$$\|g_n(s) - f(s)\| \le \|g_n(s) - f_n(s)\| + \|f_n(s) - f(s)\|\,,$$

woraus folgt, dass für fast alle $s \in S$ gilt:

$$\lim_{n\to\infty} \|g_n(s) - f(s)\| = 0\,.$$

Aufgrund der Konstruktion haben wir für fast alle $s \in S$ und alle $n \in \mathbb{N}$:

$$\|g_n(s)\| \le \frac{3}{2}\,\|f(s)\|\,,$$

und erhalten somit durch

$$\|g_n(s) - f(s)\| \le \|g_n(s)\| + \|f(s)\| \le \frac{5}{2}\,\|f(s)\|$$

eine Lebesgue–integrierbare Majorante für die Folge $\big(\|g_n(\cdot) - f(\cdot)\|\big)$. Nach dem Satz von Lebesgue über majorisierte Konvergenz folgt damit

$$\lim_{n\to\infty} \int\limits_s \|g_n(s) - f(s)\|\,ds = 0\,,$$

d.h. f ist nach Definition Bochner–integrierbar. ∎

1.14 Folgerung. *Sei $f\colon S \to X$ Bochner–integrierbar. Dann gilt:*

$$\Big\| \int\limits_S f(s)\,ds \Big\|_X \le \int\limits_S \big\|f(s)\big\|_X\,ds\,, \tag{1.15}$$

und für alle $F \in X^*$

$$\Big\langle F, \int\limits_S f(s)\,ds \Big\rangle_X = \int\limits_S \big\langle F, f(s)\big\rangle_X\,ds\,. \tag{1.16}$$

Beweis. 1. Nach der Definition des Bochner–Integrals gilt für geeignete Treppenfunktionen (f_n):

$$\begin{aligned}
\Big\| \int\limits_S f(s)\,ds \Big\| &= \lim_{n\to\infty} \Big\| \int\limits_S f_n(s)\,ds \Big\| \\
&\le \lim_{n\to\infty} \int\limits_S \|f_n(s)\|\,ds \\
&\le \lim_{n\to\infty} \Big(\int\limits_S \|f_n(s) - f(s)\|\,ds + \int\limits_S \|f(s)\|\,ds \Big) \\
&= \int\limits_S \|f(s)\|\,ds\,,
\end{aligned}$$

wobei wir die Definition des Integrals von Treppenfunktionen und (1.5) benutzt haben.

2. Nach dem Beweis von Satz 1.12 gibt es eine Folge (f_n) von Treppenfunktionen, die fast überall gegen f konvergiert und für die gilt:

$$\|f_n(s)\| \le \tfrac{3}{2}\|f(s)\| \,. \tag{1.17}$$

Aufgrund der Definition des Bochner–Integrals haben wir für alle $F \in X^*$

$$\begin{aligned}
\Big\langle F, \int_S f(s)\,ds \Big\rangle &= \lim_{n\to\infty} \Big\langle F, \int_S f_n(s) \Big\rangle\,ds \\
&= \lim_{n\to\infty} \int_S \langle F, f_n(s)\rangle\,ds \\
&= \int_S \langle F, f(s)\rangle\,ds\,,
\end{aligned}$$

wobei im 2. Schritt benutzt wurde, dass das Integral von Treppenfunktionen eine endliche Linearkombination von Elementen aus X ist und F ein lineares stetiges Funktional ist.[1] Im letzten Schritt wurde (1.17) und der Satz von der majorisierten Konvergenz von reellwertigen Funktionen benutzt. ■

• Sei I ein beschränktes Intervall. Für Funktionen $f \in C(\bar{I};X)$ existiert das Bochner–Integral nach Satz 1.12, da $\|f(\cdot)\|\colon \bar{I} \to \mathbb{R}$ Lebesgue–integrierbar ist. Es ergibt sich als Grenzwert *Riemannscher Summen*:

$$\int_I f(s)\,ds = \lim_{n\to\infty} \sum_{j=0}^{n} f(\hat{t_j^n})(t_{j+1}^n - t_j^n)\,, \tag{1.18}$$

wobei $\hat{t_j^n} \in [t_j^n, t_{j+1}^n]$. In der Tat, da f auf $\overline{I}$ gleichmäßig stetig ist, gilt offensichtlich

$$f_n(s) := \sum_{j=0}^{n} f(\hat{t_j^n})\chi_{[t_j^n, t_{j+1}^n]}(s) \rightrightarrows f(s) \qquad (n\to\infty)\,, \tag{1.19}$$

d.h. es existieren Treppenfunktionen (f_n), die gleichmäßig gegen f konvergieren. Das Bochner–Integral dieser Treppenfunktionen ist offensichtlich die Summe auf der rechten Seite in (1.18). Aus (1.19) ergibt sich sofort, dass $(\int_{\bar{I}} f_n(s)\,ds)$ eine Cauchyfolge ist und somit die Bedingung (1.5) erfüllt ist.

