beweis.* Siehe [5]. □

**Bemerkung 6.7.5.** ★ Beachte, dass die entsprechende Äquivalenz im Falle  $A \succcurlyeq 0$  nicht mehr gilt, betrachte dazu  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ .

**Theorem 6.7.6.** 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  $D^2f(x_0) \succcurlyeq 0$  in einem lokalen Minimum  
★ bzw.  $D^2f(x_0) \preccurlyeq 0$  in einem lokalen Maximum.

*Beweis.*

- (i) Sei  $u \in E$  beliebig. Dann ist 0 eine lokale Extremalstelle der Funktion  $\varphi$  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  $\varphi$  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

$$f(x_0 + u) = f(x_0) + \frac{1}{2} D^2f(x_0)\langle u \rangle^2 + o(\|u\|^2)$$

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  $f(x_0 + u) \geq f(x_0)$  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

$$0 \leq f(x_0 + te) - f(x_0) = \frac{1}{2}t^2 D^2 f(x_0)(e, e) + o(t^2).$$

Wir dividieren nun durch  $t^2$ , lassen  $t \rightarrow 0$  und erhalten, da  $e \in E$  mit  $\|e\| = 1$  beliebig war, die Behauptung.  $\square$

**Bemerkung 6.7.7.**  $\star$  Bereits im eindimensionalen Fall haben wir gesehen, dass  $D^2 f(x_0) \succ 0$  zwar eine notwendige, jedoch keine hinreichende Bedingung für ein lokales Minimum ist, wie man am Beispiel  $\mathbb{R} \ni t \mapsto t^3 \in \mathbb{R}$  sieht.

**Theorem 6.7.8.** 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  $D^2 f(x_0)\langle u, u \rangle \geq 2c\|u\|^2$  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 Taylorsche Formel

$$f(x_0 + u) \geq f(x_0) + c\|u\|^2 + o(\|u\|^2).$$

Ist  $r$  klein genug, so erhalten wir  $|o(\|u\|^2)| \leq \frac{c}{2}\|u\|^2$  und es folgt

$$f(x_0 + u) \geq f(x_0) + \frac{c}{2}\|u\|^2$$

für alle  $u \in B_r(0)$ . Dies impliziert die Behauptung.  $\square$

**Beispiele 6.7.9.**  $\star$

- (i) Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $(x, y) \mapsto x^2 - y^2$  hat in  $(0, 0)$  einen kritischen Punkt.  $D^2 f(0)$  ist indefinit; wir nennen solch einen Punkt einen **Sattelpunkt**. Die Anzahl der negativen Eigenwerte von  $D^2 f(0)$  heißt **Index** von  $f$  im kritischen Punkt 0. Für den Index gilt hier  $\text{ind } f(0) = 1$ .
- (ii) Die Funktion  $\mathbb{R}^2 \ni (x, y) \mapsto e^{-3(x^2+y^2)}(x^2 - y^2)$  hat zwei lokale Minima, zwei lokale Maxima und einen Sattelpunkt.

## 7. EXISTENZSÄTZE

**7.1. Banachscher Fixpunktsatz.** Mit kontrahierenden Abbildungen bekommen wir eine weitere Möglichkeit, die Existenz von  $\sqrt{2}$  zu zeigen.

**Definition 7.1.1.** Sei  $E$  ein metrischer Raum und  $T: E \rightarrow E$  eine Abbildung. Dann heißt  $T$  **kontrahierend** oder eine **Kontraktion**, falls es ein  $k \in [0, 1)$  gibt, so dass für alle  $x, y \in E$

$$d(T(x), T(y)) \leq kd(x, y)$$

gilt.

**Bemerkung 7.1.2.** Jede Kontraktion ist stetig.

**Theorem 7.1.3** (Banachscher Fixpunktsatz). *Sei  $E$  ein vollständiger metrischer Raum und sei  $T$  eine kontrahierende Abbildung mit Konstante  $k \in [0, 1)$ . Dann gibt es genau ein  $\hat{x} \in E$  mit  $\hat{x} = T(\hat{x})$ , d. h.  $\hat{x}$  ist der (einzige) Fixpunkt von  $T$ .*

*Beweis.* Sei  $x_0 \in E$  beliebig. Definiere induktiv  $x_{n+1} := T(x_n)$ .

- (i)  $(x_n)_n$  **ist eine Cauchyfolge:** Es gilt

$$\begin{aligned} d(x_{n+m}, x_n) &\leq d(x_{n+m}, x_{n+m-1}) + \dots + d(x_{n+1}, x_n) \\ &\leq (k^{n+m-1} + \dots + k^n) \cdot d(x_1, x_0) \\ &= k^n (k^{m-1} + \dots + 1) \cdot d(x_1, x_0) \\ &\leq k^n \sum_{j=0}^{\infty} k^j \cdot d(x_1, x_0) \\ &= \frac{k^n}{1-k} d(x_1, x_0). \end{aligned}$$

Die rechte Seite konvergiert für  $n \rightarrow \infty$  gegen Null. Somit ist  $(x_n)_n$  eine Cauchyfolge.

- (ii) **Existenz eines Fixpunktes  $\hat{x}$ :** Da  $E$  vollständig ist, konvergiert die Cauchyfolge gegen ein  $\hat{x} \in E$ . Wegen  $x_n \rightarrow \hat{x}$  erhalten wir

$$\begin{aligned} d(\hat{x}, T(\hat{x})) &\leq d(\hat{x}, x_n) + d(x_n, T(\hat{x})) \\ &= d(\hat{x}, x_n) + d(T(x_{n-1}), T(\hat{x})) \\ &\leq d(\hat{x}, x_n) + kd(x_{n-1}, \hat{x}) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

- (iii) **Eindeutigkeit:** Sei  $\bar{x}$  ein weiterer Fixpunkt. Dann gelten  $T(\bar{x}) = \bar{x}$  und

$$d(\bar{x}, \hat{x}) = d(T(\bar{x}), T(\hat{x})) \leq kd(\bar{x}, \hat{x}).$$

Somit ist  $d(\bar{x}, \hat{x}) = 0$ . Daher ist der Fixpunkt eindeutig bestimmt.  $\square$

**Bemerkung 7.1.4.**  $\star$  Mit  $m \rightarrow \infty$  erhält man aus der längsten Abschätzung aus dem Beweis des Banachschen Fixpunktsatzes die Ungleichung

$$d(\hat{x}, x_n) \leq \frac{k^n}{1-k} d(x_1, x_0).$$

Damit kann man abschätzen, wie genau man einen Fixpunkt nach  $n$  Iterationen erreicht hat.

**Beispiel 7.1.5.** Auf dem vollständigen metrischen Raum  $[1, 2] \subset \mathbb{R}$  betrachten wir die Abbildung  $T: [1, 2] \rightarrow \mathbb{R}$  mit  $x \mapsto x + c(2 - x^2)$  für unterschiedliche Werte von  $c$ .

- (i) Es gilt  $T(1) = 1 + c$ . Somit kann  $T$  höchstens für  $c \in [0, 1]$  eine Selbstabbildung sein.
- (ii) Es gilt  $T(2) = 2 - 2c$ . Daher kann  $T$  höchstens für  $c \in [0, \frac{1}{2}]$  eine Selbstabbildung sein.
- (iii) Für  $c = 0$  gilt  $T(x) = x$ , also ist  $T$  keine Kontraktion.
- (iv) Quadratisches Ergänzen liefert

$$T(x) = x + 2c - cx^2 = -c \left( x^2 - \frac{1}{c}x - 2 \right) = -c \left( \left( x - \frac{1}{2c} \right)^2 - \frac{1}{4c^2} - 2 \right).$$

Die Funktion  $T$ , auf  $\mathbb{R}$  fortgesetzt, besitzt also an der Stelle  $x = \frac{1}{2c}$  ein lokales Maximum. Ist  $1 \leq \frac{1}{2c} \leq 2$ , so ist dies im Definitionsbereich von  $T$ . Dann gilt also  $\frac{1}{4} \leq c \leq \frac{1}{2}$ . Der Wert in diesem Maximum ist dann also durch

$$-c \left( -\frac{1}{4c^2} - 2 \right) = \frac{1}{4c} + 2c \leq 1 + 1 = 2$$

nach oben abgeschätzt.

(v) Somit gilt für  $c \in [0, \frac{1}{2}]$  die Inklusion  $T([1, 2]) \subset [1, 2]$ . Betrachte daher  $T$  als Selbstabbildung  $T: [1, 2] \rightarrow [1, 2]$ .

(vi) Um zu sehen, wann  $T$  eine Kontraktion ist, betrachten wir

$$\begin{aligned} |T(x) - T(y)| &= |x + c(2 - x^2) - y - c(2 - y^2)| = |x - cx^2 - y + cy^2| \\ &= |x - y| \cdot |1 - c(x + y)|. \end{aligned}$$

Wegen  $x + y \in [2, 4]$  und  $c(x + y) \in (0, 2)$  für  $0 < c < \frac{1}{2}$  ist  $T$  für  $c \in (0, \frac{1}{2})$  eine Kontraktion.

(vii) Für den Fixpunkt  $\hat{x} \in [1, 2]$  gilt also  $\hat{x} = \hat{x} + c(2 - \hat{x}^2)$ , also  $\hat{x}^2 = 2$  im Falle  $0 < c < \frac{1}{2}$ .

Die Abbildung  $T$  liefert auch im Falle  $c = \frac{1}{2}$ , wenn  $T$  keine Kontraktion mehr ist, eine Möglichkeit,  $\sqrt{2}$  zu berechnen. Es gilt allgemeiner

**Theorem 7.1.6.** *★ Sei  $E$  ein kompakter metrischer Raum. Sei  $T: E \rightarrow E$  eine Abbildung und gelte*

$$d(T(x), T(y)) < d(x, y)$$

*für alle  $x \neq y \in E$ . Dann besitzt  $T$  einen eindeutig bestimmten Fixpunkt.*

*Beweis.* Übung.

