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 $x_0 \in X$ 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 A' beide diese Bedingung erfüllen, dann gibt es für alle $\epsilon > 0$ ein $\delta > 0$ mit

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

Also ist $A = A'$ 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: Weil $f(x) - f(x_0) = A(x - x_0) + (f(x) - f(x_0) - A(x - x_0))$ 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 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 $\mathbb{R} \to \mathcal{L}(\mathbb{R},\mathbb{R})$, $\quad x \mapsto$ Multiplikation mit x besitzt offenbar die Umkehrabbildung $\mathcal{L}(\mathbb{R},\mathbb{R}) \to \mathbb{R}$, $\quad A \mapsto A(1)$ und ist ein isometrischer Isomorphismus von normierten Vektorräumen. Deshalb können wir die Ableitungen von differenzierbaren reellen Funktionen f auf offenen Intervallen $(a,b) \subset \mathbb{R}$ mit reellen Funktionen auf diesen Intervallen identifizieren, die wir auch mit f' 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)** *Sei $f, g : U \subset X \to Y$ 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)*
$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0) \quad \text{mit}$$
$$(f'(x_0) \cdot g(x_0))x = (f'(x_0)x) \cdot g(x_0) \quad \text{und} \quad (f(x_0) \cdot g'(x_0))x = f(x_0) \cdot (g'(x_0)x).$$

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

(ii) Weil in einer normierten Algebra $\|AB\| \leq \|A\| \cdot \|B\|$ gilt, folgt
$$\begin{aligned} &\frac{\|f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - f'(x_0)(x - x_0)g(x_0) - f(x_0)g'(x_0)(x - x_0)\|}{\|x - x_0\|} \\ &\quad\leq \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|}\|g(x)\| + \|f'(x_0)\| \cdot \|g(x) - g(x_0)\| + \\ &\qquad + \|f(x_0)\| \frac{\|g(x) - g(x_0) - g'(x_0)(x - x_0)\|}{\|x - x_0\|}. \end{aligned}$$

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).

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

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

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

Daraus und aus Satz 10.2 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 einem 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 nicht gilt $\|f'\| \leq \phi'$. Sei $\epsilon > 0$ und $A_\epsilon \subset [a, b]$ die Menge

$$\left\{ y \,\middle|\, \text{für alle } x \in (a, y) \text{ gilt } \|f(x) - f(a)\| \leq \phi(x) - \phi(a) + \epsilon(x - a) + \epsilon \sum_{x_n < x} 2^{-n} \right\}.$$

Aus der Stetigkeit von f und ϕ folgt, dass auch für $y = \sup A_\epsilon$ gilt

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

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

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

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

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

Für $x \in (y, y + \delta)$ folgt $\phi(x) \geq \phi(y)$ aus $\phi'(y) \geq 0$ und damit auch

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

Woraus $y + \delta \in A_\epsilon$ folgt, im Widerspruch zu $y = \sup A_\epsilon$. Wenn es andererseits ein $N \in \mathbb{N}$ gibt mit $x_N = y$, 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 $\quad \|f(x) - f(a)\| \leq \|f(x) - f(y)\| + \|f(y) - f(a)\|$

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

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

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

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

Beweis: Sei für ein $t_0 \in (0,1)$ die Abbildung f in $x_t = (1-t)a + tb \in U$ differenzierbar. Dann ist die Abbildung $[0,1] \to Y, \quad t \mapsto f(x_t)$ im Punkt t_0 differenzierbar, und es gilt

$$t \to \begin{cases} \frac{\|f(x_t)-f(x_{t_0})-(t-t_0)f'(x_{t_0})(b-a)\|}{|t-t_0|} & \text{für } t \neq t_0 \\ 0 & \text{für } t = t_0 \end{cases}$$

ist stetig. Also ist die Ableitung von dieser Funktion in t_0 gleich $s \mapsto sf'(x_{t_0})(b-a) \in \mathcal{L}(\mathbb{R}, Y)$. Dann folgt die erste Behauptung aus Lemma 10.5 mit den beiden Funktionen

