

Kapitel 10

Differentialrechnung von Funktionen mehrerer Veränderlicher

10.1 Ableitungen von $f : X \rightarrow Y$

Definition 10.1. (Ableitung) Seien X und Y normierte Vektorräume. Eine Abbildung f von einer offenen Menge $U \subset X$ nach Y heißt im Punkt $x_0 \in U$ differenzierbar, wenn es ein $A \in \mathcal{L}(X, Y)$ gibt, so dass die folgende Abbildung in x_0 stetig ist:

$$U \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} \frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} & \text{für } x \neq x_0 \\ 0 & \text{für } x = x_0. \end{cases}$$

A heißt Ableitung von f bei x_0 und wird mit $f'(x_0)$ oder $\frac{df}{dx}(x_0)$ bezeichnet. Wenn A und B beide diese Bedingung erfüllen, dann gibt es für alle $\epsilon > 0$ ein $\delta > 0$ mit

$$\frac{\|(A - B)(x - x_0)\|}{\|x - x_0\|} \leq \frac{\|A(x - x_0) - f(x) + f(x_0)\|}{\|x - x_0\|} + \frac{\|f(x) - f(x_0) - B(x - x_0)\|}{\|x - x_0\|} < 2\epsilon$$

für alle $x \in B(x_0, \delta) \setminus \{x_0\}$. Mit $x - x_0 = \delta y \in X$ für $0 < \|y\| < 1$ folgt $\|(A - B)y\| = \frac{\|(A - B)\delta y\|}{\delta} < 2\epsilon\|y\| < 2\epsilon$ und $\|A - B\| \leq 2\epsilon$. Weil das für alle $\epsilon > 0$ gilt, ist $A = B$.

Satz 10.2. Sei $f : U \subset X \rightarrow Y$ in $x_0 \in U$ differenzierbar. Dann ist f in x_0 stetig.

Beweis: Wegen $f(x) - f(x_0) = A(x - x_0) + (f(x) - f(x_0) - A(x - x_0))$ folgt aus der Differenzierbarkeit von f in x_0 , dass es ein $\delta > 0$ gibt, so dass $\|f(x) - f(x_0)\| \leq (\|A\| + 1)\|x - x_0\|$ gilt für alle $\|x - x_0\| < \delta$. Dann ist f in x_0 auch stetig. **q.e.d.**

Beispiel 10.3. (i) Sei f konstant. Dann ist
$$\frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} = 0 \quad \text{für } x \neq x_0.$$
 Also ist f differenzierbar mit $f'(x_0) = 0$.

(ii) Für $f \in \mathcal{L}(X, Y)$ und $x \neq x_0 \in X$ gilt
$$\frac{\|f(x) - f(x_0) - f(x - x_0)\|}{\|x - x_0\|} = 0.$$

Also ist f in x_0 differenzierbar und $f'(x_0) = f$.

(iii) Die Abbildung $Y \rightarrow \mathcal{L}(\mathbb{R}, Y)$, $y \mapsto$ Multiplikation mit y besitzt offenbar die Umkehrabbildung $\mathcal{L}(\mathbb{R}, Y) \rightarrow Y$, $A \mapsto A(1)$ und ist ein isometrischer Isomorphismus von normierten Vektorräumen. Deshalb können wir die Ableitungen von differenzierbaren $f : (a, b) \rightarrow Y$ auf $(a, b) \subset \mathbb{R}$ auch durch $f' : (a, b) \rightarrow Y$ beschreiben. Für $x \neq x_0$ gilt

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

Dann sind für $f : (a, b) \rightarrow \mathbb{K}$ und $x_0 \in (a, b)$ die Definitionen 7.1 und 10.1 äquivalent.

Satz 10.4. (i) Sei $f, g : U \subset X \rightarrow Y$ in $x_0 \in U$ differenzierbar. Dann sind $f + g$ und $\lambda \cdot f$ für alle $\lambda \in \mathbb{K}$ in x_0 differenzierbar und es gilt

$$(f + g)'(x_0) = f'(x_0) + g'(x_0) \quad \text{bzw.} \quad (\lambda f)'(x_0) = \lambda f'(x_0).$$

(ii) Seien $f, g : U \subset X \rightarrow Y$ in x_0 differenzierbar und Y eine normierte Algebra. Dann ist $f \cdot g$ in x_0 differenzierbar und es gilt (Leibnizregel)

$$\begin{aligned} (f \cdot g)'(x_0) &= f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \quad \text{mit} \\ (f'(x_0) \cdot g(x_0))(x) &= f'(x_0)(x) \cdot g(x_0) \quad \text{und} \quad (f(x_0) \cdot g'(x_0))(x) = f(x_0) \cdot g'(x_0)(x). \end{aligned}$$

(iii) Seien $f : U \subset X \rightarrow Y$ in $x_0 \in U$ und $g : V \subset Y \rightarrow Z$ in $f(x_0) \in V$ differenzierbar mit $f[U] \subset V$. Dann ist $g \circ f$ im Punkt x_0 differenzierbar mit

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

Beweis:(i) Mit den drei Eigenschaften von Normen folgt (i) für $x \neq x_0$ aus

$$\begin{aligned} & \frac{\|f(x) + g(x) - f(x_0) - g(x_0) - f'(x_0)(x - x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ & \leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} + \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}. \\ & \frac{\|\lambda f(x) - \lambda f(x_0) - \lambda f'(x_0)(x - x_0)\|}{\|x - x_0\|} = |\lambda| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}. \end{aligned}$$

(ii) Weil in einer normierten Algebra $\|AB\| \leq \|A\| \cdot \|B\|$ gilt, folgt für $x \neq x_0$

$$\begin{aligned} & \frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x - x_0)g(x_0) - f(x_0)g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ & \leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \|g(x)\| + \|f'(x_0)\| \cdot \|g(x) - g(x_0)\| + \\ & \quad + \|f(x_0)\| \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}. \end{aligned}$$

Weil g in x_0 stetig ist, gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass aus $\|x - x_0\| < \delta$ folgt $\|g(x) - g(x_0)\| < \epsilon$ bzw. $\|g(x)\| \leq \|g(x_0)\| + \epsilon$. Dann folgt (ii).

$$\begin{aligned}
 \text{(iii)} \quad & \text{Aus Satz 9.61 folgt für } x \neq x_0 \quad \frac{\|(g \circ f)(x) - (g \circ f)(x_0) - g'(f(x_0))f'(x_0)(x - x_0)\|}{\|x - x_0\|} \\
 & \leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \\
 & \quad + \frac{\|g'(f(x_0))(f(x) - f(x_0)) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \\
 & \leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \\
 & \quad \cdot \left(\|f'(x_0)\| + \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} \right) \\
 & \quad + \|g'(f(x_0))\| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}.
 \end{aligned}$$

Ersetze hierbei für $f(x) = f(x_0)$ die beiden Differenzenquotienten von g wie in der Definition von $g'(f(x_0))$ durch Null. Aus Satz 10.2 und Korollar 9.32 folgt (iii). **q.e.d.**

10.2 Schrankensatz

Wir verallgemeinern in diesem Abschnitt den Schrankensatz auf differenzierbare Abbildungen zwischen normierten Vektorräumen.