Man kann den Beweis des Banachschen Fixpunktsatzes wie folgt modifizieren: Betrachte für ein  $x \in E$  die Folge  $d(x, T(x))$ ,  $d(T(x), T(T(x)))$ ,  $\dots$ . Solange alle betrachteten Punkte  $T^k(x), T^{k+1}(x) \in E$  die Ungleichung  $d(T^k(x), T^{k+1}(x)) \geq \varepsilon$  erfüllen, erhält man eine Kontraktion. Auf diese Weise findet man ein  $y \in E$  mit  $d(y, T(y)) < \varepsilon$ . Es gibt aber noch einen deutlich kürzeren Beweis.  $\square$

Für später benötigen wir noch die folgende Variante des Banachschen Fixpunktsatzes.

**Theorem 7.1.7.** *★ Sei  $E$  ein Banachraum und  $f: \bar{B}_r(0) \rightarrow E$  mit  $\bar{B}_r(0) \subset E$  eine Kontraktion mit Kontraktionskonstante  $\kappa$ . Gelte  $\|f(0)\| \leq r(1 - \kappa)$ . Dann besitzt  $f$  einen Fixpunkt.*

Im Unterschied zum Banachschen Fixpunktsatz nehmen wir hier nicht an, dass  $f: \bar{B}_r(0) \rightarrow \bar{B}_r(0)$  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  $x_n \in \bar{B}_r(0)$  für alle  $0 \leq n \leq N$  gilt. Wie beim Beweis des Banachschen Fixpunktsatzes erhalten wir für alle  $0 < n \leq N + 1$

$$\begin{aligned} \|x_n - x_{n-1}\| &= \|f(x_{n-1}) - f(x_{n-2})\| \leq \kappa \cdot \|x_{n-1} - x_{n-2}\| \\ &= \kappa \cdot \|f(x_{n-2}) - f(x_{n-3})\| \leq \kappa^2 \cdot \|x_{n-2} - x_{n-3}\| \leq \dots \\ &\leq \kappa^{n-1} \cdot \|x_1 - x_0\| = \kappa^{n-1} \cdot \|f(0)\|, \\ \|x_{N+1}\| &= \|x_0 - x_{N+1}\| \leq \|x_0 - x_1\| + \dots + \|x_N - x_{N+1}\| \\ &\leq \sum_{i=0}^N \kappa^i \cdot \|f(0)\| \leq \frac{1}{1 - \kappa} \cdot \|f(0)\| \leq r. \end{aligned}$$

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  $\bar{B}_r(0)$ .  $\square$

**7.2. Satz über die Umkehrabbildung.** Wir beginnen mit etwas Vorbereitung. Da einige Resultate auch für Hölderräume gelten, definieren wir diese.

**Definition 7.2.1.** ★

- (i) Sei  $E$  ein metrischer Raum und sei  $F$  ein Banachraum. Eine Funktion  $f: E \rightarrow F$  heißt **lokal hölderstetig** mit Exponent  $\alpha$ ,  $0 < \alpha \leq 1$ , falls

$$\forall x_0 \in E \quad \exists U \in \mathcal{U}(x_0) \quad \exists c \geq 0 \quad \forall x, y \in U \quad \|f(x) - f(y)\| \leq cd(x, y)^\alpha$$

gilt. Ist  $\alpha = 1$ , so sprechen wir von lokal Lipschitzstetigen Funktionen. Wir schreiben  $C^{0,\alpha}(E, F)$ ,  $C^{0,1}(E, F)$  bzw. im Fall  $F = \mathbb{R}$  auch  $C^{0,\alpha}(E)$  und  $C^{0,1}(E)$ .

- (ii) Sind  $E, F$  Banachräume und sei  $\Omega \subset E$  offen. Sei  $m \in \mathbb{N}_{>0}$  und  $0 < \alpha \leq 1$ . Dann definieren wir

$$C^{m,\alpha}(\Omega, F) := \{f \in C^m(\Omega, F) : D^m f \in C^{0,\alpha}(\Omega, L_m(E; F))\}$$

und schreiben  $C^{m,\alpha}(\Omega, \mathbb{R}) \equiv C^{m,\alpha}(\Omega)$ .

Wichtig wird in diesem Kapitel die Menge der Diffeomorphismen.

**Definition 7.2.2** (Diffeomorphismus). Seien  $E, F$  Banachräume und  $U \subset E$  sowie  $V \subset F$  offen.

- (i) Eine Abbildung  $f: U \rightarrow V$  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  $\text{Diff}^k(U, V)$ .
- (ii) Eine Abbildung  $f: U \rightarrow F$  heißt **lokaler  $C^k$ -Diffeomorphismus**, falls  $f$  eine offene Abbildung ist (dies ist nötig um später die Definition von  $\text{Diff}^k$  anwenden zu können) und jeder Punkt  $x_0 \in U$  eine offene Umgebung  $B_r(x_0) \subset U$  besitzt, so dass  $f \in \text{Diff}^k(B_r(x_0), f(B_r(x_0)))$  gilt. Wir bezeichnen die Menge aller lokalen  $C^k$ -Diffeomorphismen von  $U$  nach  $F$  mit  $\text{Diff}_{\text{loc}}^k(U, F)$ .
- (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  $\varphi^{-1}$ . Weiterhin bemerken wir, dass  $L_{\text{top}}(E, E) = GL(E)$  gilt.)

Für hölderstetige Funktionen gilt

**Lemma 7.2.3.** ★ Seien  $E_i$ ,  $1 \leq i \leq 3$ , Banachräume und  $\Omega_i \subset E_i$  offen. Seien  $m \in \mathbb{N}_{>0}$  und  $0 < \alpha \leq 1$ .

- (i) Seien  $f \in C^{m,\alpha}(\Omega_1, E_2)$  und  $g \in C^{m,\alpha}(\Omega_2, E_3)$  mit  $f(\Omega_1) \subset \Omega_2$ . Dann gilt  $g \circ f \in C^{m,\alpha}(\Omega_1, E_3)$ .
- (ii) Sei  $f \in \text{Diff}^m(\Omega_1, \Omega_2)$  und gelte  $f \in C^{m,\alpha}(\Omega_1, E_2)$ . Dann folgt für die Umkehrabbildung  $f^{-1}$  auch  $f^{-1} \in C^{m,\alpha}(\Omega_2, E_1)$ .

Daher bezeichnen wir solche Diffeomorphismen auch mit  $\text{Diff}^{m,\alpha}(\Omega_1, \Omega_2)$  und im Fall lokaler Diffeomorphismen mit  $\text{Diff}_{\text{loc}}^{m,\alpha}(\Omega_1, \Omega_2)$ .

*Beweis.* Übung. □

**Theorem 7.2.4** (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  $x_0 \in \Omega$ . Ist  $Df(x_0) \in L_{\text{top}}(E, F)$ , so gibt es eine offene Umgebung  $U = B_r(x_0) \subset \Omega$ , so dass  $V = f(U)$  offen ist und  $f \in \text{Diff}^k(U, V)$  gilt.

Gilt also  $Df(x) \in L_{\text{top}}(E, F)$  für alle  $x \in \Omega$ , so folgt  $f \in \text{Diff}_{\text{loc}}^k(\Omega, F)$ .

*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  $\Phi: GL(E) \rightarrow GL(E)$  mit  $A \mapsto A^{-1}$  von der Klasse  $C^\infty$  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  $f: U \rightarrow V := f(B_r(x_0))$  ein Homöomorphismus ist (und zeigen dies später noch). Sei  $g := (f|_U)^{-1}$ . Nach Theorem 6.1.12, 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.

- (iii) **Behauptung:** Es gibt ein  $r > 0$ , so dass  $f|_U: U \rightarrow V$  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.
- (iv) Surjektivität und Stetigkeit sind klar.
- (v) Injektivität: Wir gehen ähnlich wie beim Beweis von Theorem 6.1.12 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  $\|Df(y)\langle v \rangle\| \geq c\|v\|$  gilt. (Erklärung dazu: Aufgrund der Stetigkeit von  $(Df(x_0))^{-1}$  gibt es eine Konstante  $c > 0$ , so dass  $\|(Df(x_0))^{-1}\langle u \rangle\| \leq \frac{1}{2c}\|u\|$  für alle  $u \in E$  gilt. Wir setzen nun insbesondere  $u = Df(x_0)\langle v \rangle$  und erhalten  $2c \cdot \|v\| \leq \|Df(x_0)\langle v \rangle\|$  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  $\|Df(y) - Df(x_0)\| \leq c$  in der Operatornorm gilt. Daher folgt für solche  $y$

$$\begin{aligned} \|Df(y)\langle v \rangle\| &\geq \|Df(x_0)\langle v \rangle\| - \|Df(x_0)\langle v \rangle - Df(y)\langle v \rangle\| \\ &\geq 2c \cdot \|v\| - \|Df(x_0) - Df(y)\| \cdot \|v\| \geq c \cdot \|v\| \end{aligned}$$

wie oben behauptet.)

Indem wir die Umgebung gegebenenfalls weiter verkleinern, können wir erreichen, dass  $\|o(\|x - y\|)\| \leq \frac{c}{2}\|x - y\|$  für alle  $x, y$  in dieser Umgebung gilt, da wir aufgrund des Mittelwertsatzes die Abschätzung

$$\begin{aligned} \|o(\|x - y\|)\| &= \|f(x) - f(y) - Df(y)\langle x - y \rangle\| \\ &\leq \|x - y\| \cdot \sup_{\xi \in [x, y]} \|Df(\xi) - Df(y)\| \end{aligned}$$

haben und  $Df$  stetig ist. Wir erhalten somit für  $x, y \in B_r(x_0)$  für geeignete kleine  $r > 0$

$$\|f(x) - f(y)\| \geq \frac{c}{2}\|x - y\|$$

und somit die Injektivität.

