

X. Differentiation vektorwertiger Funktionen

X.1. Partielle und stetige Differenzierbarkeit

Definition X.1. Seien $M \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ offen, $x \in U$, $F : U \rightarrow \mathbb{R}$ und $j \in \mathbb{Z}_1^M$.

(i) Die Funktion F heißt **bei $x \in U$ partiell nach x_j differenzierbar**

$$:\Leftrightarrow \frac{\partial F(x)}{\partial x_j} := \lim_{t \rightarrow 0} \left\{ \frac{F(x + te_j) - F(x)}{t} \right\} \text{ existiert,} \quad (\text{X.1})$$

wobei $e_j = (0, \dots, 0, 1, 0, \dots, 0)^T \in \mathbb{R}^M$ der j . kanonische Basisvektor ist. In diesem Fall heißt $\frac{\partial F(x)}{\partial x_j} =: \partial_{x_j} F(x)$ **partielle Ableitung von F nach x_j bei x .**

(ii) Die Funktion F heißt **bei $x \in U$ partiell differenzierbar**

$$:\Leftrightarrow \forall j \in \mathbb{Z}_1^M : F \text{ ist bei } x \text{ partiell nach } x_j \text{ differenzierbar.} \quad (\text{X.2})$$

In diesem Fall heißt der reelle Vektor

$$\nabla F(x) := \begin{pmatrix} \frac{\partial F(x)}{\partial x_1} \\ \frac{\partial F(x)}{\partial x_2} \\ \vdots \\ \frac{\partial F(x)}{\partial x_M} \end{pmatrix} \in \mathbb{R}^M \quad (\text{X.3})$$

der partiellen Ableitungen **Gradient von F bei x .**

(iii) Die Funktion F heißt **auf U partiell differenzierbar**

$$:\Leftrightarrow \forall x \in U : F \text{ ist bei } x \text{ partiell differenzierbar.} \quad (\text{X.4})$$

Bemerkungen und Beispiele.

- Da U eine offene Menge ist, ist $x \in U$ ein innerer Punkt, und es gibt $r > 0$, sodass $B_{\text{eukl}}(x, r) \subseteq U$. Diese Eigenschaft sichert, dass $F(x + te_j)$ und damit der Differenzenquotient in (X.1) für $t \in (-r, r)$ überhaupt definiert ist. Dies unterstreicht die Bedeutung des topologischen Begriffs der *offenen Menge*.
- Die partielle Ableitung einer Funktion F nach einer Variablen x_j bedeutet konkret, dass man $F(x_1, x_2, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_M)$ nach x_j differenziert und alle anderen Variablen, $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_M$, als Parameter betrachtet.
- Wir betrachten $U = \mathbb{R}^3$ und $F(x) := x_1 x_2^2 + 10 x_3^5 x_2$, wobei $x = (x_1, x_2, x_3)^T$. Dann sind

$$\frac{\partial F(x)}{\partial x_1} = x_2^2, \quad \frac{\partial F(x)}{\partial x_2} = 2x_1 x_2 + 10 x_3^5, \quad \frac{\partial F(x)}{\partial x_3} = 50 x_3^4 x_2, \quad (\text{X.5})$$

die partiellen Ableitungen, und der Gradient von F bei x ist

$$\nabla F(x) := \begin{pmatrix} x_2^2 \\ 2x_1 x_2 + 10 x_3^5 \\ 50 x_3^4 x_2 \end{pmatrix} \in \mathbb{R}^3. \quad (\text{X.6})$$

- Wir betrachten $U = \mathbb{R}^2$ und $F(x) := \sin(x_1^2 - x_2^2)$, wobei $x = (x_1, x_2)^T$. Dann sind

$$\frac{\partial F(x)}{\partial x_1} = 2x_1 \cos(x_1^2 - x_2^2), \quad \frac{\partial F(x)}{\partial x_2} = -2x_2 \cos(x_1^2 - x_2^2), \quad (\text{X.7})$$

die partiellen Ableitungen, und der Gradient von F bei x ist

$$\nabla F(x) := \begin{pmatrix} 2x_1 \cos(x_1^2 - x_2^2) \\ -2x_2 \cos(x_1^2 - x_2^2) \end{pmatrix} \in \mathbb{R}^2. \quad (\text{X.8})$$

Definition X.2. Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ offen, $V \subseteq \mathbb{R}^N$ offen und $F = (F_1, F_2, \dots, F_N) : U \rightarrow V$.

(i) F ist bei $x \in U$ **partiell differenzierbar**

$$:\Leftrightarrow \quad \forall i \in \mathbb{Z}_1^N : \quad F_i \text{ ist bei } x \text{ partiell differenzierbar.} \quad (\text{X.9})$$

(ii) F ist **auf U partiell differenzierbar**

$$:\Leftrightarrow \quad \forall x \in U : \quad F \text{ ist bei } x \text{ partiell differenzierbar.} \quad (\text{X.10})$$

In diesem Fall heißt die reelle $N \times M$ -Matrix

$$J_F(x) := \begin{pmatrix} \frac{\partial F_1(x)}{\partial x_1} & \cdots & \frac{\partial F_1(x)}{\partial x_M} \\ \vdots & & \vdots \\ \frac{\partial F_N(x)}{\partial x_1} & \cdots & \frac{\partial F_N(x)}{\partial x_M} \end{pmatrix} \in \mathfrak{M}_{N \times M}(\mathbb{R}) \quad (\text{X.11})$$

der partiellen Ableitungen **Jacobimatrix von F bei x**.