Lemma 10.5. *Seien f eine stetige Abbildung von einem kompakten Intervall $[a, b]$ mit $a < b$ in einen normierten Vektorraum Y und ϕ eine stetige reelle Funktion auf $[a, b]$. Wenn im Komplement einer abzählbaren Teilmenge von (a, b) sowohl f als auch ϕ differenzierbar sind und dort gilt $\|f'\| \leq \phi'$, dann gilt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$.*

Beweis: Sei $(x_n)_{n \in \mathbb{N}}$ eine Abzählung der Punkte in $[a, b]$, an denen entweder f oder ϕ nicht differenzierbar ist oder $\|f'\| > \phi'$. Mit kontinuierlicher Induktion zeigen wir

$$\|f(x) - f(a)\| \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n} \quad \text{für alle } \epsilon > 0 \text{ und } x \in [a, b].$$

Sei also $\epsilon > 0$ und A_ϵ die Menge aller $y \in [a, b]$, so dass diese Ungleichung für alle $x \in [a, y]$ gilt. Wegen der Stetigkeit von f und ϕ gilt auch für $y = \sup A_\epsilon$

$$\begin{aligned}
 \|f(y) - f(a)\| & \leq \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sup \left\{ \sum_{x_n < x} 2^{-n} \mid x \in (a, y) \right\} \\
 & = \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n}.
 \end{aligned}$$

Deshalb ist A_ϵ ein Intervall von der Form $A_\epsilon = [a, y]$. Wenn $y \in (a, b)$ und f und ϕ in y differenzierbar sind und $\|f'(y)\| \leq \phi'(y)$ gilt, dann gibt es ein $\delta > 0$, so dass

$$\phi'(y) - \frac{\epsilon}{2} < \frac{\phi(x) - \phi(y)}{x - y} < \phi'(y) + \frac{\epsilon}{2} \quad \text{und} \quad \frac{\|f(x) - f(y) - f'(y)(x - y)\|}{|x - y|} < \frac{\epsilon}{2}$$

für alle $x \in (y - \delta, y) \cup (y, y + \delta)$ gilt. Dann folgt für $x \in (y, y + \delta)$

$$\begin{aligned} \|f(x) - f(y)\| &\leq \|f'(y)(x - y)\| + \|f(x) - f(y) - f'(y)(x - y)\| \\ &< \left(\|f'(y)\| + \frac{\epsilon}{2}\right) |x - y| \leq \left(\phi'(y) - \frac{\epsilon}{2} + \epsilon\right) (x - y) \\ &< \left(\frac{\phi(x) - \phi(y)}{x - y} + \epsilon\right) (x - y) = \phi(x) - \phi(y) + \epsilon(x - y) \quad \text{und} \\ \|f(x) - f(a)\| &\leq \|f(x) - f(y)\| + \|f(y) - f(a)\| \\ &< \phi(x) - \phi(y) + \epsilon(x - y) + \phi(y) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \\ &\leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n}. \end{aligned}$$

Woraus $y + \delta \in A_\epsilon$ folgt, im Widerspruch zu $y = \sup A_\epsilon$. Wenn es andererseits ein $N \in \mathbb{N}$ gibt mit $x_N = y < b$, dann gibt es ein $\delta > 0$, so dass aus $x \in (y, y + \delta)$ folgt

$$\|f(x) - f(y)\| - (\phi(x) - \phi(y)) < \epsilon 2^{-N}.$$

Dann folgt für dieselben x wieder $\|f(x) - f(a)\| \leq \|f(x) - f(y)\| + \|f(y) - f(a)\|$

$$< \epsilon 2^{-N} + \phi(x) - \phi(a) + \epsilon(y - a) + \epsilon \sum_{x_n < y} 2^{-n} \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n}.$$

Also gilt wieder $y + \delta \in A_\epsilon$, was $y = \sup A_\epsilon$ widerspricht. Dann muß aber $\sup A_\epsilon = b$ gelten. Weil das für alle $\epsilon > 0$ gilt, folgt auch $\|f(b) - f(a)\| \leq \phi(b) - \phi(a)$. **q.e.d.**

Korollar 10.6. (Schranksatz) Sei f eine stetige Abbildung von einer offenen Teilmenge U des normierten Vektorraumes X in den normierten Vektorraum Y . Wenn f im Komplement einer abzählbaren Teilmenge S von $D = \{(1 - t)a + tb \mid t \in [0, 1]\} \subset U$ differenzierbar ist mit $a, b \in U$, und die Ableitung auf $D \setminus S$ beschränkt ist, dann gilt

$$\begin{aligned} \|f(b) - f(a)\| &\leq \|b - a\| \sup\{\|f'(x)\| \mid x \in D \setminus S\} \quad \text{und} \\ \|f(b) - f(a) - A(b - a)\| &\leq \|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} \quad \text{für } A \in \mathcal{L}(X, Y). \end{aligned}$$

Auf konvexen Mengen ist jede obere Schranke an $\|f'\|$ eine Lipschitzkonstante von f .

