Kapitel 10

Differentialrechnung von Funktionen mehrerer Veränderlicher

10.1 Ableitungen von $f : X \to Y$

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:*

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

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

$$\frac{\|(A-B)(x-x_0)\|}{\|x-x_0\|} \leq \frac{\|f(x)-f(x_0)-A(x-x_0)\|}{\|x-x_0\|} + \frac{\|f(x)-f(x_0)-B(x-x_0)\|}{\|x-x_0\|} \leq 2\epsilon \text{ für alle } x \in B(x_0, \delta) \setminus \{0\}.$$

Also ist $A = B$ und damit die Ableitung, wenn sie existiert, eindeutig.

Satz 10.2. *Sei $f : U \subset X \to Y$ 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* $\dfrac{\|f(x)-f(x_0)-f(x-x_0)\|}{\|x-x_0\|} = 0.$
Also ist f in x_0 differenzierbar und $f'(x_0) = f$.
(iii) *Die Abbildung $Y \to \mathcal{L}(\mathbb{R},Y)$, $\quad x \mapsto$ Multiplikation mit x besitzt offenbar die Umkehrabbildung $\mathcal{L}(\mathbb{R},Y) \to Y$, $\quad A \mapsto A(1)$ und ist ein isometrischer Isomorphismus von normierten Vektorräumen. Deshalb können wir die Ableitungen von differenzierbaren Funktionen $f : (a,b) \to Y$ auf offenen Intervallen $(a,b) \subset \mathbb{R}$ mit Funktionen auf diesen Intervallen identifizieren, die wir auch mit $f' : (a,b) \to Y$ bezeichnen. Für $x \neq x_0$ gilt*

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

Deshalb ist f als Funktion von der Teilmenge (a,b) des normierten Vektorrraumes $\mathbb{R}$ in den normierten Vektorraum $\mathbb{R}$ genau dann in x_0 differenzierbar, wenn f als reelle Funktion in x_0 im Sinne von Definition 7.1 differenzierbar ist.

Satz 10.4. **(i)** *Seien $f, g : U \subset X \to Y$ 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 $f, g : U \subset X \to Y$ in x_0 differenzierbar und Y eine normierte Algebra. Dann ist $f \cdot g$ in x_0 differenzierbar und es gilt (Leibnizregel)*

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

(iii) *Seien $f : U \subset X \to Y$ in $x_0 \in U$ und $g : V \subset Y \to Z$ 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) Wegen der Dreiecksungleichung gilt für $x \neq x_0$

$$\frac{\|f(x)+g(x)-f(x_0)-g(x_0)-f'(x_0)(x-x_0)-g'(x_0)(x-x_0)\|}{\|x-x_0\|}$$
$$\leq \frac{\|f(x)-f(x_0)-f'(x_0)(x-x_0)\|}{\|x-x_0\|} + \frac{\|g(x)-g(x_0)-g'(x_0)(x-x_0)\|}{\|x-x_0\|}.$$

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

$$\frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x-x_0)g(x_0) - f(x_0)g'(x_0)(x-x_0)\|}{\|x-x_0\|}$$
$$\leq \frac{\|f(x)-f(x_0)-f'(x_0)(x-x_0)\|}{\|x-x_0\|}\|g(x)\| + \|f'(x_0)\| \cdot \|g(x)-g(x_0)\| +$$
$$+ \|f(x_0)\| \frac{\|g(x)-g(x_0)-g'(x_0)(x-x_0)\|}{\|x-x_0\|}.$$

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).

(iii) Aus Satz 9.61 folgt für $x \neq x_0$ $\dfrac{\|(g \circ f)(x) - (g \circ f)(x_0) - g'(f(x_0))f'(x_0)(x - x_0)\|}{\|x - x_0\|}$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} + \frac{\|g'(f(x_0))(f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|}$$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - g'(f(x_0))(f(x) - f(x_0))\|}{\|f(x) - f(x_0)\|} \cdot \left(\|f'(x_0)\| + \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}\right) + \|g'(f(x_0))\| \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0))\|}{\|x - x_0\|}.$$

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 $\|f'(y)\| \le \phi'(y)$ 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 $x \in (y, y + \delta)$

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

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 $\quad \|f(x) - f(a)\| \le \|f(x) - f(y)\| + \|f(y) - f(a)\|$

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

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 $\quad \|f(b) - f(a)\| \le \phi(b) - \phi(a)$. **q.e.d.**

Korollar 10.6. *(Schrankensatz) 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*

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

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

Beweis: Die Abbildung $x : \mathbb{R} \to X$, $t \mapsto x(t) = (t - 1)a + tb = a + t(b - a)$ ist wegen Beispiel 10.3 (i)-(ii) und Satz 10.4 (i) diffferenzierbar 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