(vi)  $f|_U$  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  $B_\varepsilon(f(\bar{x})) \subset f(\tilde{U})$  gilt.

Ohne Einschränkung dürfen wir (nach Anwendung von Translationen) annehmen, dass  $\bar{x} = 0$  und  $f(\bar{x}) = 0$  gelten.

Sei  $\kappa$  mit  $0 < \kappa < 1$  beliebig. Da  $Df$  stetig ist, gibt es ein  $\rho > 0$ , so dass  $B_\rho(0) \subset \tilde{U}$  und

$$\left\| \text{id} - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y) dt \right\| \leq \kappa$$

für alle  $x, y \in \bar{B}_\rho(0)$  gelten. (Erklärung dazu: Es gilt

$$\begin{aligned} & \left\| \text{id} - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y) dt \right\| \\ &= \left\| (Df(0))^{-1} \circ (Df(0)) - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y) dt \right\| \\ &= \left\| (Df(0))^{-1} \circ \left( Df(0) - \int_0^1 Df(tx + (1-t)y) dt \right) \right\| \\ &\leq \|(Df(0))^{-1}\| \cdot \left\| \int_0^1 Df(0) dt - \int_0^1 Df(tx + (1-t)y) dt \right\|, \end{aligned}$$

da  $Df(0)$  von  $t$  unabhängig ist,

$$\begin{aligned} &= \|(Df(0))^{-1}\| \cdot \left\| \int_0^1 Df(0) - Df(tx + (1-t)y) dt \right\| \\ &\leq \|(Df(0))^{-1}\| \cdot \sup_{z \in [x, y]} \|Df(0) - Df(z)\| \cdot 1 \\ &\leq \|(Df(0))^{-1}\| \cdot \sup_{z \in B_\rho(0)} \|Df(0) - Df(z)\| \end{aligned}$$

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  $\varphi: \bar{B}_\rho(0) \rightarrow E$  für  $\rho > 0$  wie oben durch

$$\varphi(x) := x + (Df(0))^{-1} \langle u - f(x) \rangle$$

und wollen zeigen, dass  $\varphi$  für kleine  $\varepsilon > 0$  die Voraussetzungen von Theorem 7.1.7 erfüllt. Dann erhalten wir hieraus nämlich einen Fixpunkt, was gleichbedeutend mit  $u - f(x) = 0$  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

$$\begin{aligned} \varphi(x) - \varphi(y) &= x - y - (Df(0))^{-1} \langle f(x) - f(y) \rangle \\ &= x - y - (Df(0))^{-1} \left\langle \int_0^1 \frac{d}{dt} f(tx + (1-t)y) dt \right\rangle \end{aligned}$$

$$\begin{aligned}
&= x - y - (Df(x))^{-1} \left\langle \int_0^1 Df(tx + (1-t)y) \langle x - y \rangle dt \right\rangle \\
&= \left( \text{id} - (Df(0))^{-1} \circ \int_0^1 Df(tx + (1-t)y) dt \right) \langle x - y \rangle.
\end{aligned}$$

Somit erhalten wir aufgrund der Wahl von  $\rho$

$$\|\varphi(x) - \varphi(y)\| \leq \kappa \cdot \|x - y\|$$

für alle  $x, y \in \bar{B}_\rho(0)$ . Somit ist  $\varphi$  wie gewünscht eine Kontraktion.

Mit Theorem 7.1.7 folgt nun die Behauptung, falls wir noch  $\|\varphi(0)\| \leq \rho(1 - \kappa)$  beweisen können. Aufgrund unserer Annahme  $f(0) = 0$  erhalten wir aber

$$\|\varphi(0)\| = \|(Df(0))^{-1} \langle u \rangle\| \leq \|(Df(0))^{-1}\| \cdot \|u\| \leq \|(Df(0))^{-1}\| \cdot \varepsilon.$$

Somit erhalten wir die behauptete Abschätzung, indem wir  $\varepsilon > 0$  falls nötig verkleinern, so dass

$$\varepsilon \leq \|(Df(0))^{-1}\|^{-1} \rho(1 - \kappa)$$

gilt. □

**Bemerkung 7.2.5.** ★ Beachte, dass wir in der letzten Abschätzung den Ausdruck  $\|(Df(0))^{-1}\|^{-1}$  nicht durch  $\|Df(0)\|$  ersetzen dürfen. Dies sind im Allgemeinen unterschiedliche reelle Zahlen.

**Korollar 7.2.6.** ★ Gilt im Satz über die Umkehrfunktion zusätzlich  $f \in C^{k,\alpha}(\Omega, F)$ , so folgt auch  $f \in \text{Diff}_{loc}^{k,\alpha}(\Omega, F)$ .

*Beweis.* Kombiniere den Satz über die Umkehrfunktion, Theorem 7.2.4 mit Lemma 7.2.3. □

Als weiteres Korollar zum Satz über die Umkehrfunktion erhalten wir die folgende globale Variante

**Korollar 7.2.7.** ★ 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  $f \in \text{Diff}^k(\Omega, f(\Omega))$ .

*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  $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Dann folgt aus der Bijektivität von  $Df(x_0)$  bereits  $n = m$  und wir erhalten

**Korollar 7.2.8.** Sei  $\Omega \subset \mathbb{R}^n$  offen und  $f \in C^k(\Omega, \mathbb{R}^n)$ ,  $k \geq 1$ . Gelte für die Jacobimatrix  $\det J_f(x_0) \neq 0$  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  $f \in \text{Diff}^k(U, V)$  gilt.

Also folgt aus  $\det J_f(x) \neq 0$  für alle  $x \in \Omega$ , dass  $f \in \text{Diff}_{loc}^k(\Omega, f(\Omega))$  gilt.

Eine Variante des Satzes über die Umkehrfunktion gilt auch, wenn  $Df(x_0)$  lediglich surjektiv ist. Wir empfehlen hier dringend, den Fall allgemeiner Banachräume bis zur Vorlesung über Funktionalanalysis zu überblättern, da es nicht so ganz einfach ist, sich abgeschlossene Teilräume vorzustellen, die nicht splitten (ein deutscher Begriff dazu scheint nicht üblich zu sein).

**Definition 7.2.9.** ★ Sei  $E$  ein normierter Raum und  $M \subset E$  ein abgeschlossener Teilraum. Dann **splittet**  $M$  den Raum  $E$ , falls  $M$  ein **topologisches Komplement** besitzt, d. h. es gibt einen abgeschlossenen Teilraum  $N \subset E$ , so dass  $E$  die **algebraische direkte Summe**, d. h. direkte Summe im Sinne der linearen Algebra, ist, also  $E = M \oplus N$  gilt und  $E$  als topologischer Raum zum Produktraum  $M \times N$  isomorph ist, d. h. die Abbildung

$$j: M \times N \rightarrow E \quad \text{mit} \quad (x, y) \mapsto x + y$$

ist ein topologischer Isomorphismus, oder, äquivalent dazu, falls es eine Konstante  $c > 0$  mit

$$\frac{1}{c} \|x + y\| \leq \max(\|x\|, \|y\|) \leq c \cdot \|x + y\|$$

für alle  $(x, y) \in M \times N$  gibt. In diesem Falle heißt  $N$  ein **topologisches Komplement** von  $M$ .  $E$  heißt **topologische direkte Summe** von  $M$  und  $N$ .

**Bemerkung 7.2.10.** ★ Sei  $E = M \oplus N$ . Betrachten wir  $M, N$  als Teilmengen von  $E$ , so können wir jeden Vektor  $z \in E$  in eindeutiger Weise als  $z = x + y$  mit  $x \in M$  und  $y \in N$  schreiben. Dies definiert lineare Abbildungen  $P_M: E \rightarrow E$  mit  $P_M z = x$  und  $P_N: E \rightarrow E$  mit  $P_N z = y$ . Die Abbildungen  $P_M$  und  $P_N$  sind idempotent, d. h. es gelten  $P_M^2 = P_M$  und  $P_N^2 = P_N$ .

**Definition 7.2.11.** ★ Sei  $E$  ein normierter Raum. Dann heißt  $P \in L(E)$  ein **Projektor**, falls  $P^2 = P$  gilt,  $P$  also **idempotent** ist.

**Lemma 7.2.12.** ★ Sei  $E$  ein Banachraum und  $M \subset E$  ein abgeschlossener Teilraum. Dann **splittet**  $M$  den Raum  $E$  genau dann, wenn es einen Projektor  $P \in L(E)$  mit  $R(P) = M$  gibt.

*Beweis.*

„ $\Rightarrow$ “: Sei  $E = M \oplus N$  eine topologische direkte Summe. Wir definieren  $P := P_M$ . Dann ist  $P$  idempotent und es gilt  $R(P) = M$ . Wir müssen noch nachweisen, dass  $P$  stetig ist, d. h. dass  $P \in L(E)$  gilt. Nach Voraussetzung gibt es ein  $c > 0$ , so dass für alle  $z = x + y$  mit  $(x, y) \in M \times N$  die Abschätzung  $\max(\|x\|, \|y\|) \leq c \cdot \|x + y\|$  gilt. Somit folgt  $\|Pz\| = \|x\| \leq \max(\|x\|, \|y\|) \leq c \cdot \|z\|$  für alle  $z \in E$ . Also ist  $P$  stetig.

„ $\Leftarrow$ “: Sei  $P \in L(E)$  ein Projektor mit  $R(P) = M$ . Wir definieren  $N := P^{-1}(\{0\})$ .