Bemerkungen und Beispiele.

- Wir betrachten $U = \mathbb{R}^+ \times (-\pi, \pi) \ni (r, \varphi)$ und

$$F(r, \varphi) := \begin{pmatrix} F_1(r, \varphi) \\ F_2(r, \varphi) \end{pmatrix} := \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix}. \quad (\text{X.12})$$

Dann ist

$$J_F(x) = \begin{pmatrix} \frac{\partial F_1(r, \varphi)}{\partial r} & \frac{\partial F_1(r, \varphi)}{\partial \varphi} \\ \frac{\partial F_2(r, \varphi)}{\partial r} & \frac{\partial F_2(r, \varphi)}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix}. \quad (\text{X.13})$$

Definition X.3. Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$, $V \subseteq \mathbb{R}^N$ offene Teilmengen und $F : U \rightarrow V$.

- (i) F heißt **stetig differenzierbar auf U** $:\Leftrightarrow$

$$F \text{ ist partiell differenzierbar auf } U \text{ und } \forall i \in \mathbb{Z}_1^N, j \in \mathbb{Z}_1^M : \frac{\partial F_i}{\partial x_j} \text{ ist stetig auf } U. \quad (\text{X.14})$$

Die Klasse der auf U stetig differenzierbaren Funktionen mit Werten in V bezeichnet man mit

$$C^1(U; V) := \{F : U \rightarrow V \mid F \text{ ist stetig differenzierbar auf } U\}. \quad (\text{X.15})$$

- (ii) Sei $k \in \mathbb{N}$. F heißt **k -mal stetig differenzierbar auf U** $:\Leftrightarrow$

F ist stetig differenzierbar auf U und

$$\forall j_1 \in \mathbb{Z}_1^M : \frac{\partial F}{\partial x_{j_1}} \text{ ist stetig differenzierbar auf } U \text{ und} \quad (\text{X.16})$$

$$\forall j_1, j_2 \in \mathbb{Z}_1^M : \frac{\partial}{\partial x_{j_2}} \left(\frac{\partial F}{\partial x_{j_1}} \right) \text{ ist stetig differenzierbar auf } U \text{ und}$$

\vdots

$$\forall j_1, \dots, j_{k-1} \in \mathbb{Z}_1^M : \frac{\partial}{\partial x_{j_{k-1}}} \left(\dots \frac{\partial}{\partial x_{j_2}} \left(\frac{\partial F}{\partial x_{j_1}} \right) \right) \text{ ist stetig differenzierbar auf } U.$$

In diesem Fall bezeichnet

$$\frac{\partial^k F(x)}{\partial x_{j_k} \partial x_{j_{k-1}} \dots \partial x_{j_1}} := \frac{\partial}{\partial x_{j_k}} \left[\frac{\partial}{\partial x_{j_{k-1}}} \left(\dots \left(\frac{\partial}{\partial x_{j_1}} F \right) \right) \right](x) \quad (\text{X.17})$$

die **k . partielle Ableitung von F nach x_{j_1}, \dots, x_{j_k}** . Die Klasse der auf U k -mal stetig differenzierbaren Funktionen mit Werten in V bezeichnet man mit

$$C^k(U; V) := \{F : U \rightarrow V \mid F \text{ ist } k\text{-mal stetig differenzierbar auf } U\}. \quad (\text{X.18})$$

(iii) Die Klasse der **unendlich oft (stetig) differenzierbaren** oder auch **glatten Funktionen auf U mit Werten in V** wird mit

$$C^\infty(U; V) := \bigcap_{k=1}^{\infty} C^k(U; V) \quad (\text{X.19})$$

bezeichnet, und

$$\begin{aligned} C_0^\infty(\mathbb{R}^M; V) &:= \{F \in C^\infty(\mathbb{R}^M; V) \mid \text{supp}(F) \text{ ist kompakt}\} \\ &= \{F \in C^\infty(\mathbb{R}^M; V) \mid \exists R < \infty \ \forall \vec{x} \in \mathbb{R}^M, |\vec{x}| \geq R: F(\vec{x}) = 0\} \end{aligned} \quad (\text{X.20})$$

sind die **unendlich oft differenzierbaren Funktionen mit kompaktem Träger**, wobei **Träger** $\text{supp}(F)$ **von F** die Menge $\text{supp}(F) := \{x \in \mathbb{R}^M \mid F(x) \neq 0\}$ bezeichnet, auf der F nicht verschwindet.

Bemerkungen und Beispiele.

- Wir betrachten abermals $U = \mathbb{R}^3$, $N = 1$, $V = \mathbb{R}$ und $F(x) := x_1 x_2^2 + 10 x_3^5 x_2$, wobei $x = (x_1, x_2, x_3)^T$. Dann sind

$$\frac{\partial F(x)}{\partial x_2} = 2x_1 x_2 + 10 x_3^5, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} = 2x_2, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_1 \partial x_2} = 0, \quad (\text{X.21})$$

$$\frac{\partial F(x)}{\partial x_1} = x_2^2, \quad \frac{\partial^2 F(x)}{\partial x_2 \partial x_1} = 2x_2, \quad \frac{\partial^2 F(x)}{\partial x_1 \partial x_2 \partial x_1} = 0. \quad (\text{X.22})$$

Der folgende Satz zeigt, dass stetige Differenzierbarkeit die Vertauschbarkeit der partiellen Ableitungen garantiert.