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

Die zweite Behauptung folgt aus diesem Lemma mit den 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 $x \mapsto f(x_1, \ldots, x_i, x, x_{i+1}, \ldots, x_n)$ im Punkt $x = x_i$ differenzierbar sind.*

Die wichtigsten Beispiele sind reelle Funktionen auf offenen Teilmengen des $\mathbb{R}^n$. Weil für jede stetige lineare Abbildung $A \in \mathcal{L}(X_1 \times X_2, Y)$ die Abbildungen

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

stetig und linear sind, und weil für $x \in X_1$ mit $(x, x_2) \in U$ bzw. $x \in X_2$ mit $(x_1, x) \in U$

$$\frac{\|f(x,x_2) - f(x_1,x_2) - A((x,x_2) - (x_1,x_2))\|}{\|(x,x_2) - (x_1,x_2)\|} = \frac{\|f(x,x_2) - f(x_1,x_2) - A(x - x_1, 0)\|}{\|x - x_1\|}$$
$$\frac{\|f(x_1,x) - f(x_1,x_2) - A((x_1,x) - (x_1,x_2))\|}{\|(x_1,x) - (x_1,x_2)\|} = \frac{\|f(x_1,x) - f(x_1,x_2) - A(0, x - x_2)\|}{\|x - x_1\|}$$

gilt, folgt aus den Definitionen der folgende

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:*

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

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

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. *Sei $f : U \to Y$ eine partiell differenzierbare Funktion von einer offenen Teilmenge U des kartesischen Produktes zwei normierter Vektorräume $X_1 \times X_2$ in dem normierten Vektorraum Y. Dann sind die partiellen Ableitungen von f*

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

genau dann auf U stetig, wenn f auf U stetig 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

$$X_1 \times X_2 \to Y, \quad (x_1, x_2) \to A_1x_1 \text{ bzw.} \quad X_1 \times X_2 \to Y, \quad (x_1, x_2) \to A_2x_2$$

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

$$A_1 \times A_2 : X_1 \times X_2 \to Y, \quad (x_1, x_2) \mapsto A_1x_1 + A_2x_2$$

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

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

Also ist die Abbildung $\mathcal{L}(X_1, Y) \times \mathcal{L}(X_2, Y) \to \mathcal{L}(X_1 \times X_2, Y), \quad (A_1, A_2) \mapsto A$ 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 $\frac{\partial f}{\partial x_1}$ stetig ist, dann folgt

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

Wegen der Stetigkeit von $\frac{\partial f}{\partial x_1}$ gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass für $z_1 \in B(x_1, \delta)$ und $z_2 \in B(x_2, \delta)$ gilt

$$\left\| \frac{\partial f}{\partial x_1}(z_1, z_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right\| < \frac{\epsilon}{2}.$$

Aus dem Schrankensatz Korollar 10.6 folgt dann für $y_1 \in B(x_1, \delta)$ und $y_2 \in B(x_1, \delta)$

$$\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f(x_1, x_2)}{\partial x_1}(y_1 - x_1) \right\| < \frac{\epsilon}{2} \|y_1 - x_1\|.$$

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

$$\left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f(x_1, x_2)}{\partial x_2}(y_2 - x_2) \right\| < \frac{\epsilon}{2} \|y_2 - x_2\|.$$

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

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

wobei $\frac{df(x_1,x_2)}{dx} = (\frac{\partial f(x_1,x_2)}{\partial x_1}, \frac{\partial f(x_1,x_2)}{\partial x_2})$ 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)**
$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{2xy}{x^2+y^2} & \textit{für } (x,y) \neq 0 \\ 0 & \textit{für } (x,y) = 0. \end{cases}$$

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

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

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.