Wir behaupten, dass  $E = M \oplus N$  gilt: Sei dazu  $z \in E$ . Wir definieren dazu  $x := P(z)$  und  $y := z - P(z)$ . Es gilt  $(x, y) \in M \times N$ ; die Inklusion für  $x$  ist klar, die für  $y$  folgt aus  $P(y) = P(z - P(z)) = P(z) - P^2(z) = P(z) - P(z) = 0$ . Die Eindeutigkeit der Darstellung folgt, wenn in einer Darstellung  $0 = x + y$  mit  $(x, y) \in M \times N$  stets  $x = y = 0$  gilt. Aus  $x \in M$  folgt, dass es ein  $\bar{x} \in E$  mit  $P(\bar{x}) = x$  gibt. Damit folgt  $P(x) = P(P(\bar{x})) = P(\bar{x}) = x$ . Wir wenden nun  $P$  auf  $0 = x + y$  an und erhalten  $0 = x + 0$  nach Definition von  $N$ . Somit ist  $x = 0$  und folglich gilt auch  $y = 0$ . Also ist  $E$  die algebraische direkte Summe von  $M$  und  $N$ .

Zu den Normabschätzungen: Es gilt nach Dreiecksungleichung  $\|x + y\| \leq \|x\| + \|y\| \leq 2 \max(\|x\|, \|y\|)$ . Die Stetigkeit der Abbildungen  $P$  und  $\text{id} - P$  impliziert  $\|x\| = \|Pz\| \leq \|P\| \cdot \|z\|$  sowie  $\|y\| = \|(\text{id} - P)z\| \leq \|\text{id} - P\| \cdot \|z\|$ . Beide Ungleichungen zusammen liefern  $\max(\|x\|, \|y\|) \leq c \cdot \|z\|$ . Somit **splittet**  $M$ .  $\square$

Das folgende Lemma gilt insbesondere für  $H = \mathbb{R}^n$ .

**Lemma 7.2.13.** ★ Sei  $H$  ein Hilbertraum und  $M \subset H$  ein abgeschlossener Teilraum. Dann **splittet**  $M$  den Raum  $H$ .

*Beweisidee.* Das orthogonale Komplement  $M^\perp$  ist auch ein topologisches Komplement. Für  $z = x + y$  mit  $\langle x, y \rangle = 0$  folgt insbesondere  $\|x + y\|^2 = \|x\|^2 + \|y\|^2$  und somit die zweite Ungleichung in der Definition eines splittenden Unterraumes.

Endlichdimensional können wir eine Orthonormalbasis  $(a_1, \dots, a_k)$  von  $M$  zu einer Orthonormalbasis  $(a_1, \dots, a_n)$  von  $H$  ergänzen und erhalten das orthogonale Komplement als  $\langle a_{k+1}, \dots, a_n \rangle$ .  $\square$

**Proposition 7.2.14.** *★ Sei  $E$  ein Banachraum und  $M \subset E$  ein Unterraum mit  $\dim M < \infty$ . Dann besitzt  $M$  ein topologisches Komplement in  $E$ .*

*Beweis.* Seien  $\dim M = n$  und  $(e_1, \dots, e_n)$  eine Basis von  $M$ . Definiere  $M_i := \langle e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n \rangle$ ,  $1 \leq i \leq n$ . Dann gibt es aufgrund des Satzes von Hahn-Banach stetige Linearformen  $\varphi_i \in E^*$  mit  $\varphi_i|_{M_i} = 0$  und  $\langle e_i, \varphi_i \rangle = 1$ . Wir definieren nun eine Abbildung  $P: E \rightarrow E$  durch

$$P(z) := \sum_{i=1}^n e_i \langle z, \varphi_i \rangle.$$

Man rechnet direkt nach, dass  $P(e_i) = e_i$ ,  $R(P) \subset M$  und damit  $R(P) = M$  und  $P^2 = P$  gelten. Die Stetigkeit von  $P$  ist klar. Somit splittet  $M$  den Raum  $E$ .  $\square$

★ Bei der folgenden Folgerung aus dem Satz über die Umkehrabbildung empfehlen wir, zunächst nur den Fall zu betrachten, dass  $E = \mathbb{R}^n$  gilt oder  $E$  ein Hilbertraum ist.

**Theorem 7.2.15.** *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  $Df(x_0): E \rightarrow F$  surjektiv ist und eine der folgenden Bedingungen erfüllt sind:*

- (i)  $E = \mathbb{R}^n$ ,
- (ii)  $E$  ist ein Hilbertraum oder ★
- (iii)  $\ker(Df(x_0))$  splittet  $E$ . ★

*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  $x = (x^1, x^2)$ .

Im Falle  $E = \mathbb{R}^n$  erreichen wir dies beispielsweise, indem wir eine Orthonormalbasis von  $E_1$  mit  $(a_k, \dots, a_n)$  zu einer Orthonormalbasis von  $E$  ergänzen und  $E_2 := \langle a_k, \dots, a_n \rangle$  definieren.

Nun gilt für  $u \in E$

$$Df(x_0)\langle u \rangle = D_1f(x_0)\langle u^1 \rangle + D_2f(x_0)\langle u^2 \rangle.$$

(Die Notation  $D_i$  für partielle Ableitungen zu Unterräumen führen wir in Kapitel 7.3 bzw. Anhang 9 ein.) Nach Voraussetzung gilt  $D_1f(x_0)\langle u^1 \rangle = Df(x_0)\langle (u^1, 0) \rangle = 0$  und  $D_2f(x_0): E_2 \rightarrow F$  ist surjektiv, da  $Df(x_0)$  nach Voraussetzung surjektiv ist.

Dann ist die Abbildung  $D_2f(x_0)$  offen, da

- (i)  $D_2f(x_0)$  als injektive und surjektive Abbildung bijektiv ist.
- (ii) (im Fall mit ★) dies aus dem Satz von der offenen Abbildung folgt, siehe Funktionalanalysis.

Somit folgt in allen Fällen  $D_2f(x_0) \in L_{\text{top}}(E_2, F)$ . Die Abbildung  $g := f(x_0^1, \cdot)$  erfüllt daher die Voraussetzungen des Satzes von der inversen Abbildung. Somit gibt es eine Umgebung  $U \in \mathcal{U}(x_0^2)$  in  $E_2$ , so dass  $g: U \rightarrow g(U)$  ein  $C^1$ -Diffeomorphismus ist. Insbesondere ist  $g(U) \subset f(\Omega)$  offen.  $\square$

Wir wollen damit eine nichtlineare gewöhnliche Differentialgleichung lösen. Diese Methode kann man später auch auf nichtlineare partielle Differentialgleichungen anwenden.

**Beispiel 7.2.16.** ★ Seien  $E = C^1([0, 1])$  und  $F = C^0([0, 1])$  mit der  $C^1$ - bzw. der  $C^0$ -Norm. Wir definieren die nichtlineare Abbildung  $\varphi: E \rightarrow F$  durch

$$\varphi(u) := \dot{u} + u^2$$

für  $u \in E$ . Dann gilt  $\varphi \in C^1(E, F)$ , ja sogar  $\varphi \in C^\infty$ , und aus

$$\varphi(u + u_0) = \varphi(u_0) + \dot{u} + 2u_0u + u^2$$

erhalten wir

$$D\varphi(u_0)\langle u \rangle = \dot{u} + 2u_0u$$

für  $u, u_0 \in E$ . Wir betrachten nun speziell  $u_0 = 0$ . Dann ist  $D\varphi(0) = \frac{d}{dt} \in L(E, F)$ . Aufgrund des Hauptsatzes der Differential- und Integralrechnung ist  $D\varphi(0)$  surjektiv, da nämlich  $\left. \frac{d}{ds} \int_0^s f(\tau) d\tau \right|_{s=t} = f(t)$  für alle  $f \in F$  gilt. Weiterhin besteht  $\ker D\varphi(0)$  gerade aus den konstanten Funktionen, ist daher eindimensional und besitzt somit nach Proposition 7.2.14 ein topologisches Komplement.

Wir können daher Theorem 7.2.15 anwenden und erhalten, dass es ein  $\varepsilon > 0$  gibt, so dass für alle  $f \in F$  mit  $\|f\|_{C^0([0,1])} < \varepsilon$  ein  $u \in E$  mit

$$\dot{u} + u^2 = f$$

existiert. Wir haben damit eine nichtlineare gewöhnliche Differentialgleichung gelöst.

### 7.3. Satz von der impliziten Funktion.

**Bemerkung 7.3.1.** Der Satz von der impliziten Funktion verallgemeinert das folgende Resultat aus der linearen Algebra: Sei  $A = \begin{pmatrix} A_1 & A_2 \end{pmatrix}$  eine Blockmatrix, so dass  $A_2$  bijektiv ist. Dann gibt es für jedes  $x^1$  ein  $x^2 = \varphi(x^1)$ , so dass  $A \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = 0$  gilt.

#### Bemerkung 7.3.2.

- (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  $f: \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}^m$  und

$$D_1 f(x, y) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f^k}{\partial x^1} & \cdots & \frac{\partial f^k}{\partial x^n} \end{pmatrix} (x, y)$$

sowie

$$D_2 f(x, y) = \begin{pmatrix} \frac{\partial f^1}{\partial y^1} & \cdots & \frac{\partial f^1}{\partial y^m} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial y^1} & \cdots & \frac{\partial f^m}{\partial y^m} \end{pmatrix} (x, y).$$

Die Bedingung  $D_2 f(x_0, y_0) \in L_{\text{top}}(\mathbb{R}^m, \mathbb{R}^m)$  wird dabei zu

$$\det D_2 f(x_0, y_0) \neq 0.$$

- (ii) ★ Beachte, dass beim Definitionsbereich und Zielraum von  $f$ , also in

$$\mathbb{R}^n \times \mathbb{R}^m \supset \Omega_1 \times \Omega_2 \ni (x, y) \xrightarrow{f} f(x, y) \in \mathbb{R}^m$$

zweimal  $m$  auftritt. Dies ist nötig, damit  $D_2 f(x_0, y_0) \in L_{\text{top}}$  überhaupt erfüllt sein kann.