Satz X.4. Seien $M, k \in \mathbb{N}$, $k \geq 2$, $U \subseteq \mathbb{R}^M$ und $V \subseteq \mathbb{R}$ offene Teilmengen und $F \in C^k(U; V)$. Seien weiterhin $j_1, j_2, \dots, j_k \in \mathbb{Z}_1^M$ und sei $\pi \in \mathcal{S}_k$ eine Permutation von k Elementen. Dann gilt für jedes $x \in U$ dass

$$\frac{\partial^k F(x)}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_k}} = \frac{\partial^k F(x)}{\partial x_{j_{\pi(1)}} \partial x_{j_{\pi(2)}} \cdots \partial x_{j_{\pi(k)}}}. \quad (\text{X.23})$$

Beweis. Wir zeigen (X.23) zunächst für $k = 2$. Seien dazu $F \in C^2(U; V)$ sowie $\alpha, \beta \in \mathbb{Z}_1^M$ mit $\alpha < \beta$. Dann ist

$$\begin{aligned} \frac{\partial^2 F(x)}{\partial x_\alpha \partial x_\beta} &= \lim_{t \rightarrow 0} \left\{ \frac{1}{t} \left(\frac{\partial F}{\partial x_\beta}(x + te_\alpha) - \frac{\partial F}{\partial x_\beta}(x) \right) \right\} \\ &= \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \left\{ \frac{1}{st} \left[F(x + te_\alpha + se_\beta) - F(x + te_\alpha) - F(x + se_\beta) + F(x) \right] \right\} \\ &= \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \{G(t, s)\}, \end{aligned} \quad (\text{X.24})$$

und mit Vertauschen von α und β auch $\frac{\partial^2 F(x)}{\partial x_\beta \partial x_\alpha} = \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} \{G(t, s)\}$, wobei

$$G(t, s) := \frac{1}{st} \left[F(x + te_\alpha + se_\beta) - F(x + se_\beta) - F(x + te_\alpha) + F(x) \right], \quad (\text{X.25})$$

für festes $x \in U$. Wir definieren weiterhin, für festes $s \neq 0$,

$$\tilde{G}(t) := F(x + te_\alpha + se_\beta) - F(x + te_\alpha), \quad (\text{X.26})$$

so dass,

$$\begin{aligned} G(t, s) &= \frac{1}{st} [\tilde{G}(t) - \tilde{G}(0)] = \frac{1}{s} \frac{d\tilde{G}(\theta_t t)}{dt} \\ &= \frac{1}{s} \left[\frac{\partial F}{\partial x_\alpha}(x + \theta_t te_\alpha + se_\beta) - \frac{\partial F}{\partial x_\alpha}(x + \theta_t te_\alpha) \right] \\ &= \frac{\partial^2 F}{\partial x_\beta \partial x_\alpha}(x + \theta_t te_\alpha + \theta_s se_\beta), \end{aligned} \quad (\text{X.27})$$

nach zweimaliger Anwendung des Mittelwertsatzes der Differentialrechnung, wobei $\theta_t, \theta_s \in (0, 1)$. Gleichung (X.27) zeigt, dass $G(t, s)$ stetig fortsetzbar in (t, s) bei $(0, 0)$ ist, also gleichzeitig in $t \rightarrow 0$ und $s \rightarrow 0$ und unabhängig von der Reihenfolge der Limiten, da $F \in C^2(U; V)$. Also ist

$$\frac{\partial^2 F(x)}{\partial x_\alpha \partial x_\beta} = \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \{G(t, s)\} = \lim_{(t,s) \rightarrow (0,0)} G(t, s) = \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} \{G(t, s)\} = \frac{\partial^2 F(x)}{\partial x_\beta \partial x_\alpha}. \quad (\text{X.28})$$

Seien nun $k \geq 3$, $F \in C^k(U; V)$ und $1 \leq \ell \leq k - 1$. Dann ist

$$\hat{F} := \frac{\partial^{k-\ell-1} F}{\partial x_{j_{\ell+2}} \partial x_{j_{\ell+3}} \dots \partial x_{j_k}} \in C^{\ell+1}(U; V) \subseteq C^2(U; V), \quad (\text{X.29})$$

und damit gilt nach (X.28)

$$\frac{\partial^2 \hat{F}}{\partial x_{j_\ell} \partial x_{j_{\ell+1}}} = \frac{\partial^2 \hat{F}}{\partial x_{j_{\ell+1}} \partial x_{j_\ell}}. \quad (\text{X.30})$$

Setzen wir \hat{F} ein, so erhalten wir

$$\frac{\partial^k F}{\partial x_{j_1} \dots \partial x_{j_\ell} \partial x_{j_{\ell+1}} \dots \partial x_{j_k}} = \frac{\partial^k F}{\partial x_{j_1} \dots \partial x_{j_{\ell+1}} \partial x_{j_\ell} \dots \partial x_{j_k}}, \quad (\text{X.31})$$

also die Gültigkeit von (X.23) im Fall, dass π eine Transposition ist. Da sich jede Permutation als Komposition von Transpositionen schreiben lässt, folgt somit (X.23) auch für den allgemeinen Fall $\pi \in \mathcal{S}_K$. \square

Bemerkungen und Beispiele.