[1] Es gilt:

$$\langle F, \int_S f_n(s)\,ds\rangle = \langle F, \sum_{i=1}^{N} m(B_i^n)\,x_i^n\rangle = \sum_{i=1}^{N} m(B_i^n)\langle F, x_i^n\rangle = \int_S \langle F, f_n(s)\rangle_X\,ds\,.$$

2.1.1 L^p–Räume mit Werten in Banachräumen

Das Bochner–Integral ist analog zum Lebesgue–Integral definiert. Man kann deshalb viele Resultate für Lebesgue–Integrale auf Bochner–Integrale übertragen, wobei die obigen Sätze, die einen Zusammenhang zwischen dem Bochner– und dem Lebesgue–Integral herstellen, nützlich sind. Wir wollen dies an einigen Beispielen illustrieren.

1.20 Definition. *Wir bezeichnen mit*

$$L^p(S;X)\,, \qquad\qquad 1 \le p < \infty\,,$$

die Menge aller Bochner–messbaren Funktionen, für die

$$\int_S \|f(s)\|_X^p \, ds < \infty\,.$$

Die Menge aller Bochner–messbaren Funktionen, für die eine Konstante M existiert so, dass für fast alle $s \in S$ gilt:

$$\|f(s)\|_X \le M$$

bezeichnen wir mit $L^\infty(S;X)$.

• Zur Illustration wollen wir für ein glattes beschränktes Gebiet $\Omega \subset \mathbb{R}^n$ und ein beschränktes Intervall $I = [0,T]$ den Raum $L^p(Q_T)$ betrachten, wobei $Q_T := \Omega \times I$ und $1 \le p < \infty$. Für $f \in L^p(Q_T)$ erhalten wir mithilfe des Satzes von Fubini, dass für fast alle $t \in I$ die Funktion

$$f(t)\colon \Omega \to \mathbb{R}\colon x \mapsto f(x,t)\,,$$

zum Raum $L^p(\Omega)$ gehört. In der Tat haben wir

$$\int_0^T \int_\Omega |f(x,t)|^p \, dx \, ds = \int_0^T \|f(t)\|_{L^p(\Omega)}^p \, ds\,, \tag{1.21}$$

was insbesondere zeigt, dass die Funktion

$$f\colon I \to L^p(\Omega)\colon t \mapsto f(t)$$

zum Raum $L^p(I;L^p(\Omega))$ gehört. Andererseits erhalten wir mithilfe von (1.21) sofort, dass für $f \in L^p(I;L^p(\Omega))$ gilt: $f \in L^p(Q_T)$. Wir haben somit gezeigt, dass $L^p(Q_T) = L^p(I;L^p(\Omega))$. Weitere Beispiele für L^p–Räume mit Werten in Banachräumen werden wir kennenlernen, wenn wir uns mit parabolischen Differentialgleichungen beschäftigen.

1.22 Satz. *Die Menge* $L^p(S;X), 1 \le p \le \infty$*, bildet einen Banachraum bezüglich der Norm*

$$\|f\|_{L^p(S;X)} := \left(\int_S \|f(s)\|_X^p \, ds \right)^{\frac{1}{p}}$$

bzw.

$$\|f\|_{L^\infty(S;X)} := \operatorname*{ess\,sup}_S \|f(s)\|_X .$$

Beweis. Die Eigenschaften der Norm sind leicht nachzurechnen. Die Vollständigkeit der Räume $L^p(S;X)$ kann wie folgt auf die Vollständigkeit der Räume $L^p(S;\mathbb{R})$ zurückgeführt werden: Sei (f_n) eine Cauchy–Folge in $L^p(S;X)$, d.h.

$$\|f_n - f_k\|_{L^p(S;X)}^p = \int_S \|f_n(s) - f_k(s)\|_X^p \, ds \to 0 \qquad (n,k \to \infty) .$$

Also ist nach Satz 1.12 die Folge der Funktionen $\big(\|f_n(\cdot) - f_k(\cdot)\|_X\big)$ eine Cauchyfolge in $L^p(S;\mathbb{R})$. Ab hier kann man dem Beweis im Falle von $L^p(S;\mathbb{R})$ folgen (cf. [**?**]). ■