**Theorem 7.3.3** (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

$$f(x_0, y_0) = 0 \quad \text{sowie} \quad D_2 f(x_0, y_0) \in L_{\text{top}}(E_2, F).$$

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  $\varphi$  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  $\varphi$  zu zeigen, definieren wir

$$g: \Omega_1 \times \Omega_2 \rightarrow E_1 \times F \quad \text{durch} \quad (x, y) \mapsto (x, f(x, y)).$$

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

$$Dg = (D\pi_1, Df).$$

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

$$Dg(x, y) = \begin{pmatrix} \text{id}_{E_1} & 0 \\ D_1 f(x, y) & D_2 f(x, y) \end{pmatrix}$$

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

$$(u, v) := (\bar{u}, (D_2 f(x_0, y_0))^{-1} \langle \bar{v} - D_1 f(x_0, y_0)\langle \bar{u} \rangle \rangle).$$

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

$$(Dg(x_0, y_0))^{-1} \langle \bar{u}, \bar{v} \rangle = (\bar{u}, (D_2 f(x_0, y_0))^{-1} \langle \bar{v} - D_1 f(x_0, y_0)\langle \bar{u} \rangle \rangle).$$

- (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

$$\begin{aligned} \|(Dg(x_0, y_0))^{-1} \langle \bar{u}, \bar{v} \rangle\| &= \max(\|\bar{u}\|, \|(D_2 f(x_0, y_0))^{-1} \langle \bar{v} - D_1 f(x_0, y_0)\langle \bar{u} \rangle \rangle\|) \\ &\leq c \cdot \max(\|\bar{u}\|, \|\bar{v} - D_1 f(x_0, y_0)\langle \bar{u} \rangle\|) \end{aligned}$$

$$\begin{aligned} &\leq c \cdot \max(\|\bar{u}\|, \|\bar{v}\| + \|D_1 f(x_0, y_0)(\bar{u})\|) \\ &\leq c \cdot \max(\|\bar{u}\|, \|\bar{v}\| + \|\bar{u}\|) \leq c \cdot \max(\|\bar{u}\|, \|\bar{v}\|). \end{aligned}$$

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  $h$

$$h: W_0 \rightarrow U_0 \times V_0.$$

Wir schreiben  $h = (h^1, h^2)$  und erhalten für  $(x, z) \in W_0 \subset E_1 \times F$

$$(x, z) = g \circ h(x, z) = g(h^1(x, z), h^2(x, z)) = (h^1(x, z), f(h^1(x, z), h^2(x, z))).$$

Somit gelten für alle  $(x, z) \in W_0$  die Identitäten  $h^1 = \pi_1$  (Projektion auf die erste Komponente) und  $f(x, h^2(x, z)) = z$ .

Wir definieren nun  $U := \{x \in U_0 : (x, 0) \in W_0\}$ . 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  $\varphi: U \rightarrow V_0$  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.  $\varphi \in C^{k,\alpha}(U, E_2)$ .

Aus  $g(x_0, y_0) = (x_0, f(x_0, y_0)) = (x_0, 0)$  erhalten wir  $h(x_0, 0) = (x_0, y_0)$ ,  $h^2(x_0, 0) = y_0$  und somit  $\varphi(x_0) = y_0$ .

Ohne Einschränkung können wir durch Verkleinern von  $U$  schließlich noch annehmen, dass  $U$  wegzusammenhängend ist und  $\|\varphi(x) - y_0\| \leq r/2$  für alle  $x \in U$  gilt.

- (iv) Eindeutigkeit: Wir zeigen, dass für die gewählte Menge  $U$  die Abbildung  $\varphi: U \rightarrow E_2$  die einzige stetige Abbildung ist, so dass  $f(x, \varphi(x)) = 0$  und  $\varphi(x_0) = y_0$  gelten.

Zunächst beschränken wir uns dabei auf den Fall von Werten in  $V_0$ : 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  $g: U_0 \times V_0 \rightarrow W_0$  injektiv ist, kann es zu festem  $x \in U_0 \supset U$  maximal ein solches  $y \in V_0$  geben.

Wir müssen nun noch den Fall ausschließen, dass solch eine Abbildung Werte außerhalb von  $V_0$  annimmt: Angenommen,  $\psi: U \rightarrow E_2$  wäre eine weitere stetige Abbildung mit  $f(x, \psi(x)) = 0$  für alle  $x \in U$  und  $\psi(x_0) = y_0$ . Wir nehmen weiterhin an, dass es  $x_1 \in U$  mit  $\varphi(x_1) \neq \psi(x_1)$  gibt. Aufgrund der gerade gezeigten Eindeutigkeit innerhalb von  $V_0 = B_r(y_0)$  und  $\|\varphi(x_1) - y_0\| \leq r/2$  erhalten wir  $\|\psi(x_1) - y_0\| \geq r$ . Da  $U$  wegzusammenhängend ist, gibt es einen Weg  $\gamma: [0, 1] \rightarrow U$  mit  $\gamma(0) = x_0$  und  $\gamma(1) = x_1$ . Da weiterhin  $d: [0, 1] \rightarrow \mathbb{R}$  mit  $t \mapsto \|\psi(\gamma(t)) - y_0\|$  stetig ist und  $d(0) = 0$  sowie  $d(1) \geq r$  erfüllt, gibt es  $t_{3/4} \in [0, 1]$  mit  $d(t_{3/4}) = \|\psi(\gamma(t_{3/4})) - y_0\| = \frac{3}{4}r$ . Wir definieren  $x_2 := \gamma(t_{3/4}) \in U$  und erhalten somit  $f(x_2, \psi(x_2)) = 0 = f(x_2, \varphi(x_2))$  mit zwei verschiedenen Werten  $\varphi(x_2), \psi(x_2) \in V_0$  im Widerspruch zur oben gezeigten Eindeutigkeit in  $V_0$ .

Daher ist  $\varphi$  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: Wir setzen dazu speziell Vektoren ein, die nur in der zweiten Komponente von Null verschiedene Einträge haben und sehen, dass sich Ungleichungen der Form  $\frac{1}{c}\|v\| \leq \|Bv\| \leq c\|v\|$  von  $Dg$  auf  $D_2 f$  übertragen.

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 7.3.4.**  $\star$

- (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  $\text{rk } A \leq \min(m, n)$ . Gilt  $\text{rk } A = \min(m, n)$ , 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  $f: \mathbb{R}^n \times \mathbb{R}^m \equiv \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$  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  $f: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$  gegeben und  $Df(x_0, y_0)$  surjektiv bzw.  $J_f(x_0, y_0)$  hat vollen Rang, so können wir durch Ummummerierung der Koordinaten im  $\mathbb{R}^{n+m}$  erreichen, dass wir eine Abbildung  $\tilde{f}: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$  mit bijektivem  $D_2 \tilde{f}(x_0, y_0)$  erhalten.

**Aufgabe 7.3.5.** Zeige, dass es unter den Voraussetzungen des Satzes über implizite Funktionen eine Umgebung  $A$  von  $(x_0, y_0)$  mit

$$\{(x, y) \in A: f(x, y) = 0\} = \{(x, \varphi(x)): x \in U\} \cap A \equiv \text{graph } \varphi \cap A$$

gibt, d. h. dass sich lokal die gesamte Menge  $\{f = 0\}$  als  $\text{graph } \varphi$  beschreiben lässt.

*Hinweis.* Benutze den Teil des Beweises des Satzes über implizite Funktionen, in dem wir die Eindeutigkeit von  $\varphi$  gezeigt haben.  $\square$

In diesem Zusammenhang verwendet man häufiger die folgende Definition. Sie verallgemeinert auch die bisherige Definition eines kritischen Punktes.

**Definition 7.3.6.** 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**.

**Bemerkung 7.3.7.**  $\star$

- (i) Beachte, dass  $\mathcal{C}f(\Omega)$  aus regulären Werten besteht.
- (ii) Da der Zielraum endlichdimensional ist, gibt es nach Proposition 7.2.14 stets einen splittenden endlichdimensionalen Teilraum  $M$ , so dass  $Df(x)|_M$  bijektiv ist.

In der folgenden Bemerkung führen wir die später in der Differentialgeometrie noch wichtigen Hyperflächen ein.

**Bemerkung 7.3.8.**

- (i) Sei  $f \in C^k(\Omega)$ ,  $\Omega \subset \mathbb{R}^n$  offen und  $k \geq 1$ . Die Bedingung, dass  $\nabla f$  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  $\Omega$ , 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  $\hat{x}_0 \equiv (x_0^1, \dots, x_0^{n-1})$  und ein  $\varepsilon > 0$  sowie eine Funktion  $\varphi \in C^k(U)$  mit  $\varphi(\hat{x}_0) = x_0^n$  und

$$M_c \cap (U \times (x_0^n - \varepsilon, x_0^n + \varepsilon)) = \text{graph } \varphi$$

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 7.3.9.** 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 7.3.10.** Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  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  $M_c \cap \{x^n > 0\} = \text{graph}(\hat{x} \mapsto \sqrt{c - |\hat{x}|^2})$ .

**Aufgabe 7.3.11.** 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)  $M_c \cap U = (TO \text{ graph } \varphi) \cap U$ ,
- (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. Das Rangtheorem \*** Sei  $r \leq \min(m, n)$  und sei  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine Abbildung mit  $\varphi(x) := (x^1, \dots, x^r, 0, \dots, 0)$ . Dann hat  $D\varphi$  überall den konstanten Rang  $r$ . Das Rangtheorem besagt nun, dass lokal bis auf Diffeomorphismen auch die Umkehrung hiervon gilt.