- Für $k = 2$ und $\alpha, \beta \in \{1, \dots, M\}$ behauptet (X.23), dass

$$\frac{\partial^2 F(x)}{\partial x_\alpha \partial x_\beta} = \frac{\partial^2 F(x)}{\partial x_\beta \partial x_\alpha}. \quad (\text{X.32})$$

- Seien $M = k = 2$ und $f(x_1, x_2) = x_1^3 x_2 + 5x_1 x_2^6$. Dann sind

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 3x_1^2 x_2 + 5x_2^6, \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1^3 + 30x_1 x_2^5, \quad (\text{X.33})$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = 3x_1^2 + 30x_2^5 = \frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}. \quad (\text{X.34})$$

X.2. Beschränkte lineare Operatoren

Definition X.5. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei normierte \mathbb{K} -Vektorräume. Wir bezeichnen mit

$$\mathcal{B}(X; Y) := \{ A : X \rightarrow Y \mid A \text{ linear, } \|A\|_{\mathcal{B}(X; Y)} < \infty \} \quad (\text{X.35})$$

den Raum der **beschränkten linearen Operatoren (von X nach Y)**, wobei

$$\|A\|_{\mathcal{B}(X; Y)} := \sup_{x \in X \setminus \{0\}} \left\{ \frac{\|Ax\|_Y}{\|x\|_X} \right\} = \sup_{x \in X, \|x\|_X=1} \{ \|Ax\|_Y \} \quad (\text{X.36})$$

die **Operatornorm von A** ist.

Bemerkungen und Beispiele.

- Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ drei Banachräume und $x \in X$, $A \in \mathcal{B}(X; Y)$ sowie $B \in \mathcal{B}(Y; Z)$. Aus der Definition (X.36) der Operatornorm folgt sofort, dass

$$\|Ax\|_Y \leq \|A\|_{\mathcal{B}(X; Y)} \cdot \|x\|_X, \quad (\text{X.37})$$

$$\|BA\|_{\mathcal{B}(X; Z)} \leq \|B\|_{\mathcal{B}(Y; Z)} \cdot \|A\|_{\mathcal{B}(X; Y)}. \quad (\text{X.38})$$

- Sind $(X, \|\cdot\|_X)$ ein normierter Raum und $(Y, \|\cdot\|_Y)$ ein Banachraum, so ist auch $(\mathcal{B}(X; Y), \|\cdot\|_{\mathcal{B}(X; Y)})$ ein Banachraum.
- Sind $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei Banachräume über \mathbb{K} und $A : X \rightarrow Y$ linear, so gilt folgende Äquivalenz:

$$\{ \|A\|_{\mathcal{B}(X; Y)} < \infty \} \quad \Leftrightarrow \quad \{ A : X \rightarrow Y \text{ ist stetig} \}. \quad (\text{X.39})$$

- Seien $M, N \in \mathbb{N}$ und $X = \mathbb{K}^M$, $Y = \mathbb{K}^N$. Wir wählen auf X und Y jeweils die ℓ^1 -Norm

$$\|x\|_X := |x_1| + \dots + |x_M|, \quad \|y\|_Y := |y_1| + \dots + |y_N|, \quad (\text{X.40})$$

wobei $x = (x_1, \dots, x_M)^T \in \mathbb{K}^M$ und $y = (y_1, \dots, y_N)^T \in \mathbb{K}^N$. Sei nun $A : \mathbb{K}^M \rightarrow \mathbb{K}^N$ eine lineare Abbildung mit der Matrixdarstellung $(A_{k,\ell})_{k \in \mathbb{Z}_1^N, \ell \in \mathbb{Z}_1^M} \in \mathfrak{M}_{N \times M}(\mathbb{K})$ bezüglich der Standardbasen in \mathbb{K}^M und \mathbb{K}^N . Für $x = (x_1, \dots, x_M)^T \in \mathbb{K}^M$ ist dann

$$\begin{aligned} \|Ax\|_Y &= \sum_{k=1}^N |(Ax)_k| = \sum_{k=1}^N \left| \sum_{\ell=1}^M A_{k,\ell} x_\ell \right| \leq \sum_{\ell=1}^M \sum_{k=1}^N |A_{k,\ell}| |x_\ell| \quad (\text{X.41}) \\ &\leq \left(\max_{1 \leq m \leq M} \sum_{k=1}^N |A_{k,m}| \right) \sum_{\ell=1}^M |x_\ell| = \left(\max_{1 \leq m \leq M} \left\{ \sum_{k=1}^N |A_{k,m}| \right\} \right) \|x\|_X. \end{aligned}$$

Also ist

$$\|A\|_{\mathcal{B}(X; Y)} = \sup_{x \in X \setminus \{0\}} \left\{ \frac{\|Ax\|_Y}{\|x\|_X} \right\} \leq \max_{1 \leq m \leq M} \left\{ \sum_{k=1}^N |A_{k,m}| \right\}. \quad (\text{X.42})$$

- Insbesondere ist jede lineare Abbildung von \mathbb{K}^M nach \mathbb{K}^N ein beschränkter linearer Operator,

$$\mathcal{B}(\mathbb{K}^M; \mathbb{K}^N) = \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N), \quad (\text{X.43})$$

und zwar unabhängig von der auf \mathbb{K}^M und \mathbb{K}^N definierten Norm.

X.3. Totale Differenzierbarkeit

Definition X.6. Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ zwei Banachräume, $U \subseteq X$ und $V \subseteq Y$ zwei nichtleere offene Teilmengen und $x \in U$.

- (i) Eine Abbildung $F : U \rightarrow V$ heißt **total differenzierbar bei x** .

$$:\Leftrightarrow \exists A \in \mathcal{B}(X; Y) : \lim_{z \rightarrow 0} \left\{ \frac{\|F(x+z) - F(x) - Az\|_Y}{\|z\|_X} \right\} = 0. \quad (\text{X.44})$$

In diesem Fall heißt

$$F'(x) := dF(x) := A \quad (\text{X.45})$$

totale Ableitung von F bei x .