1.23 Satz (Hölder–Ungleichung). *Sei* $f \in L^p(S;X)$, $g \in L^{p'}(S;X^*)$, $\frac{1}{p} + \frac{1}{p'} = 1$. *Dann ist* $\langle g(\cdot), f(\cdot)\rangle_X \in L^1(S;\mathbb{R})$*, und es gilt:*

$$\int_S \langle g(s), f(s)\rangle_X \, ds \le \|g\|_{L^p(S;X^*)} \|f\|_{L^{p'}(S;X)} .$$

Beweis. Seien (f_n), (g_n) Folgen von Treppenfunktionen so, dass für fast alle $s \in S$ gilt:

$$f_n \to f \text{ in } X , \quad g_n \to g \text{ in } X^* , \qquad (n \to \infty) .$$

Somit gilt für fast alle $s \in S$

$$\langle g_n(s), f_n(s)\rangle_X \to \langle g(s), f(s)\rangle_X \qquad (n \to \infty) ,$$

d.h. auch die Funktion $\langle g(\cdot), f(\cdot)\rangle_X$ ist Lebesgue–messbar, und es gilt

$$\begin{aligned} \int_S \big\langle g(s), f(s)\big\rangle_X \, ds &\le \int_S \|g(s)\|_{X^*} \|f(s)\|_X \, ds \\ &\le \|g\|_{L^p(S;X^*)} \|f\|_{L^{p'}(S;X)} \end{aligned}$$

aufgrund der Hölder–Ungleichung für Funktionen mit Werten in $\mathbb{R}$. ■

1.24 Satz. *Sei* X *ein reflexiver Banachraum und* $1 < p < \infty$. *Dann besitzt jedes Funktional* $F \in (L^p(S;X))^*$ *genau eine Darstellung der Form*

$$F(u) = \int_S \langle v(s), u(s) \rangle_X \, ds \qquad \textit{für alle } u \in L^p(S; X)\,,$$

wobei $v \in L^{p'}(S; X^*)$, $\frac{1}{p} + \frac{1}{p'} = 1$.

Beweis. Der sehr technische Beweis verläuft analog zum Beweis im Falle von Funktionen $f\colon S \to \mathbb{R}$ (cf. [**?**, Kap. 4]). ∎

2.2 Differentiation von Funktionen mit Werten in Banachräumen

Bevor wir uns mir Ableitungen von Funktionen mit Werten in Banachräumen beschäftigen wollen wir an die Definition des Raumes der stetigen Funktionen erinnern. Seien X, Y Banachräume und sei $U \subseteq X$ eine Teilmenge. Die Menge aller stetigen Funktionen $f\colon U \to Y$ bezeichnen wir mit $C(U; Y)$, welche in natürlicher Weise einen Vektorraum bildet. Man beachte, dass z.B. für offene Mengen U eine stetige Funktion $f\colon U \to Y$ *nicht* beschränkt sein muss. Falls man die Menge der beschränkten, stetigen Funktionen mit der kanonischen Norm

$$\|f\|_0 := \sup_{x \in U} \|f(x)\|_Y$$

versieht, ist

$$\big(C_b(U; Y), \|\cdot\|_0\big) := \big(\{f\colon U \to Y \,\big|\, \|f\|_0 < \infty\}, \|\cdot\|_0\big)$$

ein Banachraum. Dies beweist man analog zum Fall reellwertiger Funktionen. Im Falle einer *kompakten* Menge $U \subset X$ gilt: $C_b(U; Y) = C(U; Y)$.

Die Konvergenz einer Folge (f_n) von Funktionen aus $C_b(U; Y)$ bezüglich der Norm $\|\cdot\|_0$ ist nichts anderes als die auf U gleichmäßige Konvergenz von f_n gegen f, $(n \to \infty)$, im Banachraum Y.