**Definition 7.4.1.** Ein (**achsenparalleler**) **Würfel** im  $\mathbb{R}^n$  mit Kantenlänge  $2r$  und Mittelpunkt  $x_0$  ist die Menge

$$W_r(x_0) \equiv W_r^n(x_0) := \{x \in \mathbb{R}^n: |x^i - x_0^i| < r \text{ für alle } 1 \leq i \leq n\}.$$

Schreiben wir  $\mathbb{R}^n = \mathbb{R}^l \times \mathbb{R}^m$ , so erhalten wir  $W_r^n(x_0, y_0) = W_r^l(x_0) \times W_r^m(y_0)$ .

**Theorem 7.4.2.** Sei  $\Omega \subset \mathbb{R}^n$  offen und  $f \in C^k(\Omega, \mathbb{R}^m)$  mit  $k \geq 1$ . Nehme an, dass die Ableitung der Abbildung  $f$  in  $\Omega$  den konstanten Rang  $r$  hat,  $\text{rk } Df(x) = r$  für alle  $x \in \Omega$ . (In diesem Fall gilt automatisch  $r \leq \min(n, m)$ .) Sei  $x_0 \in \Omega$  beliebig. Dann gibt es offene Umgebungen  $U \in \mathcal{U}(x_0)$  und  $V \in \mathcal{U}(f(x_0))$  und  $C^k$ -Diffeomorphismen  $\varphi_n \in \text{Diff}^k(W_1^n(0), U)$  und  $\varphi_m \in C^k(V, W_1^m(0))$ , so dass die Abbildung  $\varphi_m \circ f \circ \varphi_n: W_1^n(0) \rightarrow W_1^m(0)$  durch

$$(x^1, \dots, x^n) \mapsto (x^1, \dots, x^r, 0, \dots, 0)$$

gegeben ist.

Ist zusätzlich  $f \in C^{k, \alpha}$ , so können wir auch  $\varphi_n$  und  $\varphi_m$  in  $C^{k, \alpha}$  wählen.

*Beweis.*

- (i) Ohne Einschränkung wollen wir annehmen, dass  $x_0 = 0$ ,  $f(x_0) = 0$  sowie  $Df(0)\langle u \rangle = (u^1, \dots, u^r, 0, \dots, 0)$  für alle  $u \in \mathbb{R}^n$  gelten. Sonst verknüpfen wir  $f$  mit Translationen um die ersten beiden Bedingungen zu erfüllen oder Elementen  $B \in GL(m)$  bzw.  $A \in GL(n)$ , so dass  $B \circ Df(0) \circ A$  das Differential  $B \circ Df(0) \circ A\langle u \rangle = (u^1, \dots, u^r, 0, \dots, 0)$  für alle  $u \in \mathbb{R}^n$  hat.
- (ii) Wir definieren eine Funktion  $g \in C^k(\Omega, \mathbb{R}^n)$  durch

$$g(x) = (f^1(x), \dots, f^r(x), x^{r+1}, \dots, x^n),$$

wobei wie ganz bewusst einige der Komponenten von  $f = (f^1, \dots, f^m)$  weglassen. Man rechnet leicht nach, dass  $Dg(0) = \text{id}_{\mathbb{R}^n}$  gilt.

- (iii) Aufgrund des Satzes über die Umkehrabbildung gibt es eine Umgebung  $U \in \mathcal{U}(0)$  in  $\mathbb{R}^n$  und einen Würfel  $W_\varepsilon^n \in \mathcal{U}(0)$ , so dass  $g|_U \in \text{Diff}^k(U, W_\varepsilon^n)$  gilt. Wir definieren  $\varphi'_n$  als Inverse von  $g|_U$  und definieren weiterhin  $\psi := f \circ \varphi'_n: W_\varepsilon^n \rightarrow \mathbb{R}^m$ . Dann gilt  $\text{rk } D\varphi'_n = n$  und somit folgt  $\text{rk } D\psi = \text{rk}(Df \circ D\varphi'_n) = \text{rk } Df = r$  in  $W_\varepsilon^n$ , wobei wir die Argumente weggelassen haben. Nach Definition von  $\psi$  folgt

$$(7.1) \quad \psi(y) = (y^1, \dots, y^r, h^{r+1}(y), \dots, h^m(y))$$

für alle  $y \in W_\varepsilon^n$ , wobei  $h^j$ ,  $r+1 \leq j \leq m$ ,  $C^k$ -Funktionen sind.

- (iv) Wir behaupten, dass die Funktionen  $h^j$  nur von  $(y^1, \dots, y^r)$  abhängen: Seien  $e_i$  bzw.  $e'_j$  die kanonischen Basisvektoren von  $\mathbb{R}^n$  bzw.  $\mathbb{R}^m$  und sei  $(a_i^j) = (a_i^j(y))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} = J_\psi(y)$  die Jacobimatrix von  $\psi$ . Dann folgt für alle  $1 \leq i \leq n$  nach Definition der Komponenten einer Matrix  $D\psi(y)\langle e_i \rangle = \sum_{j=1}^m a_i^j(y)e'_j$ .

Unter Benutzung der Darstellung (7.1) erhalten wir

$$D\psi(y)\langle e_i \rangle = e'_i \quad \text{für alle } 1 \leq i \leq r$$

und

$$D\psi(y)\langle e_i \rangle = \sum_{j=r+1}^m \frac{\partial h^j}{\partial y^i} e'_j \quad \text{für alle } r+1 \leq i \leq n.$$

Nun gilt  $\text{rk } D\psi(y) = r$ , die Vektoren  $e'_i$ ,  $1 \leq i \leq r$ , erzeugen bereits ein  $r$ -dimensionales Bild und die Vektoren  $e'_j$ ,  $r+1 \leq j \leq m$ , stehen orthogonal dazu. Folglich dürfen letztere keinen weiteren Beitrag zu  $R(D\psi(y))$  liefern. Dies bedeutet, dass

$$\frac{\partial h^j}{\partial y^i}(y) = 0 \quad \text{für alle } i, j \text{ mit } r+1 \leq i \leq n \text{ und } r+1 \leq j \leq m$$

für alle  $y \in W_\varepsilon^n$  gilt.

Schreiben wir  $W_\varepsilon^n = W_\varepsilon^r \times W_\varepsilon^{n-r}$  mit Elementen  $y = (y_1, y_2)$ , so haben wir  $D_{y_2} h^j = 0$  für alle  $r+1 \leq j \leq m$  gezeigt. Da die Würfel zusammenhängend sind, folgt  $h^j(y_1, y_2) = h^j(y_1) = h^j(y^1, \dots, y^r)$  für alle  $r+1 \leq j \leq m$ .

- (v) Wir definieren eine Funktion  $h: W_\varepsilon^r \times \mathbb{R}^{m-r} \rightarrow W_\varepsilon^r \times \mathbb{R}^{m-r}$  durch

$$h = (0, \dots, 0, h^{r+1}, \dots, h^m),$$

wobei sämtliche Komponenten nur von den ersten  $r$  Variablen  $(y^1, \dots, y^r) \in W_\varepsilon^r$  abhängen. Wir definieren weiterhin  $\alpha: W_\varepsilon^r \times \mathbb{R}^{m-r} \rightarrow W_\varepsilon^r \times \mathbb{R}^{m-r}$  durch  $\alpha(y) := y - h(y)$ . Zerlegen wir Vektoren  $y \in W_\varepsilon^r \times \mathbb{R}^{m-r}$  (ein anderes  $y$  als oben in  $W_\varepsilon^r \times W_\varepsilon^{n-r} = W_\varepsilon^n$ ) als  $y = (y_1, y_2)$ , so folgt  $\alpha(y_1, y_2) = (y_1, y_2) - h(y_1)$ . Die Abbildung  $\alpha$  ist von der Klasse  $C^k$ , wie wir direkt an der Definition von  $h$  ablesen können, auch injektiv, und für  $u = (u_1, u_2) \in \mathbb{R}^r \times \mathbb{R}^{m-r}$  gilt

$$D\alpha(y)\langle u_1, u_2 \rangle = (u_1, u_2) - D_{y_1} h(y)\langle u_1 \rangle.$$

In Matrixschreibweise mit  $h = (0, h_2)$  erhalten wir

$$D\alpha(y) = \begin{pmatrix} \text{id}_{\mathbb{R}^r} & 0 \\ D_{y_1} h_2(y) & \text{id}_{\mathbb{R}^{m-r}} \end{pmatrix}$$

und somit  $D\alpha(y) \in GL(m)$  für beliebige  $y \in W_\varepsilon^r \times \mathbb{R}^{m-r}$ .

Aus der Darstellung  $\alpha(y_1, y_2) = (y_1, y_2 - h_2(y_1))$  sieht man leicht, dass  $\alpha$  auch surjektiv ist.

Nach Korollar 7.2.7 ist  $\alpha$  somit ein Diffeomorphismus.

- (vi) Aus (7.1) und der Darstellung  $\alpha(y_1, y_2) = (y_1, y_2) - h(y_1)$  lesen wir direkt ab, dass

$$\alpha \circ \psi(y) = (y^1, \dots, y^r, 0, \dots, 0)$$

für alle  $y \in W_\varepsilon^n$  gilt. Insbesondere erhalten wir daraus  $\alpha \circ \psi(W_\varepsilon^n) \subset W_\varepsilon^m$ .

Wir definieren  $V := \alpha^{-1}(W_\varepsilon^m)$  und  $\varphi'_m := \alpha|_V$ . Dann erhalten wir durch Anwenden von  $\alpha^{-1}$  auf die gerade bewiesene Mengeninklusion  $\psi(W_\varepsilon^n) \subset V$ .

Nach Definition von  $\psi$  gilt  $\psi = f \circ \varphi'_n$ . Somit erhalten wir aus der obigen Darstellung für  $\alpha \circ \psi$

$$\varphi'_m \circ f \circ \varphi'_n(y) = \alpha \circ \psi(y) = (y^1, \dots, y^r, 0, \dots, 0) \in W_\varepsilon^r \times \mathbb{R}^{m-r}$$

für alle  $y \in W_\varepsilon^n$ .