- (ii) Die Abbildung $F : U \rightarrow V$ heißt **total differenzierbar auf U**

$$:\Leftrightarrow \forall x \in U : F \text{ ist differenzierbar bei } x. \quad (\text{X.46})$$

In diesem Fall ist $F' : U \rightarrow \mathcal{B}(X; Y)$.

Bemerkungen und Beispiele.

- Für $(X, \|\cdot\|_X) = (Y, \|\cdot\|_Y) = (\mathbb{R}, |\cdot|)$ sind die linearen Abbildungen von \mathbb{R} nach \mathbb{R} reelle 1×1 -Matrizen, also reelle Zahlen, $\mathcal{B}(\mathbb{R}; \mathbb{R}) \cong \mathbb{R}$. Nach Definition X.6 ist $f : \mathbb{R} \rightarrow \mathbb{R}$ somit genau dann total differenzierbar bei $x \in \mathbb{R}$, wenn es eine Zahl $f'(x) \in \mathbb{R}$ gibt, sodass

$$\lim_{z \rightarrow 0} \left\{ \frac{|f(x+z) - f(x) - f'(x) \cdot z|}{|z|} \right\} = \left| \lim_{z \rightarrow 0} \left\{ \frac{f(x+z) - f(x)}{z} \right\} - f'(x) \right| = 0. \quad (\text{X.47})$$

Dies ist aber offensichtlich gleichwertig mit der Existenz von

$$f'(x) := \lim_{z \rightarrow 0} \left\{ \frac{f(x+z) - f(x)}{z} \right\} \in \mathbb{R}. \quad (\text{X.48})$$

Für $X = Y = \mathbb{R}$ fällt also totale Differenzierbarkeit mit dem üblichen Begriff der Differenzierbarkeit einer reellen Funktion einer reellen Variablen zusammen.

- Weiterhin bemerken wir, dass die Ableitung einer total differenzierbaren Funktion eindeutig ist. Gilt nämlich (X.44) sowohl für $A \in \mathcal{B}(X; Y)$ als auch für $\tilde{A} \in \mathcal{B}(X; Y)$, so setzen wir $B := A - \tilde{A}$. Sind nun $z \in X \setminus \{0\}$ und $t > 0$, so beobachten wir, dass

$$\frac{\|Bz\|_Y}{\|z\|_X} = \frac{\|B(tz)\|_Y}{\|tz\|_X} = \lim_{t \rightarrow 0} \frac{\|B(tz)\|_Y}{\|tz\|_X}, \quad (\text{X.49})$$

da der zweite Ausdruck gar nicht von t abhängt. Somit ist aber

$$\begin{aligned} \frac{\|B(z)\|_Y}{\|z\|_X} &= \lim_{t \rightarrow 0} \left\{ \frac{\|(A - \tilde{A})(tz)\|_Y}{\|tz\|_X} \right\} \\ &\leq \lim_{t \rightarrow 0} \left\{ \frac{\|F(x + tz) - F(x) - A(tz)\|_Y}{\|tz\|_X} \right\} \\ &\quad + \lim_{t \rightarrow 0} \left\{ \frac{\|F(x + tz) - F(x) - \tilde{A}(tz)\|_Y}{\|tz\|_X} \right\} \\ &= 0 \end{aligned} \quad (\text{X.50})$$

also ist $\|Bz\|_Y = 0$, für alle $z \in X$, was $Bz = 0$ und daher $B = A - \tilde{A} = 0$ impliziert.

- Wir stellen fest, dass die totale Differenzierbarkeit (X.44) von $F : U \rightarrow V$ bei $x \in U$ auch äquivalent durch

$$\begin{aligned} \exists F'(x) \in \mathcal{B}(X; Y), \quad \delta > 0, \quad r_x : B(0, \delta) \rightarrow Y \quad \forall z \in B(0, \delta) : \quad (\text{X.51}) \\ F(x + z) - F(x) = F'(x)z + r_x(z) \quad \|z\|_X, \quad \lim_{z \rightarrow 0} \|r_x(z)\|_Y = 0, \end{aligned}$$

dargestellt werden kann.

Lemma X.7. *Sind $U \subseteq X$, $V \subseteq Y$ offene Teilmengen zweier Banachräume $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $F : U \rightarrow V$ total differenzierbar bei $x \in U$, so ist F auch stetig in $x \in U$.*

Satz X.8 (Kettenregel). *Seien $U \subseteq X$, $V \subseteq Y$ und $W \subseteq Z$ offene Teilmengen dreier Banachräume $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$, und seien $G : U \rightarrow V$ total differenzierbar bei $x \in U$ und $F : V \rightarrow W$ total differenzierbar bei $y := G(x) \in V$. Dann ist auch $F \circ G : U \rightarrow W$ total differenzierbar bei x , und es gilt*

$$(F \circ G)'(x) = F'[G(x)] \circ G'(x). \quad (\text{X.52})$$

Wir stellen nun den Zusammenhang zwischen der partiellen und der totalen Differenzierbarkeit einer Funktion her. Der folgende Satz zeigt, dass diese Begriffe zusammenfallen, wenn man zusätzlich die Stetigkeit der Ableitung fordert.