Beweis: Die Abbildung $x : \mathbb{R} \rightarrow X$, $t \mapsto x(t) = (t - 1)a + tb = a + t(b - a)$ ist wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 (i) differenzierbar mit $x'(t) : s \mapsto s(b - a)$

als Element von $\mathcal{L}(\mathbb{R}, X)$ bzw. mithilfe von Beispiel 10.3 (iii) mit $x'(t) = (b - a) \in X$. Wegen Satz 10.4 (iii) ist dann auch die Abbildung $t \mapsto f(x(t))$ bei den $t \in [0, 1]$ differenzierbar, für die f bei $x(t) \in U$ differenzierbar ist, mit der Ableitung $(f \circ x)'(t) : s \mapsto sf'(x(t))(b - a)$ bzw. mithilfe von Beispiel 10.3 (iii) $(f \circ x)'(t) = f'(x(t))(b - a) \in Y$. Dann folgt die erste Behauptung aus Lemma 10.5 mit den beiden Funktionen

$$f \circ x : [0, 1] \rightarrow Y, \quad t \mapsto f(x(t)), \quad \phi : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto t\|b - a\| \sup\{\|f'(x)\| \mid x \in D \setminus S\}.$$

Die zweite Behauptung folgt aus diesem Lemma mit den beiden Funktionen

$$t \mapsto f(x(t)) - tA(b - a) \in Y, \quad t \mapsto t\|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} \in \mathbb{R}. \quad \mathbf{q.e.d.}$$

10.3 Partielle Ableitungen

Definition 10.7. Eine Abbildung f von einer offenen Teilmenge U eines normierten Vektorraumes X in einen normierten Vektorraum Y heißt stetig differenzierbar, wenn

- (i) f in allen $x_0 \in U$ differenzierbar ist, und
- (ii) die Abbildung $f' : U \rightarrow \mathcal{L}(X, Y)$, $x \mapsto f'(x)$ stetig ist.

Wegen Beispiel 10.3 (iii) stimmt im Falle von $X = Y = \mathbb{R}$ diese Definition mit der Definition von stetig differenzierbaren reellen Funktionen auf Intervallen in $\mathbb{R}$ überein.

Definition 10.8. (partielle Ableitung) Sei f eine Funktion von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes der normierten Vektorräume X_1 und X_2 in den normierten Vektorraum Y . Dann heißt f im Punkt $(x_1, x_2) \in U \subset X_1 \times X_2$ partiell differenzierbar, falls die Abbildung $x \mapsto f(x, x_2)$ im Punkt $x = x_1$, und die Abbildung $x \mapsto f(x_1, x)$ im Punkt $x = x_2$ differenzierbar ist. Die Ableitungen heißen partielle Ableitungen an der Stelle (x_1, x_2) und werden mit $\frac{\partial f}{\partial x_1}(x_1, x_2)$ und $\frac{\partial f}{\partial x_2}(x_1, x_2)$ bezeichnet. Allgemeiner heißt eine Abbildung von einer offenen Menge $U \subset X_1 \times \dots \times X_n$ eines n -fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum Y im Punkt $(x_1, \dots, x_n) \in U$ partiell differenzierbar, wenn für $i = 1, \dots, n$ die Abbildungen $x \mapsto f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$ bei $x = x_i$ differenzierbar sind.

Die wichtigsten Beispiele sind reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$. Für jedes $A \in \mathcal{L}(X_1 \times X_2, Y)$ sind die beiden folgenden Abbildungen stetig und linear

$$A_1 : X_1 \rightarrow Y, \quad x_1 \mapsto A((x_1, 0)) \quad \text{und} \quad A_2 : X_2 \rightarrow Y, \quad x_2 \mapsto A((0, x_2)).$$

Weil für $x \in X_1 \setminus \{x_1\}$ mit $(x, x_2) \in U$ bzw. $x \in X_2 \setminus \{x_2\}$ mit $(x_1, x) \in U$

$$\begin{aligned} \frac{\|f(x, x_2) - f(x_1, x_2) - A((x, x_2) - (x_1, x_2))\|}{\|(x, x_2) - (x_1, x_2)\|} &= \frac{\|f(x, x_2) - f(x_1, x_2) - A_1(x - x_1)\|}{\|x - x_1\|} \\ \frac{\|f(x_1, x) - f(x_1, x_2) - A((x_1, x) - (x_1, x_2))\|}{\|(x_1, x) - (x_1, x_2)\|} &= \frac{\|f(x_1, x) - f(x_1, x_2) - A_2(x - x_2)\|}{\|x - x_2\|} \end{aligned}$$

gilt, folgt aus den Definitionen der folgende Satz:

Satz 10.9. Eine im Punkt $(x_1, x_2) \in U$ differenzierbare Funktion f von einer offenen Teilmenge $U \subset X_1 \times X_2$ des kartesischen Produktes zweier normierter Vektorräume in einen normierten Vektorraum Y ist in (x_1, x_2) auch partiell differenzierbar. **q.e.d.**

Beispiel 10.10. Die Umkehrung dieses Satzes gilt nicht:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \text{besitzt die partiellen Ableitungen}$$

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \frac{\partial f}{\partial y} = \begin{cases} \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases}$$

Für alle $r \in (0, \infty)$ und alle $\phi \in \mathbb{R}$ gilt $f(r \cos \phi, r \sin \phi) = 2 \sin \phi \cos \phi = \sin(2\phi)$, und deshalb $\lim_{r \rightarrow 0+} f(r \cos \phi, r \sin \phi) = \sin(2\phi)$. Also ist f im Punkt $(x, y) = 0$ nicht stetig.

Aber es gilt folgende Umkehrung.

Satz 10.11. Seien X_1, X_2, Y normierte Vektorräume, $U \subset X_1 \times X_2$ offen und $f : U \rightarrow Y$ bei $x = (x_1, x_2) \in U$ partiell differenzierbar. Wenn eine der partiellen Ableitungen

$$\frac{\partial f}{\partial x_1} : U \rightarrow \mathcal{L}(X_1, Y) \quad \text{oder} \quad \frac{\partial f}{\partial x_2} : U \rightarrow \mathcal{L}(X_2, Y)$$

auf U existiert und bei x stetig ist, dann ist f bei x differenzierbar. Also ist f genau dann auf U stetig differenzierbar, wenn f auf U stetig partiell differenzierbar ist.