2.1 Definition. *Seien X, Y Banachräume, $f\colon X \to Y$ und $h \in X$, $x_0 \in X$ gegeben. Falls die Abbildung $\varphi\colon \mathbb{R} \to Y$, definiert durch*

$$\varphi(t) := f(x_0 + th),$$

in $t = 0$ differenzierbar ist, d.h.

$$\varphi'(0) = \lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t} \in Y\,, \tag{2.2}$$

*sagen wir, dass f im Punkt x_0 eine **Ableitung in Richtung** h besitzt, die wir mit $\delta f(x_0, h)$ bezeichnen. Falls $\delta f(x_0, h)$ für alle $h \in X$ existiert und die Abbildung*

$$Df(x_0)\colon X \to Y : h \mapsto \delta f(x_0, h)$$

*stetig und linear ist, sagen wir, dass f im Punkt x_0 **Gâteaux–differenzierbar** ist und nennen $Df(x_0)$ die **Gâteaux–Ableitung** von f im Punkt x_0.*

• Die Definition (2.1) stimmt im Fall von $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ mit der Definition der Richtungsableitung reeller vektorwertiger Funktionen überein. Insbesondere gilt im Falle $m = 1$

$$\delta f(x_0, \mathbf{e}_i) = \frac{\partial f}{\partial x_i}(x_0)\,,$$

wobei $\mathbf{e}$, $i = 1, \ldots, n$, die kanonische Orthonormalbasis des $\mathbb{R}^n$ ist.

• Falls $\delta f(x_0, h)$ existiert, ist f im Punkt x_0 stetig in Richtung h, d.h.

$$\lim_{t\to 0} f(x_0 + th) = f(x_0)\,.$$

• Indem wir die Bezeichnung $r(x) = o(\|x\|)$ einführen, können wir die Definition der Gâteaux–Ableitung umschreiben. Wir definieren:

$$r(x) = o(\|x\|) \qquad \Leftrightarrow \qquad \frac{\|r(x)\|}{\|x\|} \to 0\,, \qquad (x \to 0)\,.$$

Damit schreibt sich die Bedingung (2.2) in Definition 2.1 wie folgt:

$$f(x_0 + th) - f(x_0) = \varphi'(0)\,t + o(t)\,.$$

• Falls $\delta f(x_0, h)$ existiert, dann gilt für alle $y^* \in Y^*$:

$$\frac{d}{dt}\langle y^*, f(x_0 + th)\rangle_Y\Big|_{t=0} = \langle y^*, \delta f(x_0, h)\rangle_Y\,.$$

2.3 Satz (Mittelwertsatz). *Seien X, Y Banachräume und seien $x_0, h \in X$ gegeben. Ferner sei $f\colon X \to Y$ für $0 \le t \le 1$ in $f(x_0 + th)$ Gâteaux-differenzierbar und sei die Abbildung $t \mapsto \delta f(x_0 + th, h)$ auf den Intervall $[0, 1]$ stetig. Dann gilt:*

$$f(x_0 + h) - f(x_0) = \int_0^1 \delta f(x_0 + th, h)\,dt\,.$$

Beweis. Seien $y^* \in Y^*$ und $h \in X$ gegeben. Wir definieren Funktionen $g\colon [0, 1] \to Y$ und $j\colon [0, 1] \to \mathbb{R}$ durch:

$$g(t) = f(x_0 + th)\,, \qquad j(t) = \langle y^*, g(t)\rangle_Y\,.$$

Nach Voraussetzung sind beide Funktionen nach t differenzierbar und es gilt:

$$g'(t) = \delta f(x_0 + th, h)\,, \qquad j'(t) = \langle y^*, g'(t)\rangle_Y\,.$$

Für die Funktion j gilt nach dem Mittelwertsatz in $\mathbb{R}$

$$j(1) - j(0) = \int_0^1 j'(t)\,dt\,,$$

was wir unter Benutzung der Definitionen von j und g umschreiben können als

$$\begin{aligned}\langle y^*, f(x_0+h) - f(x_0)\rangle_Y &= \int_0^1 \big\langle y^*, \delta f(x_0+th, h)\big\rangle_Y \, dt \\ &= \Big\langle y^*, \int_0^1 \delta f(x_0+th, h)\, dt\Big\rangle_Y ,\end{aligned}$$

wobei wir (1.16) benutzt haben. Da $y^* \in Y^*$ beliebig war, folgt aufgrund einer Folgerung des Satzes von Hahn–Banach

$$f(x_0+h) - f(x_0) = \int_0^1 \delta f(x_0+th, h)\, dt ,$$

also die Behauptung. ∎

2.4 Definition. *Seien X, Y Banachräume und sei $f\colon B_r(x_0) \subset X \to Y$ eine Funktion. Dann ist f* ***Fréchet–differenzierbar*** *im Punkt $x_0 \in X$ genau dann, wenn eine stetige lineare Abbildung $A\colon X \to Y$ existiert so, dass*

$$f(x_0+h) - f(x_0) = Ah + o(\|h\|) , \qquad (h \to 0) .$$

Wenn diese Abbildung existiert, nennen wir sie ***Fréchet–Ableitung*** *von f in x_0 und bezeichnen sie mit $f'(x_0) =: A$.*