Satz X.9. *Seien $M, N \in \mathbb{N}$, $U \subseteq \mathbb{R}^M$ offen, $V \subseteq \mathbb{R}^N$ offen und $F : U \rightarrow V$. Dann sind folgende Aussagen äquivalent:*

$$F \text{ ist total differenzierbar auf } U, \text{ und } F' : U \rightarrow \mathcal{B}(\mathbb{R}^M; \mathbb{R}^N) \text{ ist stetig in } U. \quad (\text{X.53})$$

$$\Leftrightarrow F \in C^1(U; V), \text{ d.h. } F \text{ ist stetig differenzierbar auf } U. \quad (\text{X.54})$$

Gilt (X.53) oder (X.54), so ist

$$\forall x \in U : F'(x) = J_F(x), \quad (\text{X.55})$$

d.h. die Ableitung $F'(x)$ von F ist gegeben durch die Matrix $(\partial_{x_j} F_i)_{i \in \mathbb{Z}_1^N, j \in \mathbb{Z}_1^M}$ der partiellen Ableitungen, der Jacobimatrix $J_F(x)$ von F bei x .

Beweis. Wir beweisen zunächst (X.54) \Rightarrow (X.53), und zwar für den Spezialfall $N = 1$. Dazu wählen wir für \mathbb{R}^M wieder die ℓ^1 -Norm, $\|(x_1, \dots, x_M)^T\| := |x_1| + \dots + |x_M|$. Wir schreiben $x = \sum_{j=1}^M x_j e_j$, $z = \sum_{j=1}^M z_j e_j$ und $w_m = \sum_{j=1}^m z_j e_j$, für alle $0 \leq m \leq M$, wobei $e_i \in \mathbb{R}^M$ die kanonischen Basivektoren sind, sowie $\partial_m F(x) := \frac{\partial F(x)}{\partial x_m}$. und beachten, dass $w_M = z$ und $w_0 = 0$. Damit ist

$$\begin{aligned} & \left| F(x+z) - F(x) - \sum_{m=1}^M \partial_m F(x) z_m \right| \\ &= \left| \{F(x+w_M) - F(x+w_{M-1})\} + \{F(x+w_{M-1}) - F(x+w_{M-2})\} \right. \\ & \quad \left. + \dots + \{F(x+w_1) - F(x+w_0)\} - \sum_{m=1}^M \partial_m F(x) z_m \right| \\ &\leq \left| \sum_{m=1}^M \{F(x+w_m) - F(x+w_{m-1}) - \partial_m F(x) z_m\} \right| \\ &\leq \sum_{m=1}^M \left| F(x+w_{m-1} + z_m e_m) - F(x+w_{m-1}) - \partial_m F(x) z_m \right| \end{aligned} \quad (\text{X.56})$$

Weiterhin sind alle $\partial_m F$ auf einer offenen Umgebung $U \ni x$ stetig, und daher sichert eine genügend kleine Wahl von $\delta > 0$, dass

$$\forall 1 \leq m \leq M : \partial_m F \text{ ist stetig auf } x + K_\delta, \quad (\text{X.57})$$

wobei $K_\delta := [-\delta, \delta]^M$. Für $1 \leq m \leq M$ und $z \in K_\delta$ impliziert der Mittelwertsatz der Differentialrechnung die Existenz einer Zahl $0 < \vartheta_m < 1$, so dass

$$F(x+w_{m-1} + z_m e_m) - F(x+w_{m-1}) = \partial_m F(x+w_{m-1} + \vartheta_m z_m e_m) \cdot z_m, \quad (\text{X.58})$$

und wegen $w_{m-1} + \vartheta_m z_m e_m \in K_\delta$ erhalten wir daraus

$$\left| F(x+w_{m-1} + z_m e_m) - F(x+w_{m-1}) - \partial_m F(x) z_m \right| \leq Q(\delta) |z_m|, \quad (\text{X.59})$$

wobei

$$Q(\delta) := \max_{1 \leq m \leq M} \max_{y \in K_\delta} |\partial_m F(x+y) - \partial_m F(x)| \rightarrow 0, \quad (\text{X.60})$$

für $\delta \rightarrow 0$, wegen der Stetigkeit von $\partial_m F$ auf $U \supset x + K_\delta$. Setzen wir (X.60) in (X.56) ein, so erhalten wir

$$\frac{1}{\|z\|} \left| F(x+z) - F(x) - \sum_{m=1}^M \partial_m F(x) z_m \right| \leq \frac{Q(\delta)}{\|z\|} \left(\sum_{m=1}^M |z_m| \right) = Q(\delta) \rightarrow 0, \quad (\text{X.61})$$

für $\delta \rightarrow 0$. Dies beweist die totale Differenzierbarkeit von F bei x und $F'(x) = J_F(x)$ für $N = 1$. Letztere Gleichung ergibt auch die Stetigkeit von $x \mapsto F'(x)$, da $x \mapsto J_F(x)$ nach (X.54) stetig ist.

Um (X.54) \Rightarrow (X.53) für $N \geq 1$ zu beweisen, wenden wir (X.61) auf alle Komponenten F_1, F_2, \dots, F_N an und erhalten

$$\forall 1 \leq n \leq N : \quad \lim_{z \rightarrow 0} \left\{ \frac{1}{\|z\|} \left| F_n(x+z) - F_n(x) - \sum_{m=1}^M \partial_m F_n(x) z_m \right| \right\} = 0. \quad (\text{X.62})$$

Damit erhalten wir aber

$$\begin{aligned} & \lim_{z \rightarrow 0} \left\{ \frac{1}{\|z\|_{\mathbb{R}^M}} \left\| F_n(x+z) - F_n(x) - \sum_{m=1}^M \partial_m F_n(x) z_m \right\|_{\mathbb{R}^N} \right\} \\ &= \lim_{z \rightarrow 0} \left\{ \frac{1}{\|z\|_{\mathbb{R}^M}} \sum_{n=1}^N \left| F_n(x+z) - F_n(x) - \sum_{m=1}^M \partial_m F_n(x) z_m \right| \right\} = 0. \end{aligned} \quad (\text{X.63})$$

Dies beweist für $N \geq 1$ die totale Differenzierbarkeit von F bei x , $F'(x) = J_F(x)$ und die Stetigkeit von F' wie für $N = 1$.