Beweis: Wenn $A_1 \in \mathcal{L}(X_1, Y)$ und $A_2 \in \mathcal{L}(X_2, Y)$, dann sind auch die Abbildungen

$$X_1 \times X_2 \rightarrow Y, \quad (x_1, x_2) \mapsto A_1(x_1) \text{ bzw. } X_1 \times X_2 \rightarrow Y, \quad (x_1, x_2) \mapsto A_2(x_2)$$

lineare und stetige Abbildungen in $\mathcal{L}(X_1 \times X_2, Y)$. Also ist

$$A_1 \times A_2 : X_1 \times X_2 \rightarrow Y, \quad (x_1, x_2) \mapsto A_1(x_1) + A_2(x_2)$$

eine stetige lineare Abbildung in $\mathcal{L}(X_1 \times X_2, Y)$. Umgekehrt sind für jede stetige lineare Abbildung $A \in \mathcal{L}(X_1 \times X_2, Y)$ die Abbildungen $A_1 : X_1 \rightarrow Y, x_1 \mapsto A((x_1, 0))$ und $A_2 : X_2 \rightarrow Y, x_2 \mapsto A((0, x_2))$ stetig und linear. Und es gilt

$$\|A((x_1, x_2))\| = \|A((x_1, 0)) + A((0, x_2))\| \leq \|A((x_1, 0))\| + \|A((0, x_2))\|.$$

Aufgrund der Definition der Norm von $\mathcal{L}(X_1 \times X_2, Y)$ folgt dann

$$\|A_1\| \leq \|A\| \quad \|A_2\| \leq \|A\| \quad \|A\| \leq \|A_1\| + \|A_2\| \leq 2\|A\|.$$

Also ist die Abbildung $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y) \rightarrow \mathcal{L}(X_1 \times X_2, Y)$, $(A_1, A_2) \mapsto A$ eine bijektive Abbildung von normierten Vektorräumen und die beiden Normen von $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y)$ und $\mathcal{L}(X_1 \times X_2, Y)$ sind bezüglich dieser Identifikation äquivalent. Daraus folgt, dass für jede stetig differenzierbare Funktion $f : U \rightarrow Y$, die beiden partiellen Ableitungen $\frac{\partial f}{\partial x_1} : U \rightarrow \mathcal{L}(X_1, Y)$ und $\frac{\partial f}{\partial x_2} : U \rightarrow \mathcal{L}(X_2, Y)$ stetig sind.

Wenn umgekehrt f in (x_1, x_2) partiell differenzierbar ist, dann folgt

$$\begin{aligned} \left\| f(y_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| &\leq \\ &\leq \left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| + \\ &\quad + \left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\|. \end{aligned}$$

Wenn $\frac{\partial f}{\partial x_1}$ auf U existiert und bei (x_1, x_2) stetig ist, dann gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass

$$\left\| \frac{\partial f(z_1, z_2)}{\partial x_1} - \frac{\partial f(x_1, x_2)}{\partial x_1} \right\| < \epsilon$$

für $(z_1, z_2) \in B((x_1, x_2), \delta)$ gilt. Aus Korollar 10.6 folgt für $(y_1, y_2) \in B((x_1, x_2), \delta)$

$$\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| < \epsilon \|y_1 - x_1\|.$$

Weil $\frac{\partial f}{\partial x_2}$ in (x_1, x_2) existiert, gibt es auch ein $\delta' > 0$, so dass für $y_2 \in B(x_2, \delta')$ folgt

$$\left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| < \epsilon \|y_2 - x_2\|.$$

Dann folgt für $(y_1, y_2) \in B((x_1, x_2), \min\{\delta, \delta'\})$ auch

$$\|f(y_1, y_2) - f(x_1, x_2) - f'(x_1, x_2)((y_1, y_2) - (x_1, x_2))\| < \epsilon(\|y_1 - x_1\| + \|y_2 - x_2\|),$$

wobei $f'(x_1, x_2) = \frac{\partial f(x_1, x_2)}{\partial x_1} \times \frac{\partial f(x_1, x_2)}{\partial x_2}$ durch die partiellen Ableitungen gegeben ist. Also ist f differenzierbar, und mit den partiellen Ableitungen stetig differenzierbar. **q.e.d.**

Durch mehrmaliges Anwenden erhalten wir dann auch die entsprechende Aussage für Abbildungen von offenen Teilmengen U des n -fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum. Unsere wichtigsten Beispiele sind wieder reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$.

Beispiel 10.12. (i) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$

$$\frac{\partial f(x, y)}{\partial x} = \begin{cases} \frac{2y(y^2-x^2)}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases} \quad \frac{\partial f(x, y)}{\partial y} = \begin{cases} \frac{2x(x^2-y^2)}{(x^2+y^2)^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0 \end{cases}$$

Für $y = 0$ ist $\frac{\partial f}{\partial x} = 0$ und für $x = 0$ ist $\frac{\partial f}{\partial y} = 0$, so dass diese partiellen Ableitungen für alle $(x, y) \in \mathbb{R}^2$ existieren. Allerdings sind sie in keiner Umgebung von $(x, y) = (0, 0)$ beschränkt, und deshalb auch nicht stetig. Wir hatten schon im Beispiel 10.10 gesehen, dass f bei $(0, 0)$ nicht stetig und deshalb auch nicht differenzierbar ist.

$$(ii) \quad f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{x^3 - y^3}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Offenbar gilt $|f(x, y)| \leq \frac{|x|^3}{x^2 + y^2} + \frac{|y|^3}{x^2 + y^2} \leq |x| + |y|$. Also ist f stetig.

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{3x^2(x^2 + y^2) - 2x(x^3 - y^3)}{(x^2 + y^2)^2} = \frac{x(x^3 + 3xy^2 + 2y^3)}{(x^2 + y^2)^2} & \text{für } (x, y) \neq 0 \\ 1 & \text{für } (x, y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} -\frac{3y^2(x^2 + y^2) + 2y(x^3 - y^3)}{(x^2 + y^2)^2} = -\frac{y(y^3 + 3yx^2 + 2x^3)}{(x^2 + y^2)^2} & \text{für } (x, y) \neq 0 \\ -1 & \text{für } (x, y) = 0. \end{cases}$$

Wegen $\frac{\partial f(x, 0)}{\partial x} = 1$ und $\frac{\partial f(0, y)}{\partial y} = -1$ ist f partiell differenzierbar. In Beispiel 10.14 (iii) werden wir sehen, dass f in $(0, 0)$ nicht differenzierbar ist.

(iii) Alle Polynome in endlich vielen Variablen sind partiell unendlich oft stetig differenzierbar, und deshalb differenzierbar.