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

Die zweite Behauptung folgt aus diesem Lemma mit den beiden Funktionen

$$t \mapsto f(x(t)) - tA(b - a) \in Y, \quad t \mapsto t\|b - a\| \sup\{\|f'(x) - A\| \mid x \in D \setminus S\} \in \mathbb{R}.$$ **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 $f' : U \to \mathcal{L}(X, Y), \quad x \mapsto f'(x)$ 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 $U \subset X_1 \times \ldots \times X_n$ eines n-fachen kartesischen Produktes von normierten Vektorräumen in einen normierten Vektorraum Y im Punkt $(x_1, \ldots, x_n) \in U$ partiell differenzierbar, wenn für $i = 1, \ldots, n$ die Abbildungen $x \mapsto f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$ 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

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

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

$$\frac{\|f(x, x_2) - f(x_1, x_2) - A((x, x_2) - (x_1, x_2))\|}{\|(x, x_2) - (x_1, x_2)\|} = \frac{\|f(x, x_2) - f(x_1, x_2) - A_1(x - x_1)\|}{\|x - x_1\|}$$

$$\frac{\|f(x_1, x) - f(x_1, x_2) - A((x_1, x) - (x_1, x_2))\|}{\|(x_1, x) - (x_1, x_2)\|} = \frac{\|f(x_1, x) - f(x_1, x_2) - A_2(x - x_2)\|}{\|x - x_2\|}$$

gilt, folgt aus den Definitionen der folgende Satz:

Satz 10.9. *Eine im Punkt $(x_1, x_2) \in U$ differenzierbar 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:*

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

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

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 $\lim_{r\to 0+} f(r\cos\phi, r\sin\phi) = \sin(2\phi)$. 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 $f : U \to Y$ bei $x = (x_1, x_2) \in U$ partiell differenzierbar. Wenn eine der partiellen Ableitungen*

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

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

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

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

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

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 $A_1 : X_1 \to Y$, $x_1 \mapsto A((x_1, 0))$ und $A_2 : X_2 \to Y$, $x_2 \mapsto A((0, x_2))$ 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

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

Also ist die Abbildung $\quad \mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y) \to \mathcal{L}(X_1 \times X_2, Y), \quad (A_1, A_2) \mapsto A$ 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 $f : U \to Y$, die beiden partiellen Ableitungen $\frac{\partial f}{\partial x_1} : U \to \mathcal{L}(X_1, Y)$ und $\frac{\partial f}{\partial x_2} : U \to \mathcal{L}(X_2, Y)$ stetig sind.

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

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

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

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

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

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 $f : U \to Y$ 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*

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

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 $f : U \to Y$ in $x_0 \in U$ differenzierbar. Für $x_1 \in X$ ist die Abbildung $x : \mathbb{R} \to X$, $t \mapsto x(t) = x_0 + tx_1$ 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 $(-\epsilon, \epsilon) \to Y$, $t \mapsto f(x(t))$ 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)
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq 0 \\ 0 & \text{für } (x, y) = 0. \end{cases}$$

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

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. Weil diese Abbildung nicht linear ist, ist* f *im Punkt* $(0,0)$ *nicht differenzierbar.*

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

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

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

Wegen $\lim_{x\to 0}\frac{|f(x,x^2)|}{\|(x,x^2)\|}=\lim_{x\to 0}\frac{1}{\sqrt{1+x^2}}=1\neq 0$ *ist* f *in* $(0,0)$ *nicht differenzierbar.*

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

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

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

bei $x=x_i$ *differenzierbar sind. Die entsprechenden Ableitungen heißen partielle Ableitungen von* f *und werden mit* $\frac{\partial f_j}{\partial x_i}(x_1,\ldots,x_n)$ *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*

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

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 Vektofeld ist, dann ist die Divergenz von* f *folgende reelle Funktion*

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

Die lineare Abbildung $\triangle:\quad f\mapsto\triangle f=\operatorname{div}\operatorname{grad} f=\frac{\partial^2 f}{\partial x_1^2}+\ldots+\frac{\partial^2 f}{\partial x_n^2}$

auf den zweimal partiell differenzierbaren reellen Funktionen heißt Laplaceoperator. Im Fall von $n=3$ *ist die Rotation eines differenzierbaren Vektorfeldes* f *definiert durch*

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

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 $e_1, \ldots, e_n$ aus dem Beweis von Satz 9.37. Wegen der Linearität der Ableitung ist die Ableitung die lineare Abbildung:

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

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:

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

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:

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

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 $f : U \to \mathbb{R}^m$ 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 einesBanachraumes 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 Abbleitung $f''(x_0)$ 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 $A : V \times V \to W$ heißt bilinear, wenn für alle $v, v', v'' \in V$ und $\lambda \in \mathbb{K}$ gilt*

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

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 lineare 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 Abbilungen 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 $A : V \times V \to W$ ist genau dann bilinear, wenn*

$$B : V \to \{\textit{Abbildungen } V \to W\}, \quad v \mapsto B(v), \quad B(v) : V \to W, \quad B(v)(v') = A(v, v')$$

eine lineare Abbildung von V in die linearen Abbildungen von V nach W 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 $f''(x_0) : X \times X \to Y$ symmetrisch, d.h.*

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

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

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

gelten. Daraus folgt für $t \in [0,1]$ $\qquad \|g'(t) - (f''(x_0)y)x\| \le \epsilon\|x\|(2\|x\| + \|y\|)$.
Die Anwendung von Korollar 10.6 auf die Funktion $t \mapsto g(t) - t(f''(x_0)y)x$ ergibt dann

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

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 $\|g(1) - g(0) - (f''(x_0)x)y\| \le \epsilon\|y\|(2\|y\| + \|x\|)$. Daraus folgt

$$\|(f''(x_0)y)x - (f''(x_0)x)y\| \le 2\epsilon(\|x\|^2 + \|x\| \cdot \|y\| + \|y\|^2).$$

Diese Ungleichung gilt wegen der Linearität nichtnur für $x, y \in B(0, \delta)$, sondern für $x, y \in X$. Im Grenzwert $\epsilon \to 0$ 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 $f : U \to \mathbb{R}^m$ 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 $i, j = 1, \ldots, n$ und $k = 1, \ldots, m$ gilt*

$$\partial_i \partial_j f_k = \frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i} = \partial_j \partial_i f_k \text{ und } \operatorname{rot} \operatorname{grad} f = 0 \text{ für } n = 3, m = 1.$$ **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 $X \times X \times \ldots \times X$ nach Y. D.h. für jede Permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ und $x_1, \ldots, x_n \in X$ gilt*

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

Beispiel 10.25. *Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto f(x,y)$ 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.*

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

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

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

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 $f : U \to \mathbb{R}$ eine reelle Funktion auf einer offenenTeilmenge U des normierten Vektorraumes X, die bei $x_0 \in U$ zweimal differenzierbar ist. Dann definiert die zweite Ableitung eine symmetrische Bilinearform auf X:*

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

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 $f : U \to \mathbb{R}$ 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 $\delta > 0$

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

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

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

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

Auf endlichdimensionalen Räumen ist die Bedingung $f''(x_0)(x,x) \geq \epsilon\|x\|^2$ äquivalent zu $f''(x_0)(x,x) > 0$ für alle $x \in X \setminus \{0\}$. In unendlichdimensionalen Räumen nicht. Die zweite Bedingung ist dann auch nicht hinreichend für ein lokales Minimum.

Beispiel 10.28. *Sei* $f : C([0,1]) \to \mathbb{R},\ x \mapsto \int_0^1 x^2(t)(t-x(t))dt$*. 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* $f(sx_\epsilon) = s^2\int_0^\epsilon(\epsilon-t)^2tdt - s^3\int_0^\epsilon(\epsilon-t)^3dt = -\frac{s^2}{3}(\epsilon-t)^3t]_0^\epsilon - (\frac{s^2}{12}-\frac{s^3}{4})(\epsilon-t)^4]_0^\epsilon = (\frac{s^2}{12}-\frac{s^3}{4})\epsilon^4$*. Also ist* $x=0$ *kein lokales Minimum.*

Zum Abschluss wollen wir den Satz von Taylor auf reelle Funktionen f auf offenen konvexen Teilmengen $U \subset X$ eines normierten Vektorraumes verallgemeinern. Wenn $x_0, x \in U$ in einer solchen konvexen offenen Teilmenge U liegen, dann ist

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

eine reelle Funktion. 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

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

also die m-lineare symmetrische Form zu $f^{(m)}(x_0+t(x-x_0))$ ausgewertet auf $((x-x_0),\ldots,(x-x_0)) \in X^{\times m}$. Dann erhalten wir nach dem Satz von Taylor:

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

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

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

Das Taylorpolynom und entsprechend die Taylorreihe ist auch für glatte Funktionen in einen normierten Vektorraum definiert. Eine unendlich oft differenzierbare Funktion heißt wieder reell analytisch in x_0, wenn die entsprechende Taylorreihe auf einer Umgebung von x_0 gegen die Funktion konvergiert. So definiert auf einem Banachraum die Exponentialfunktion eine analytische Funktion von $\mathcal{L}(X)$ auf sich selber.