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

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

Deshalb 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) Sei $f : U \to Y$ eine Funktion von einer offenen Teilmenge U eines normierten Vektorraumes X in einem normierten Vektorraum Y. Für $x_0 \in U$ und $x \in X \setminus \{0\}$ heißt die Ableitung in $t_0 = 0$ der Funktion*

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

wenn sie existiert, die Richtungsableitung von f in x_0 in Richtung x.

Beispiel 10.14. **(i)** *Sei $f : U \to Y$ in x_0 differenzierbar. Für $x \in X \setminus \{0\}$ gibt es ein Intervall $(-\epsilon, \epsilon)$ im Urbild von U unter $t \mapsto x_0 + tx$. Bei $t = 0$ sind die Abbildungen*

$$t \mapsto \begin{cases} \frac{\|f(x_0+tx)-f(x_0)-tf'(x_0)x\|}{|t|\|x\|} & \text{für } t \neq 0 \\ 0 & \text{für } t = 0 \end{cases} \quad t \mapsto \begin{cases} \frac{\|f(x_0+tx)-f(x_0)-tf'(x_0)x\|}{|t|} & \text{für } t \neq 0 \\ 0 & \text{für } t = 0 \end{cases}$$

stetig. Also existiert die Richtungsableitung und es gilt $\frac{df(x_0 + tx)}{dt}\Big|_{t=0} = f'(x_0)(x).$

(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 $f(tx) = t(\cos^3\phi - \sin^3\phi)$ für $x = (\cos\phi, \sin\phi)$ und $x_0 = 0$. 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.

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

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

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:

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

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

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

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

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

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 $\frac{\partial f_i(x_0)}{\partial x_j}$ 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 $x \in X$ die entsprechende Abbildung $t \mapsto f(x_0 + tx)$ auf einer Umgebung von $0 \in \mathbb{R}$ differenzierbar und besitzt dort ein lokales Maximum (bzw. Minimum). Also verschwindet die entsprechende Richtungsableitung. Dann verschwindet auch $\frac{df}{dx}(x_0)$ auf allen x. **q.e.d.**

10.4 Höhere Ableitungen

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

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*

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

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 $U \subset V$ eines normierten Vektorraumes V in den normierten Vektorraum W. Wenn f im Punkt x_0 zweimal differenzierbar ist, dann ist die der zweiten Ableitung entsprechende bilineare Abbildung $f''(x_0) : V \times V \to W$ symmetrisch, d.h.*

$$(f''(x_0)u)v = (f''(x_0)v)u \quad \text{für alle } u, v \in V.$$

Beweis: Für $t \in [0, 1]$ und kleine $u, v \in V$ sei $g(t) = f(x_0 + tu + v) - f(x_0 + tu)$. Dann ist g differenzierbar und es gilt

$$\begin{aligned} g'(t) &= f'(x_0 + tu + v)u - f'(x_0 + tu)u \\ &= ((f'(x_0 + tu + v) - f'(x_0)) - (f'(x_0 + tu) - f'(x_0)))u \end{aligned}$$

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 $u, v \in B(0, \delta) \subset V$ und $t \in [0, 1]$ die Ungleichungen

$$\begin{aligned} &\|f'(x_0 + tu + v) - f'(x_0) - f''(x_0)(tu + v)\| \le \epsilon(\|u\| + \|v\|) \\ &\|f'(x_0 + tu) - f'(x_0) - f''(x_0)tu\| \le \epsilon\|u\| \end{aligned}$$

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

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

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

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

Diese Ungleichung gilt wegen der Linearität nicht nur für $u, v \in B(0, \delta)$, sondern für $u, v \in V$. Im Grenzwert $\epsilon \to 0$ folgt $(f''(x_0)v)u = (f''(x_0)u)v$ für alle $u, v \in V$. **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*