Als Verknüpfung mit Streckungen definieren wir schließlich  $\varphi_n(y) := \varphi'_n(\varepsilon y)$  für  $y \in W^n$  und  $\psi_m(y) = \varepsilon^{-1} \varphi'_m(y)$ . Damit folgt die Behauptung.  $\square$

**Theorem 7.4.3** (Invarianz der Dimension). *Seien  $\Omega_n \subset \mathbb{R}^n$  und  $\Omega_m \subset \mathbb{R}^m$  offen. Sei  $f: \Omega_n \rightarrow \Omega_m$  ein Homöomorphismus. Dann gilt  $m = n$ .*

Wir können mit unseren momentanen Kenntnissen nur den Fall behandeln, dass  $f$  zusätzlich von der Klasse  $C^1$  ist. Wir kommen jedoch ohne die Voraussetzung aus, dass  $f$  ein  $C^1$ -Diffeomorphismus ist; in diesem Fall wäre die Aussage vergleichsweise trivial. Der allgemeine Fall wird beispielsweise in Algebraischer Topologie oder Differentialtopologie bewiesen.

*Beweis im  $C^1$ -Fall.* Da  $\text{rk } Df(x)$  nur Werte in  $\mathbb{N}_{\leq \min(m,n)}$  annimmt, gibt es  $x_0 \in \Omega_n$  mit

$$r = \text{rk } Df(x_0) = \max_{x \in \Omega_n} \text{rk } Df(x).$$

Dann gibt es aufgrund der Unterhalbstetigkeit des Ranges einer Matrix ein  $\varepsilon > 0$  mit  $\text{rk } Df(x) = r$  für alle  $x \in B_\varepsilon(x_0)$ . Aufgrund des Rangtheorems gibt es nun Diffeomorphismen  $\varphi_n$  und  $\varphi_m$ , so dass

$$\varphi_m \circ f \circ \varphi_n = (x^1, \dots, x^r, 0, \dots, 0) \in \mathbb{R}^m$$

für alle  $x \in W_1^n$  gilt. Da  $r$  der Rang einer  $m \times n$ -Matrix ist, gilt  $r \leq \min(m, n)$ .

Die Abbildung  $\varphi_m \circ f \circ \varphi_n$  ist injektiv. Dies erfordert  $r \geq n$ . Also gilt  $r = n$ .

Andererseits ist die Abbildung  $\varphi_m \circ f \circ \varphi_n$  als Verkettung eines Homöomorphismus mit Diffeomorphismen eine offene Abbildung. Dies erfordert  $r = m$ . Wir erhalten die Behauptung.  $\square$

## 7.5. Extrema mit Nebenbedingungen.

**Theorem 7.5.1** (Lagrangesche Multiplikatorregel). *Sei  $\Omega \subset \mathbb{R}^n$  offen. Sei  $f \in C^1(\Omega)$  und sei  $\Phi = (\Phi^i) \in C^1(\Omega, \mathbb{R}^m)$  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  $\bigcap_i \{\Phi^i = 0\}$  betrachten, so nimmt  $f$  in  $z_0$  ein lokales Extremum an,

und

(iii)  $\text{rk } D\Phi(z_0) = m$ , d. h. es gilt  $\dim \text{im } D\Phi(z_0) = m$ .

Dann gibt es  $\lambda = (\lambda_i)_i \in (\mathbb{R}^m)^*$ , so dass  $z_0$  ein kritischer Punkt der (nun auf ganz  $\Omega$  definierten) Funktion  $g: \Omega \rightarrow \mathbb{R}$

$$g := f + \sum_{i=1}^m \lambda_i \Phi^i$$

ist. Insbesondere gilt also

$$0 = \nabla g(z_0) = \nabla f(z_0) + \sum_{i=1}^m \lambda_i \nabla \Phi^i(z_0).$$

★ Auch hier gilt eine entsprechend technischere Banachraumvariante. Unter den in diesem Fall üblichen Voraussetzungen funktioniert auch hier der angegebene Beweis. Auf Details verzichten wir hier jedoch.

*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 Aufspaltung 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  $\{z \in B_r(z_0) : \Phi(z) = 0\} = \{(x, \varphi(x)) : x \in U\} \cap B_r(z_0)$  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 graph  $\varphi$  geschrieben und erhalten daher, dass die Funktion  $h: U \rightarrow \mathbb{R}$  mit  $h(x) := f(x, \varphi(x))$  in  $x_0 \in U$  ein lokales Extremum annimmt. Somit folgt nach Kettenregel

$$0 = Dh(x_0) = D_1f(x_0, \varphi(x_0)) + D_2f(x_0, \varphi(x_0)) \circ D\varphi(x_0).$$

Nun benutzen wir die Darstellung für  $D\varphi$  und erhalten mit  $\varphi(x_0) = y_0$

$$0 = D_1f(x_0, y_0) - D_2f(x_0, y_0) \circ (D_2\Phi(x_0, y_0))^{-1} \circ D_1\Phi(x_0, y_0).$$

Wir definieren nun  $\lambda := -D_2f(z_0) \circ (D_2\Phi(z_0))^{-1} \in (\mathbb{R}^m)^*$ . Somit können wir aus  $Dh = 0$  letztlich folgern, dass in  $z_0$

$$0 = D_1f + \sum_{j=1}^m \lambda_j D_1\Phi^j$$

gilt. Weiterhin folgt durch Umsortieren der Definition von  $\lambda$  in  $z_0$

$$0 = D_2f + \sum_{j=1}^m \lambda_j D_2\Phi^j.$$

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 7.5.2** (Eigenwerte symmetrischer Matrizen).

Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Nehme die Funktion  $f: \mathbb{S}^{n-1} \rightarrow \mathbb{R}$  mit

$$f(x) := \langle Ax, x \rangle$$

ihre 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  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  durch  $f(x) := \langle Ax, x \rangle$  definiert ist. Dann ist  $x_0$  ein Minimum von  $f$  unter der Nebenbedingung  $\Phi = 0$  mit  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$  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.  $\square$

### Bemerkung 7.5.3. $\star$

- (i) Betrachtet man im Beweis von Theorem 7.5.2  $g := f - \lambda\Phi$ , so erhält man direkt die optisch gewohntere Eigenwertgleichung  $Ax_0 = \lambda x_0$ .
- (ii) Aufgrund der Kompaktheit von  $\mathbb{S}^{n-1} \subset \mathbb{R}^n$  und der Stetigkeit von  $f$  ist klar, dass die Funktion  $f$  ihr Minimum auf  $\mathbb{S}^{n-1}$  annimmt.
- (iii) Dieses Minimum ist der kleinste Eigenwert von  $A$ . Dies sieht man leicht, da  $f$  in größeren Eigenwerten von  $A$  auch größere Werte annimmt.
- (iv) Indem man als zusätzliche Nebenbedingung noch fordert, dass die Vektoren senkrecht auf den ersten bereits gefundenen Eigenvektoren stehen, erlaubt es diese Methode auch, weitere Eigenwerte und Eigenvektoren zu finden.
- (v) Mit min-max-Methoden, siehe z. B. [5], erhält man, dass nicht nur der kleinste Eigenwert stetig von  $A$  abhängt, siehe das nachfolgende Korollar 7.5.4, sondern alle  $n$  Eigenwerte.
- (vi) Eigenvektoren hängen i. a. nicht stetig von der Matrix  $A$  ab.

**Korollar 7.5.4.** Sei  $X$  ein metrischer Raum und  $A: X \rightarrow \mathbb{R}^{n \times n}$  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 7.5.2 und der nachfolgenden Bemerkung genügt der Nachweis, dass  $x \mapsto \min_{\xi \in \mathbb{S}^{n-1}} \langle A(x)\xi, \xi \rangle$  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  $\|A - B\| \leq \varepsilon$  für  $|\xi| = 1$

$$\min_{\zeta \in \mathbb{S}^{n-1}} \langle A\zeta, \zeta \rangle \leq \langle A\xi, \xi \rangle = \langle B\xi, \xi \rangle + \underbrace{\langle (A - B)\xi, \xi \rangle}_{\leq \varepsilon} \leq \underbrace{\langle B\xi, \xi \rangle}_{=1} + \underbrace{\|A - B\| \cdot |\xi|^2}_{=1}.$$

Nun folgt zunächst  $\min_{\zeta \in \mathbb{S}^{n-1}} \langle A\zeta, \zeta \rangle \leq \langle B\xi, \xi \rangle + \varepsilon$  für alle  $\xi \in \mathbb{S}^{n-1}$ . Da  $\xi \in \mathbb{S}^{n-1}$  beliebig war, erhalten wir auch  $\min_{\zeta \in \mathbb{S}^{n-1}} \langle A\zeta, \zeta \rangle \leq \min_{\zeta \in \mathbb{S}^{n-1}} \langle B\zeta, \zeta \rangle + \varepsilon$ . Im obigen Beweis hätten wir die Rollen von  $A$  und  $B$  auch austauschen können. Somit folgt die behauptete Stetigkeit des Minimums.  $\square$

## 8. NUMERISCHE SUCHE NACH EINEM MINIMUM ★

**8.1. Theorie zur Numerik.** Die Idee bei dieser Suche nach einem Minimum ist es, Schritt für Schritt jeweils in Richtung des negativen Gradienten weiter zu gehen, bis man bei einem Minimum bzw. einem kritischen Punkt landet. Dabei muss man jeweils weit genug gehen, darf aber auch nicht zu weit gehen. Man kann dabei keine Konvergenz erwarten (Kugelbahn an einer umgedrehten Parabel), außer man hat weitere Zusatzvoraussetzungen wie Analytizität oder Konvexität. Das Folgende ist ein rein theoretisches Resultat, lässt sich jedoch (bis auf die Teilfolgenauswahl) direkt auch numerisch umsetzen.