Definition 10.13. (Richtungsableitung) Für eine Funktion $f : U \rightarrow Y$ von einer offenen Teilmenge U eines normierten Vektorraumes X in einem normierten Vektorraum Y ist die Richtungsableitung in $x_0 \in U$ in Richtung $x_1 \in X$ die Ableitung bei $t = 0$ von

$$(-\epsilon, \epsilon) \rightarrow Y, \quad t \mapsto f(x_0 + tx_1).$$

Wie in Beispiel 10.3 (iii) identifizieren wir dabei $\mathcal{L}(\mathbb{R}, Y)$ durch $A \mapsto A(1)$ mit Y .

Beispiel 10.14. (i) Sei $U \subset X$ offen und $f : U \rightarrow Y$ in $x_0 \in U$ differenzierbar. Für $x_1 \in X$ ist die Abbildung $x : \mathbb{R} \rightarrow X$, $t \mapsto x(t) = x_0 + tx_1$ wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 differenzierbar mit $x'(t) = x_1$ im Sinne von Beispiel 10.3 (iii). Dann gibt es ein Intervall $(-\epsilon, \epsilon)$ im Urbild $x^{-1}[U]$ von U unter x , und wegen Satz 10.4 (iii) ist $(-\epsilon, \epsilon) \rightarrow Y$, $t \mapsto f(x(t))$ bei $t = 0$ differenzierbar mit der Ableitung $f'(x_0)(x_1)$ im Sinne von Beispiel 10.3 (iii). Also existiert die Richtungsableitung und es gilt

$$\frac{d}{dt} f(x_0 + tx_1)|_{t=0} = f'(x_0)(x_1).$$

$$(ii) \quad f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Für $t \neq 0$ ist dann $f(t \cos(\phi), t \sin(\phi)) = \sin(2\phi)$ und $f(t \cos(\phi), t \sin(\phi)) = 0$ für $t = 0$. Also ist f in $t = 0$ für $\phi \notin \frac{\pi}{2}\mathbb{Z}$ nicht stetig und auch nicht differenzierbar. Für $\phi \in \frac{\pi}{2}\mathbb{Z}$ existieren die Richtungsableitungen und verschwinden.

$$(iii) \quad f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{x^3 - y^3}{x^2 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Dann gilt $f(t(x, y)) = f(tx, ty) = tf(x, y)$ für $(x, y) \in \mathbb{R}^2$ und $x_0 = (0, 0)$. Also existieren alle Richtungsableitungen und setzen sich im Punkt $(0, 0)$ zu f zusammen. Wegen $2 = f(2, 0) \neq f(1, 1) + f(1, -1) = 0 + 1$ ist f im Punkt $(0, 0)$ nicht differenzierbar.

$$(iv) \quad f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2x^3y}{x^4 + y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

Die Richtungsableitungen von f in $(0, 0)$ in Richtung von (x, y) verschwinden alle:

$$\left. \frac{d}{dt} f(tx, ty) \right|_{t=0} = \begin{cases} \left. \frac{d}{dt} \frac{2t^2 x^3 y}{t^2 x^4 + y^2} \right|_{t=0} = 0 & \text{für } y \neq 0 \\ 0 & \text{für } y = 0, \end{cases} \quad |f(x, y)| = 2|x| \sqrt{\frac{x^4}{x^4 + y^2} \frac{y^2}{x^4 + y^2}} \leq 2|x|.$$

Wegen $\frac{d}{dt}(t, t^2) = (1, 2t)$ und $\frac{d}{dt} f(t, t^2) = \frac{d}{dt} t = 1 \neq 0$ ist f in $(0, 0)$ nicht differenzierbar.

Wir wollen den wichtigsten Fall von Funktionen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ genauer betrachten.

Definition 10.15. (Partielle Ableitungen in $\mathbb{R}^n$) Sei $f: U \rightarrow \mathbb{R}^m$ eine Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ in den $\mathbb{R}^m$. Dann sind die Komponenten $(f_1, \dots, f_m)$ von f offenbar reelle Funktionen auf U . Die Funktion f ist in $(x_1, \dots, x_n) \in U$ genau dann partiell differenzierbar, wenn für alle $i = 1, \dots, n$ und $j = 1, \dots, m$ die Funktionen

$$x \mapsto f_j(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

bei $x = x_i$ differenzierbar sind. Die entsprechenden Ableitungen heißen partielle Ableitungen von f und werden mit $\frac{\partial f_j}{\partial x_i}(x_1, \dots, x_n)$ bezeichnet. Wenn diese partiellen Ableitungen für alle $x \in U$ existieren, heißt f auf U partiell differenzierbar.

Definition 10.16. (Vektorfeld, Gradient, Divergenz und Rotation) Eine Abbildung von einer offenen Menge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^n$ wird Vektorfeld genannt. Das Vektorfeld

$$\text{grad } f = \nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

der partiellen Ableitungen einer partiell differenzierbaren reellen Funktion f auf einer offenen Menge $U \subset \mathbb{R}^n$ heißt Gradient von f . Wenn f ein partiell differenzierbares Vektorfeld ist, dann ist die Divergenz von f folgende reelle Funktion

$$\text{div } f = \nabla \cdot f = \frac{\partial f_1}{\partial x_1} + \dots + \frac{\partial f_n}{\partial x_n}.$$

Die lineare Abbildung $\Delta: f \mapsto \Delta f = \text{div grad } f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$

auf den zweimal partiell differenzierbaren reellen Funktionen heißt Laplaceoperator.