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

Durch mehrfaches Anwenden und differenzieren erhalten wir dann auch

Korollar 10.24. *Sei f eine Abbildung von einer offenen Teilmenge U eines normierten Vektorraumes V in den normierten Vektorraum W, die in $x_0 \in U$ n mal differenzierbar ist. Dann ist $f^{(n)}(x_0)$ eine multilineare symmetrische Abbildung von $V \times V \times \ldots \times V$ nach W. D.h. für jede Permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ und $v_1, \ldots, v_n \in V$ gilt*

$$(\ldots((f^{(n)}(x_0)v_1)v_2)\ldots)v_n = (\ldots((f^{(n)}(x_0)v_{\sigma(1)})v_{\sigma(2)})\ldots)v_{\sigma(n)}. \qquad \textbf{q.e.d.}$$

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

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

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

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

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

Definition 10.26. *Sei $f : U \to W$ eine reelle Funktion auf einer offenen Teilmenge U des normierten Vektorraumes V, die bei $x_0 \in U$ zweimal differenzierbar ist. Dann definiert die zweite Ableitung eine symmetrische Bilinearform auf V:*

$$f''(x_0) : V \times V \to \mathbb{R}, \qquad\qquad (v,w) \mapsto f''(x_0)(v,w).$$

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

$$f''(x_0)(v,w) = w^t \cdot \frac{d^2 f}{dx^2}(x_0) \cdot v \qquad = \sum_{i,j=1}^{n} w_j \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0) v_i.$$

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 $x \in V$. Gibt es umgekehrt einen kritischen Punkt $x_0 \in U$ und ein $\delta > 0$ mit*

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

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

Beweis: Wenn x_0 ein lokales Maximum bzw. Minimum von f ist, dann für alle $x \in V$ auch $t = 0$ von $t \mapsto f(x_0 + tx)$. Deshalb folgt die erste Aussage aus Korollar 7.16.

Umgekehrt folgt aus den obigen Ungleichungen, dass es ein $\epsilon > 0$ gibt, so dass

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

gilt. Dort ist $f(x_0+x) - f(x_0) = \int_0^1 f'(x_0+tx)(x)dt \geq \frac{\delta}{4}\|x\|^2$ bzw. $\leq -\frac{\delta}{4}\|x\|^2$. **q.e.d.**

Auf endlichdimensionalen Räumen ist die Bedingung $f''(x_0)(x,x) \geq \delta\|x\|^2$ äquivalent zu $f''(x_0)(x,x) > 0$ für alle $x \in V \setminus \{0\}$. 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* $x_\epsilon(t) = \begin{cases} \epsilon - t & \text{für } 0 \le t < \epsilon \\ 0 & \text{für } \epsilon \le t \le 1 \end{cases}$ *mit* $\epsilon \in (0,1)$*. Dann gilt* $\|x_\epsilon\|_\infty = \epsilon$ *und* $f(sx_\epsilon) = s^2\int_0^\epsilon(\epsilon - t)^2 tdt - s^3\int_0^\epsilon(\epsilon - t)^3 dt = -\frac{s^2}{3}(\epsilon - t)^3 t]_0^\epsilon - (\frac{s^2}{12} - \frac{s^3}{4})(\epsilon - t)^4]_0^\epsilon = (\frac{s^2}{12} - \frac{s^3}{4})\epsilon^4$*. Also ist* $x = 0$ *kein lokales Minimum.*

Zum Abschluss wollen wir den Satz von Taylor auf reelle Funktionen f auf offenen konvexen Teilmengen $U \subset V$ 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 n-te Ableitung von g gleich

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

also die n-lineare symmetrische Form zu $f^{(n)}(x_0)$ ausgewertet auf $((x - x_0), \ldots, (x - x_0)) \in V^{xm}$. 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* $V^{\times k}$ *bzw.* $V^{\times(m+1)}$ *nach* $\mathbb{R}$*. Dabei heißt der erste Term wieder Taylorpolynom 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 $\mathcal{L}(V)$ auf sich selber.