**Theorem 8.1.1.** Sei  $u \in C^1(\mathbb{R}^n)$  mit  $u(x) \rightarrow \infty$  für  $|x| \rightarrow \infty$ . Sei  $x_0 \in \mathbb{R}^n$  beliebig. Definiere induktiv

$$x_{k+1} := x_k - \frac{1}{m} \cdot \nabla u(x_k)$$

bzw.  $x_{k+1} := x_k$  für  $\nabla u(x_k) = 0$  für das minimale  $m = m_k \in \mathbb{N}_{>0}$  mit

$$u\left(x_k - \frac{1}{m} \cdot \nabla u(x_k)\right) \leq u(x_k) - \frac{1}{3} \cdot \frac{1}{m} \cdot |\nabla u(x_k)|^2.$$

Dann existiert eine Teilfolge  $(x_{k_l})_l$ , die gegen einen kritischen Punkt von  $u$  konvergiert, d. h. eine Teilfolge von  $(x_k)_k$  mit

$$\nabla u\left(\lim_{l \rightarrow \infty} x_{k_l}\right) = 0.$$

*Beweis.*

- (i) Die Folge ist wohldefiniert: Sei  $x \in \mathbb{R}^n$  mit  $\nabla u(x) \neq 0$ . Wir wollen zeigen, dass wir von jedem Punkt – wir betrachten hier  $x$  statt  $x_k$  oben – ausgehend eine nichtleere Menge an  $m \in \mathbb{N}_{>0}$  finden, so dass die obige  $m$  definierende Ungleichung erfüllt wird. Damit kann dann  $x_k$  über das minimale  $m$  wohldefiniert definiert werden. Da  $u \in C^1$  ist, gibt es  $\zeta > 0$  mit

$$\langle \nabla u(x), \nabla u(y) \rangle \geq \frac{1}{2} \cdot |\nabla u(x)|^2$$

für alle  $y \in \overline{B_\zeta(x)}$ . Setze  $\xi := \nabla u(x)$ . Es folgt für  $0 \leq t \leq \frac{\zeta}{|\nabla u(x)|}$

$$\begin{aligned} u(x - t\xi) - u(x) &= \int_0^1 \frac{d}{d\tau} u(x - \tau t\xi) d\tau \\ &= - \int_0^1 \langle \nabla u(x - \tau t\xi), \nabla u(x) \rangle d\tau \cdot t \\ &\leq -\frac{1}{2} \cdot |\nabla u(x)|^2 \cdot 1 \cdot t. \end{aligned}$$

Dabei haben wir

$$|x - \tau t\xi - x| \leq \frac{\zeta}{|\nabla u(x)|} \cdot |\xi| \leq \zeta$$

benutzt. Daher ist die gewünschte Ungleichung für hinreichend große  $m \in \mathbb{N}_{>0}$  erfüllt.

- (ii) Die Folge  $(u(x_k))_k$  ist offensichtlicherweise monoton fallend. Daher ist die Folge nach oben beschränkt. Aufgrund der Wachstumsbedingung an  $u$  folgt auch die Beschränktheit der Folge  $(x_k)_k \subset \mathbb{R}^n$ . Wir können daher ein  $R > 0$  mit  $x_k \in B_R(0)$  für alle  $k \in \mathbb{N}$  wählen.

(iii) Sei  $\varepsilon > 0$  beliebig. Betrachte die Menge

$$K_\varepsilon \equiv K := \left\{ x \in \overline{B_R(0)} : |\nabla u(x)| \geq \varepsilon \right\}.$$

Wir behaupten, dass es ein Folgenglied  $x_{k_\varepsilon}$  mit  $x_{k_\varepsilon} \notin K_\varepsilon$  gibt. Wir zeigen die in (iv) in einer Weise, dass der Index  $k_\varepsilon$  größer als ein vorgegebener Wert ist.

Induktiv erhalten wir für  $l \geq 1$  Indices  $k_l \in \mathbb{N}$  mit  $k_{l+1} > k_l$ ,  $x_{k_l} \in \overline{B_R(0)} \setminus K_{\frac{1}{l}}$  und somit  $|\nabla u(x_{k_l})| \leq \frac{1}{l}$ .

Es ist klar, dass eine Teilfolge von  $(x_{k_l})_l$  gegen einen kritischen Punkt von  $u$  konvergiert.

(iv) Zur Behauptung aus (iv): Wir führen einen Widerspruchsbeweis. Angenommen, es gibt ein festes  $\varepsilon > 0$  und  $k_0 \in \mathbb{N}$ , so dass  $x_k \in K_\varepsilon$  für alle  $k \geq k_0$  gilt.

Die Menge  $K_\varepsilon$  ist kompakt.

Zu  $x \in K_\varepsilon$  gibt es  $\delta > 0$  mit

$$|\nabla u(x) - \nabla u(y)| \leq \frac{1}{10} \cdot |\nabla u(x)|$$

für alle  $y \in B_\delta(x)$ . Wir bemerken, dass die rechte Seite nach Definition von  $K_\varepsilon$  strikt positiv ist. Seien nun  $y, z \in B_\delta(x)$ . Dann folgt

$$\begin{aligned} & |\langle \nabla u(y), \nabla u(z) \rangle - |\nabla u(y)|^2| \\ &= |\langle \nabla u(y), \nabla u(z) - \nabla u(y) \rangle| \\ &\leq |\nabla u(y)| \cdot \{|\nabla u(z) - \nabla u(x)| + |\nabla u(x) - \nabla u(y)|\} \\ &\leq \frac{2}{10} \cdot |\nabla u(y)| \cdot |\nabla u(x)|. \end{aligned}$$

Den zweiten Faktor wollen wir weiter abschätzen. Es gilt

$$\begin{aligned} |\nabla u(x)| &= |\nabla u(x) - \nabla u(y) + \nabla u(y)| \\ &\leq |\nabla u(x) - \nabla u(y)| + |\nabla u(y)| \\ &\leq \frac{1}{10} \cdot |\nabla u(x)| + |\nabla u(y)|. \end{aligned}$$

Hier addieren wir  $-\frac{1}{10} \cdot |\nabla u(x)|$  und multiplizieren anschließend das Ergebnis mit  $\frac{10}{9}$ . Dies ergibt

$$|\nabla u(x)| \leq \frac{10}{9} \cdot |\nabla u(y)|.$$

Also folgt

$$\begin{aligned} |\nabla u(y)|^2 - \langle \nabla u(y), \nabla u(z) \rangle &\leq \frac{2}{10} \cdot |\nabla u(y)| \cdot |\nabla u(x)| \\ &\leq \frac{2}{10} \cdot \frac{10}{9} \cdot |\nabla u(y)|^2 \\ &= \frac{2}{9} \cdot |\nabla u(y)|^2. \end{aligned}$$

Wir erhalten

$$\frac{1}{2} \cdot |\nabla u(y)|^2 \leq \frac{7}{9} \cdot |\nabla u(y)|^2 \leq \langle \nabla u(y), \nabla u(z) \rangle$$

für alle  $y, z \in B_\delta(x)$ . Wir setzen  $\delta(x) := \delta > 0$ .

Endlich viele Bälle der Form  $B_{\frac{\delta(x)}{2}}(x)$  überdecken die kompakte Menge  $K_\varepsilon$ . Wir betrachten die zugehörigen endlich vielen  $x_i$ 's und  $\delta_i = \delta(x_i)$ 's. Setze

$\zeta := \min_i \frac{1}{2} \delta_i > 0$ . Somit gelten die Überlegungen aus (i) unabhängig von  $x \in K_\varepsilon$ , also die Abschätzung für die Funktionswerte:

$$u(x - t \cdot \nabla u(x)) \leq u(x) - \frac{1}{2} \cdot t \cdot |\nabla u(x)|^2$$

für alle  $t$  mit  $0 \leq t \leq \frac{\zeta}{|\nabla u(x)|} \leq \frac{\zeta}{\varepsilon}$ . Ohne Einschränkung dürfen wir, gegebenenfalls nach Verkleinerung von  $\zeta > 0$ , annehmen, dass  $\frac{\zeta}{\varepsilon} \leq 1$  gilt.

Somit ist die  $m$  charakterisierende Gleichung für  $m \leq m_0$  mit  $\frac{1}{m_0} \leq \frac{\zeta}{\varepsilon}$  erfüllt. Wir dürfen  $m_0$  minimal mit  $\frac{1}{m_0} \leq \frac{\zeta}{\varepsilon}$  wählen. Somit gilt

$$\frac{\zeta}{2\varepsilon} \leq \frac{1}{m_0} \leq \frac{1}{m}$$

und wir erhalten

$$u(x_{k+1}) \leq u(x_k) - \frac{1}{2} \cdot \frac{\zeta}{2\varepsilon} \cdot \varepsilon^2.$$

Auf der rechten Seite wird nun eine feste positive Konstante abgezogen, die insbesondere nicht mehr von  $k$  abhängt. Per Induktion folgt daher

$$u(x_k) \leq u(x_0) - k \cdot \frac{\zeta \cdot \varepsilon}{4} \rightarrow -\infty$$

für  $k \rightarrow \infty$ . Dies widerspricht aber

$$u(x_k) \geq \inf_{B_R(0)} u > -\infty.$$

Somit verlässt die Folge  $x_k$  die Menge  $K_\varepsilon$ . □

## 9. ZUSATZMATERIAL ZUR DIFFERENZIERBARKEIT ★

### 9.1. Zusatzmaterial. Ist komplizierter und daher herauskommentiert.

#### 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*, 2024, Skript zur Vorlesung.
4. Oliver C. Schnürer, *Lineare Algebra I*, 2010, Skript zur Vorlesung.
5. Oliver C. Schnürer, *Lineare Algebra II*, 2013, Skript zur Vorlesung.
6. Wikipedia, <https://www.wikipedia.org>.
7. 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`