Im Fall von $n = 3$ ist die Rotation eines differenzierbaren Vektorfeldes f definiert durch

$$\text{rot } f = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right).$$

Wenn die reelle Funktion f in $x_0 \in U$ differenzierbar ist, dann ist f auch in x_0 partiell differenzierbar und die partiellen Ableitungen sind die Richtungsableitungen in Richtung der kanonischen Einheitsvektoren $e_1, \dots, e_n$ aus dem Beweis von Satz 9.37. Wegen der Linearität der Ableitung ist die Ableitung die lineare Abbildung:

$$f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto x_1 \frac{\partial f}{\partial x_1}(x_0) + \dots + x_n \frac{\partial f}{\partial x_n}(x_0).$$

Wenn wir den $\mathbb{R}^n$ mit den Spaltenvektoren bezeichnen, können wir diese Abbildung durch das Matrixprodukt des Zeilenvektors ∇f mit dem Spaltenvektor x darstellen:

$$f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto (\nabla f(x_0)) \cdot x.$$

Oder allgemeiner, für eine $\mathbb{R}^m$ -wertige Funktion können wir die Ableitung $f'(x_0)$ von f an der Stelle x_0 als lineare Abbildung mit der Jacobimatrix identifizieren:

$$\frac{\partial f(x_0)}{\partial x} = \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \dots & \frac{\partial f_1(x_0)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x_0)}{\partial x_1} & \dots & \frac{\partial f_m(x_0)}{\partial x_n} \end{pmatrix} \quad f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto \frac{\partial f(x_0)}{\partial x} \cdot x.$$

Die lineare Abbildung ist einfach die Matrixmultiplikation der Jacobimatrix, einer $m \times n$ -Matrix, mit dem Spaltenvektoren in $\mathbb{R}^n$. Insbesondere ist also die Richtungsableitung einer reellen Funktion f auf U an der Stelle $x_0 \in U$ in Richtung eines Vektors $x_1 \in \mathbb{R}^n$ das Skalarprodukt des Gradienten $\nabla f(x_0)$ von f an der Stelle x_0 mit dem Vektor x_1 :

$$\frac{d}{dt} f(x_0 + tx_1)|_{t=0} = x_1 \cdot \nabla f(x_0).$$

Satz 10.9 und Satz 10.11 zeigen insbesondere, dass folgendes gilt:

Korollar 10.17. *Sei $f : U \rightarrow \mathbb{R}^m$ eine Funktion auf einer offenen Teilmenge $U \subset \mathbb{R}^n$.*

- (i) *Wenn f in $x_0 \in U$ differenzierbar ist, dann existieren in x_0 alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(x_0)$ und setzen sich zu der Jacobimatrix zusammen.*
- (ii) *Wenn f auf U stetig differenzierbar ist, dann existieren auf U die partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ und setzen sich zusammen zu einer stetigen Funktion von U in die $m \times n$ -Matrizen in $\mathbb{R}^{m \times n}$. Diese Matrizen heißen Jacobimatrizen von f .*
- (iii) *Wenn f auf U partiell differenzierbar ist, und alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$ auf U stetig sind, dann ist f auf U stetig differenzierbar. Die Ableitung bei x_0 ist die Multiplikation der Jacobimatrix $\frac{\partial f(x_0)}{\partial x}$ mit Spaltenvektoren in $\mathbb{R}^n$. **q.e.d.***

Definition 10.18. *Eine Nullstelle $x_0 \in U$ der Ableitung f' einer auf einer offenen Menge U reellen differenzierbaren Funktion f heißt kritischer Punkt.*

Satz 10.19. *Jedes lokale Maximum (bzw. Minimum) einer differenzierbaren reellen Funktion auf einer offenen Menge ist ein kritischer Punkt.*

Beweis: Sei x_0 ein solches lokales Maximum (bzw. Minimum). Dann ist für alle $x_1 \in X$ die entsprechende Abbildung $t \mapsto f(x_0 + tx_1)$ auf einer Umgebung von $t = 0$ differenzierbar und besitzt dort ein lokales Maximum (bzw. Minimum). Also verschwindet die entsprechende Richtungsableitung. Dann verschwindet auch $f'(x_0)$ auf allen x_1 . **q.e.d.**

10.4 Höhere Ableitungen

Sei f eine auf einer offenen Teilmenge U eines Banachraumes X differenzierbare Funktion in den Banachraum Y . Wenn f zweimal differenzierbar ist, dann ist f' stetig. Die Ableitung f' ist dann eine stetige Abbildung von U nach $\mathcal{L}(X, Y)$. Die zweite Ableitung $f''(x_0)$ ist an den Stellen $x_0 \in U$, wo sie existiert, ein Element von $\mathcal{L}(X, \mathcal{L}(X, Y))$.

Definition 10.20. *Eine Abbildung $A : V \times V \rightarrow W$ heißt bilinear, wenn für alle $v, v', v'' \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$\begin{aligned} A(v + v'', v') &= A(v, v') + A(v'', v') & \text{und} & & A(v, v' + v'') &= A(v, v') + A(v, v'') \\ A(\lambda v, v') &= \lambda A(v, v') & \text{und} & & A(v, \lambda v') &= \lambda A(v, v'). \end{aligned}$$

Das kartesische Produkt $V \times V$ von (normierten) Vektorräumen ist wieder ein normierter Vektorraum. Die bilinearen Abbildungen von $V \times V$ nach W unterscheiden sich von den linearen Abbildungen von $V \times V$ nach W . Es gibt einen anderen Vektorraum $V \otimes V$, den man das Tensorprodukt von V mit V nennt, so dass die linearen Abbildungen von $V \otimes V$ nach W genau die bilinearen Abbildungen von $V \times V$ nach W sind. Allerdings besitzt $V \otimes V$ keine natürliche Norm. Für die Dimensionen gilt

$$\dim(V \times V) = \dim(V) + \dim(V) \qquad \dim(V \otimes V) = \dim(V) \cdot \dim(V).$$

Die bilinearen Abbildungen von $V \times V$ nach W lassen sich mit den linearen Abbildungen von V in die linearen Abbildungen von V nach W identifizieren:

Lemma 10.21. *Eine Abbildung $A : V \times V \rightarrow W$ ist genau dann bilinear, wenn*

$$B : V \rightarrow \{\text{Abbildungen } V \rightarrow W\}, \qquad v \mapsto B(v), \qquad B(v) : V \rightarrow W,$$

*eine lineare Abbildung von V in die linearen Abbildungen von V nach W ist, wobei B aus A als $B(v)(v') = A(v, v')$ und A aus B als $A(v, v') = B(v)(v')$ definiert ist. **q.e.d.***