Um (X.53) \Rightarrow (X.54) zu beweisen, wählen wir $j \in \mathbb{Z}_1^M$ und setzen $z := te_j$, sodass

$$\begin{aligned} 0 &= \lim_{z \rightarrow 0} \left\{ \frac{\|F(x+z) - F(x) - F'(x) \cdot z\|_{\mathbb{R}^N}}{\|z\|_{\mathbb{R}^M}} \right\} \\ &= \lim_{t \rightarrow 0} \left\{ \frac{1}{t} \|F(x+te_j) - F(x) - tF'(x)e_j\| \right\}. \end{aligned} \quad (\text{X.64})$$

Daher muss auch, für jedes $i \in \mathbb{Z}_1^N$

$$\begin{aligned} 0 &= \lim_{t \rightarrow 0} \left\{ \frac{1}{t} [F_i(x+te_j) - F_i(x) - (F'(x))_{i,j} \cdot t] \right\} \\ &= \lim_{t \rightarrow 0} \left\{ \frac{F_i(x+te_j) - F_i(x)}{t} \right\} - (F'(x))_{i,j} \end{aligned} \quad (\text{X.65})$$

gelten. Also ist F partiell differenzierbar auf U , die Jacobimatrix $J_F = F'$ ist gegeben durch die Ableitung von F , die gemäß (X.53) stetig von x abhängt. \square

Bemerkungen und Beispiele.

- Satz X.9 zeigt, dass totale und partielle Differenzierbarkeit zusammenfallen, sofern die Ableitungen stetig sind.
- Im Allgemeinen ist jedoch partielle Differenzierbarkeit der schwächere Begriff, und es gibt Funktionen, die partiell, aber nicht total differenzierbar in einem zu betrachteten Punkt $x \in U$ sind. Es gilt also nur

$$F \text{ ist total differenzierbar in } x \in U \Rightarrow F \text{ ist partiell differenzierbar in } x \in U, \quad (\text{X.66})$$

aber nicht notwendig die Umkehrung.

Satz X.9 zeigt auch, dass stetige Differenzierbarkeit auch im allgemeinen Kontext das richtige Konzept ist, und wir definieren:

Definition X.10. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei Banachräume über \mathbb{K} und $U \subseteq X$ und $V \subseteq Y$ zwei nichtleere, offene Teilmengen. Die Menge der **stetig auf U differenzierbaren Funktionen mit Werten in V** ist definiert als

$$C^1(U; V) := \left\{ F : U \rightarrow V \mid F \text{ ist total differenzierbar auf } U \text{ und } F' \in C[U; \mathcal{B}(X; Y)] \right\}. \quad (\text{X.67})$$

X.4. Ergänzungen

X.4.1. Eigenschaften der Operatornorm

Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ und $(Z, \|\cdot\|_Z)$ drei Banachräume und $x \in X$, $A \in \mathcal{B}(X; Y)$ sowie $B \in \mathcal{B}(Y; Z)$. Wir zeigen folgenden Eigenschaften:

$$\|Ax\|_Y \leq \|A\|_{\mathcal{B}(X; Y)} \cdot \|x\|_X, \quad (\text{X.68})$$

$$\|BA\|_{\mathcal{B}(X; Z)} \leq \|B\|_{\mathcal{B}(Y; Z)} \cdot \|A\|_{\mathcal{B}(X; Y)}. \quad (\text{X.69})$$

Um (X.68) zu zeigen, schätzen wir für $x \neq 0$ wie folgt ab,

$$\begin{aligned} \|Ax\|_Y &= \frac{\|Ax\|_Y}{\|x\|_X} \cdot \|x\|_X \leq \sup_{x' \neq 0} \left\{ \frac{\|Ax'\|_Y}{\|x'\|_X} \right\} \cdot \|x\|_X \\ &= \|A\|_{\mathcal{B}(X; Y)} \cdot \|x\|_X. \end{aligned} \quad (\text{X.70})$$

Weiterhin gilt für $x \in X \setminus \{0\}$ mit $Bx \neq 0$, dass

$$\begin{aligned} \frac{\|ABx\|_Z}{\|x\|_X} &= \frac{\|A(Bx)\|_Z}{\|Bx\|_Y} \cdot \frac{\|Bx\|_Y}{\|x\|_X} \leq \sup_{y \in Y \setminus \{0\}} \left\{ \frac{\|Ay\|_Z}{\|y\|_Y} \right\} \cdot \sup_{x \in X \setminus \{0\}} \left\{ \frac{\|Bx\|_Y}{\|x\|_X} \right\} \\ &= \|A\|_{\mathcal{B}(Y; Z)} \cdot \|B\|_{\mathcal{B}(X; Y)}, \end{aligned} \quad (\text{X.71})$$

was natürlich auch für $Bx = 0$ richtig ist und somit (X.69) impliziert.