Satz 10.22. (Satz von Schwarz) *Sei f eine differenzierbare Abbildung von einer offenen Teilmenge $U \subset X$ eines normierten Vektorraumes X in den normierten Vektorraum Y . Wenn f im Punkt x_0 zweimal differenzierbar ist, dann ist die der zweiten Ableitung entsprechende bilineare Abbildung $f''(x_0) : X \times X \rightarrow Y$ symmetrisch, d.h.*

$$(f''(x_0)x)y = (f''(x_0)y)x \quad \text{für alle } x, y \in X.$$

Beweis: Für $t \in [0, 1]$ und kleine $x, y \in X$ sei $g(t) = f(x_0 + tx + y) - f(x_0 + tx)$. Dann ist g differenzierbar mit der Ableitung mit Werten in $Y \simeq \mathcal{L}(\mathbb{R}, Y)$:

$$\begin{aligned} g'(t) &= f'(x_0 + tx + y)x - f'(x_0 + tx)x \\ &= ((f'(x_0 + tx + y) - f'(x_0)) - (f'(x_0 + tx) - f'(x_0)))x \end{aligned}$$

Weil f in x_0 zweimal differenzierbar ist, gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$, so dass $B(x_0, 2\delta) \subset U$ und außerdem für $x, y \in B(0, \delta) \subset X$ und $t \in [0, 1]$ die Ungleichungen

$$\begin{aligned} \|f'(x_0 + tx + y) - f'(x_0) - f''(x_0)tx - f''(x_0)y\| &\leq \epsilon \|tx + y\| \leq \epsilon(t\|x\| + \|y\|) \\ \|f'(x_0 + tx) - f'(x_0) - f''(x_0)tx\| &\leq \epsilon \|tx\| = \epsilon t\|x\| \end{aligned}$$

gelten. Daraus folgt für $t \in [0, 1]$ $\|g'(t) - (f''(x_0)y)x\| \leq \epsilon\|x\|(2\|x\| + \|y\|)$.

Die Anwendung von Korollar 10.6 auf die Funktion $t \mapsto g(t) - t(f''(x_0)y)x$ ergibt dann

$$\|g(1) - g(0) - (f''(x_0)y)x\| \leq \sup\{\|g'(t) - (f''(x_0)y)x\| \mid t \in [0, 1]\} \leq \epsilon\|x\|(2\|x\| + \|y\|).$$

Weil $g(1) - g(0) = f(x_0 + x + y) - f(x_0 + x) - f(x_0 + y) + f(x_0)$ in x und y symmetrisch ist gilt dann auch $\|g(1) - g(0) - (f''(x_0)x)y\| \leq \epsilon\|y\|(2\|y\| + \|x\|)$. Daraus folgt

$$\|(f''(x_0)y)x - (f''(x_0)x)y\| \leq 2\epsilon(\|x\|^2 + \|x\| \cdot \|y\| + \|y\|^2) \quad \text{für alle } x, y \in B(0, \delta).$$

Diese Ungleichung gilt wegen der Linearität nicht nur für $x, y \in B(0, \delta)$, sondern für $x, y \in X$. Im Grenzwert $\epsilon \rightarrow 0$ folgt $(f''(x_0)y)x = (f''(x_0)x)y$ für alle $x, y \in X$. **q.e.d.**

Zusammen mit Satz 10.11 erhalten wir

Korollar 10.23. Sei $f : U \rightarrow \mathbb{R}^m$ eine zweimal stetig partiell differenzierbare Funktion von einer offenen Teilmenge $U \subset \mathbb{R}^n$ nach $\mathbb{R}^m$. Dann vertauschen die partiellen Ableitungen, d.h. für alle $i, j = 1, \dots, n$ und $k = 1, \dots, m$ gilt

$$\partial_i \partial_j f_k = \frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i} = \partial_j \partial_i f_k \quad \text{und} \quad \text{rot grad } f = 0 \quad \text{für } n = 3, m = 1. \quad \mathbf{q.e.d.}$$

Durch mehrfaches Anwenden und differenzieren erhalten wir dann auch

Korollar 10.24. Sei f eine Abbildung von einer offenen Teilmenge U eines normierten Vektorraumes X in den normierten Vektorraum Y , die in $x_0 \in U$ n -mal differenzierbar ist. Dann ist $f^{(n)}(x_0)$ eine multilineare symmetrische Abbildung von $X \times X \times \dots \times X$ nach Y . D.h. für jede Permutation $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ und $x_1, \dots, x_n \in X$ gilt

$$(\dots ((f^{(n)}(x_0)x_1)x_2)\dots)x_n = (\dots ((f^{(n)}(x_0)x_{\sigma(1)})x_{\sigma(2)})\dots)x_{\sigma(n)}. \quad \mathbf{q.e.d.}$$

Beispiel 10.25. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ mit $f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2}$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$. Dann ist f zweimal partiell differenzierbar.

$$\frac{\partial f(x, y)}{\partial x} = y \frac{(x^2 - y^2)(x^2 + y^2) + 2x^2(x^2 + y^2) - 2x^2(x^2 - y^2)}{(x^2 + y^2)^2} = y \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}$$

$$\frac{\partial f(x, y)}{\partial y} = x \frac{(x^2 - y^2)(x^2 + y^2) - 2y^2(x^2 + y^2) - 2y^2(x^2 - y^2)}{(x^2 + y^2)^2} = x \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

$$\frac{\partial^2 f(x, 0)}{\partial x^2} = 0, \quad \frac{\partial^2 f(0, y)}{\partial y \partial x} = -1, \quad \frac{\partial^2 f(x, 0)}{\partial x \partial y} = 1, \quad \frac{\partial^2 f(0, y)}{\partial y^2} = 0.$$

Also existieren auf $\mathbb{R}^2$ alle zweiten partiellen Ableitungen, mit $\frac{\partial^2 f}{\partial y \partial x}(0, 0) \neq \frac{\partial^2 f}{\partial x \partial y}(0, 0)$.

Definition 10.26. Sei $f : U \rightarrow \mathbb{R}$ eine reelle Funktion auf einer offenen Teilmenge U des normierten Vektorraumes X , die bei $x_0 \in U$ zweimal differenzierbar ist. Dann definiert die zweite Ableitung eine symmetrische Bilinearform auf X :

$$f''(x_0) : X \times X \rightarrow \mathbb{R}, \quad (x, y) \mapsto f''(x_0)(x, y).$$

Für $X = \mathbb{R}^n$ identifizieren wir die Elemente von X wieder mit den Spaltenvektoren. Dann ist diese Bilinearform durch die sogenannte Hessematrix gegeben:

$$f''(x_0)(x, y) = y^t \cdot \frac{\partial^2 f(x_0)}{\partial x^2} \cdot x = \sum_{i,j=1}^n x_j \frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} y_i.$$

Satz 10.27. Sei $f : U \rightarrow \mathbb{R}$ eine auf einer offenen Menge zweimal differenzierbare reelle Funktion f . Dann ist die zweite Ableitung bei allen lokalen Minima (Maxima) eine nicht negative (nicht positive) Bilinearform: $f''(x_0)(x, x) \geq 0$ bzw. ≤ 0 für alle $x \in X$. Gibt es umgekehrt einen kritischen Punkt $x_0 \in U$ und ein $\epsilon > 0$ mit

$$f''(x_0)(x, x) \geq \epsilon \|x\|^2 \quad \text{bzw.} \quad f''(x_0)(x, x) \leq -\epsilon \|x\|^2 \quad \text{für alle } x \in X,$$

dann ist der kritische Punkt ein striktes lokales Minimum bzw. Maximum.