X.4.2. Äquivalenz von Stetigkeit und Beschränktheit linearer Operatoren

Wir zeigen, dass für zwei Banachräume $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ und einen linearen Operator $A : X \rightarrow Y$ folgende Eigenschaften gleichwertig sind:

$$\{ \|A\|_{\mathcal{B}(X; Y)} < \infty \} \quad \Leftrightarrow \quad \{ A : X \rightarrow Y \text{ ist stetig} \}. \quad (\text{X.72})$$

“ \Rightarrow ”: Wir kürzen $\|A\|_{\mathcal{B}(X; Y)}$ mit $a := \|A\|_{\mathcal{B}(X; Y)}$ ab. Wegen

$$\|Ax - Ay\|_Y = \|A(x - y)\|_Y \leq a \cdot \|x - y\|_X \quad (\text{X.73})$$

ist

$$A(B_X(x, r)) \subseteq B_Y(Ax, ar), \quad (\text{X.74})$$

für alle $x \in X$ und $r > 0$. Für $\varepsilon > 0$ ist mit $\delta := \varepsilon a^{-1}$ also

$$\forall x \in X : \quad A(B_X(x, \delta)) \subseteq B_Y(Ax, \varepsilon). \quad (\text{X.75})$$

“ \Leftarrow ”: Das Urbild der offenen Kugel $B_Y(0, 1) := \{y \in Y \mid \|y\| < 1\}$ um 0 in Y ist wegen der Stetigkeit von A und $A(0_X) = 0_Y$ eine offene Umgebung $U \subseteq X$ von $0 \in U$. Also gibt es ein $r > 0$, so dass $B_X(0, r) \subseteq U = A^{-1}[B_Y(1, 0)]$. D.h.

$$A[B_X(0, r)] \subseteq B_Y(0, 1) \quad \Leftrightarrow \quad \forall \tilde{x} \in X, \quad \|\tilde{x}\|_X < r : \quad \|A\tilde{x}\|_Y < 1. \quad (\text{X.76})$$

Ist nun $x \in X \setminus \{0\}$, so setzen wir

$$\tilde{x} := \frac{r}{2\|x\|_X}x \quad (\text{X.77})$$

und beobachten, dass

$$\|\tilde{x}\|_X = \frac{r}{2\|x\|_X} \cdot \|x\|_X = \frac{r}{2} < r, \quad (\text{X.78})$$

also $\tilde{x} \in B_X(0, r)$. Somit ist $\|A\tilde{x}\|_Y < 1$ gemäß (X.76). Aus der Linearität von A erhalten wir schließlich

$$\frac{\|Ax\|_Y}{\|x\|_X} = \frac{\left(\frac{r}{2\|x\|_X}\right) \cdot \|Ax\|_Y}{\left(\frac{r}{2\|x\|_X}\right) \cdot \|x\|_X} = \frac{\|A\tilde{x}\|_Y}{\|\tilde{x}\|_X} \leq \frac{2}{r}, \quad (\text{X.79})$$

und somit

$$\|A\|_{\mathcal{B}(X;Y)} \leq \frac{2}{r} < \infty. \quad (\text{X.80})$$

X.4.3. Beweis der Kettenregel

Wir setzen (X.51) sukzessiv ein und erhalten

$$\begin{aligned} & (F \circ G)(x+z) - (F \circ G)(x) \\ &= F'[G(x)](G(x+z) - G(x)) + r_{F,G(x)}[G(x+z) - G(x)] \cdot \|G(x+z) - G(x)\|_Y \\ &= (F'[G(x)] \circ G'(x))(z) + F'[G(x)](r_{G,x}(z)) \cdot \|z\|_X + \\ & \quad r_{F,G(x)}[G(x+z) - G(x)] \cdot \|G(x+z) - G(x)\|_Y \\ &= (F'[G(x)] \circ G'(x))(z) + r_{F \circ G,x}(z) \|z\|_X, \end{aligned} \quad (\text{X.81})$$

mit

$$r_{F \circ G,x}(z) := F'[G(x)](r_{G,x}(z)) + r_{F,G(x)}[G(x+z) - G(x)] \frac{\|G(x+z) - G(x)\|_Y}{\|z\|_X}, \quad (\text{X.82})$$

und es gilt zu zeigen, dass

$$\lim_{z \rightarrow 0} \|r_{F \circ G}(x, z)\|_Z = 0. \quad (\text{X.83})$$

Da $F'(G(x)) \in \mathcal{B}(Y, Z)$, können wir (X.51) anwenden und erhalten

$$\lim_{z \rightarrow 0} \|F'[G(x)](r_{G,x}(z))\|_Z \leq \|F'(G(x))\|_{\mathcal{B}(Y;Z)} \cdot \lim_{z \rightarrow 0} \|r_G(x, z)\|_Y = 0, \quad (\text{X.84})$$

wobei wir die Differenzierbarkeit von G bei x ausnutzen. Abermals wegen der Differenzierbarkeit von G bei x gilt

$$\limsup_{z \rightarrow 0} \left\{ \frac{\|G(x+z) - G(x)\|_Y}{\|z\|_X} \right\} = \|G'(x)\|_{\mathcal{B}(X;Y)}. \quad (\text{X.85})$$

Nach Lemma X.7 ist G stetig bei x , und es gilt $\lim_{z \rightarrow 0} \{G(x+z) - G(x)\} = 0$. Daher ist

$$\lim_{z \rightarrow 0} r_{F,G(x)}[G(x+z) - G(x)] = 0, \quad (\text{X.86})$$

und (X.84)–(X.86) ergeben (X.52).