Beweis: Wenn x_0 ein lokales Maximum bzw. Minimum von f ist, dann für alle $x \in X$ auch $t = 0$ von $t \mapsto f(x_0 + tx)$. Deshalb folgt die erste Aussage aus Korollar 7.17.

Umgekehrt folgt aus der zweimaligen Differenzierbarkeit, dass für ein $\delta > 0$

$$-\frac{\epsilon}{2} \|x\|^2 < f'(x_0 + x)x - f'(x_0)x - f''(x_0)(x, x) < \frac{\epsilon}{2} \|x\|^2$$

für alle $x \in B(0, \delta)$ gilt. Daraus und den obigen Bedingungen folgt für die gleichen x

$$f'(x_0 + x)x = f'(x_0 + x)x - f'(x_0)x - f''(x_0)(x, x) + f''(x_0)(x, x) > \frac{\epsilon}{2} \|x\|^2 \quad \text{bzw.} \quad < -\frac{\epsilon}{2} \|x\|^2.$$

Also ist $f(x_0 + x) - f(x_0) = \int_0^1 f'(x_0 + tx)x dt > \frac{\epsilon}{4}\|x\|^2$ bzw. $< -\frac{\epsilon}{4}\|x\|^2$. **q.e.d.**

Auf endlichdimensionalen Räumen zeigt Beweis von Satz 9.37 wie $f''(x_0)(x, x) \geq \epsilon\|x\|^2$ aus $f''(x_0)(x, x) > 0$ für alle $x \in X \setminus \{0\}$ folgt. In unendlichdimensionalen Räumen gilt das nicht, und diese Bedingung ist nicht hinreichend für ein lokales Minimum.

Beispiel 10.28. Sei $f : C([0, 1]) \rightarrow \mathbb{R}$, $x \mapsto \int_0^1 x^2(t)(t-x(t))dt$. Dann ist $f''(0)(x, x) = 2 \int_0^1 x^2(t)tdt > 0$ für alle $x \in C([0, 1]) \setminus \{0\}$. Sei $x_\epsilon(t) = \begin{cases} \epsilon - t & \text{für } 0 \leq t < \epsilon \\ 0 & \text{für } \epsilon \leq t \leq 1 \end{cases}$ mit $\epsilon \in (0, 1)$. Dann gilt $\|x_\epsilon\|_\infty = \epsilon$ und $f(sx_\epsilon) = s^2 \int_0^\epsilon (\epsilon - t)^2 dt - s^3 \int_0^\epsilon (\epsilon - t)^3 dt = -\frac{s^2}{3}(\epsilon - t)^3 \Big|_0^\epsilon - (\frac{s^2}{12} - \frac{s^3}{4})(\epsilon - t)^4 \Big|_0^\epsilon = (\frac{s^2}{12} - \frac{s^3}{4})\epsilon^4$. Also ist $x = 0$ kein lokales Minimum.

Zum Abschluss wollen wir das Taylorpolynom und die Taylorreihe einer Funktion $f : U \rightarrow Y$ auf einer offenen konvexen Teilmenge $U \subset X$ eines normierten Vektorraumes X in einem normierten Vektorraum Y betrachten. Für $x_0, x \in U$ sei

$$g : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(x_0 + t(x - x_0))$$

Wenn f auf U n -mal differenzierbar ist, dann ist auch g n -mal differenzierbar. Wegen der Kettenregel Satz 10.4 (iii) ist die m -te Ableitung von g gleich

$$g^{(m)}(t) = (\dots (f^{(m)}(x_0 + t(x - x_0))(x - x_0)) \dots (x - x_0)),$$

also die m -lineare symmetrische Form zu $f^{(m)}(x_0 + t(x - x_0))$ ausgewertet auf $((x - x_0), \dots, (x - x_0)) \in X^{\times m}$. Dann erhalten wir für n -mal differenzierbare bzw. glatte Funktionen das Taylorpolynom der Ordnung n von f bei x_0 bzw. die Taylorreihe:

$$x \mapsto \sum_{k=0}^n \frac{f^{(k)}(x_0)((x - x_0), \dots, (x - x_0))}{k!} \quad \text{bzw.} \quad x \mapsto \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)((x - x_0), \dots, (x - x_0))}{k!}.$$

Eine unendlich oft differenzierbare Funktion f heißt wieder reell analytisch in x_0 , wenn die entsprechende Taylorreihe auf einer Umgebung von x_0 gegen $f(x)$ konvergiert. Auf einem Banachraum X ist z.B. $\exp : \mathcal{L}(X) \rightarrow \mathcal{L}(X)$ eine analytische Funktion.

Für reelle Funktionen, also $Y = \mathbb{R}$ ergibt der Satz von Taylor:

Satz 10.29. (von Taylor in höheren Dimensionen) Sei $f : U \rightarrow \mathbb{R}$ eine auf einer offenen konvexen Teilmenge eines normierten Vektorraumes definierte $(n+1)$ -mal differenzierbare Funktion. Dann gibt es für jedes $x, x_0 \in U$ ein $\xi \in (0, 1)$, so dass

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

gilt. Hierbei bezeichnen wir mit $f^{(k)}(x_0)$ bzw. $f^{(n+1)}(x_0 + \xi(x - x_0))$ die entsprechende multilineare Abbildung von $X^{\times k}$ bzw. $X^{\times (n+1)}$ nach $\mathbb{R}$. Der erste Term heißt wieder Taylorpolynom von f in x_0 der Ordnung n und der zweite Term Restglied **q.e.d.**