• Die Funktion $f\colon X \to Y$ heißt **stetig differenzierbar** im Punkt x_0, falls die Abbildung

$$f'\colon B_r(x_0) \subset X \to L(X,Y)\colon x \mapsto f'(x)$$

im Punkt x_0 stetig ist. Hierbei bezeichnet $L(X,Y)$ den Raum der stetigen linearen Abbildungen $A\colon X \to Y$, der mit der kanonischen Operatornorm

$$\|A\|_{L(X,Y)} := \sup_{\|x\|_X \le 1} \|Ax\|_Y ,$$

versehen, einen Banachraum bildet. Wenn f in jedem Punkt einer *offenen* Menge $U \subseteq X$ stetig differenzierbar ist, schreiben wir $f \in C^1(U;Y)$.

2.5 Satz. *Seien X, Y Banachräume. Dann gilt für $f\colon X \to Y$:*

(i) *Ist f im Punkt x_0 Fréchet–differenzierbar, dann ist f in x_0 Gâteaux–differenzierbar.*

(ii) *Ist f Gâteaux–differenzierbar in einer Umgebung $U(x_0)$, und ist $Df(x)$ stetig in x_0, dann ist f Fréchet–differenzierbar in x_0.*

Beweis. 1. Die Behauptung (i) folgt sofort, wenn wir in der Definition der Fréchet–Ableitung h als $t\tilde{h}$, $\tilde{h}$ fest aber beliebig, wählen und dann $t \to 0$ gehen lassen.

2. Für die Funktion $g(\tau) := f(x_0 + \tau h)$ erhalten wir

$$g'(\tau) = \delta f(x_0 + \tau h, h)\,.$$

Nach dem Mittelwertsatz 2.3 ergibt sich dann

$$\begin{aligned}
\|g(1) - g(0) - g'(0)\| &= \Big\| \int_0^1 \delta f(x_0 + th, h) - \delta f(x_0, h)\, dt \Big\| \\
&\le \int_0^1 \big\| \big(Df(x_0 + th) - Df(x_0)\big) h \big\|\, dt \\
&\le \int_0^1 \big\| Df(x_0 + th) - Df(x_0) \big\|\, \|h\|\, dt \\
&= o(\|h\|)\,,
\end{aligned}$$

da Df stetig in x_0 ist. Also ist f im Punkt x_0 Fréchet–differenzierbar. ∎

Auch Ketten- und Produktregel gelten ähnlich wie im $\mathbb{R}^n$. Doch wir müssen zunächst einmal klären, was ein Produkt ist.

2.6 Definition. *Seien X, Y und W Banachräume. Eine Abbildung $B\colon X \times Y \to W$ nennen wir **Produkt**, wenn B folgende Bedingungen erfüllt:*

(i) *B ist bilinear, d.h. linear in beiden Komponenten,*

(ii) *B ist stetig, d.h. es existiert eine Konstante c, so dass für alle $x \in X$ und alle $y \in Y$ gilt:*

$$\|B(x,y)\|_W \le c\|x\|_X \|y\|_Y\,.$$

2.7 Satz (Ketten– und Produktregel). *Seien X, Y, W, Z Banachräume.*

(i) *Seien $U \subseteq X$, $V \subseteq Y$ offene Mengen und seien $f \in C^1(U;Y)$ und $g \in C^1(V;Z)$ derart, dass $f(U) \subseteq V$. Dann ist $g \circ f \in C^1(U;Z)$, und es gilt:*

$$(g \circ f)'(x) = g'\big(f(x)\big) \circ f'(x)\,.$$

(ii) *Sei $U \subseteq X$ offen. Die Funktionen $f\colon U \to Y$ und $g\colon U \to Z$ seien in U Fréchet–differenzierbar und $B\colon Y \times Z \to W$ sei ein Produkt. Dann ist die Funktion $h\colon U \to W : x \mapsto B(f(x), g(x))$ Fréchet–differenzierbar, und es gilt für alle $y \in X$*

$$h'(x)y = B(f'(x)y, g(x)) + B(f(x), g'(x)y)\,.$$

Beweis. 1. Nach Voraussetzung ist g in $y := f(x)$ Fréchet–differenzierbar, d.h.

$$g(y+k) = g(y) + g'(y)k + \|k\|\, r_1(k)\,,$$

wobei $r_1(k) \to 0$ für $k \to 0$. Wir wählen

$$k = k(h) = f(x+h) - f(x) = f'(x)h + \|h\|\, r_2(h)\,,$$

wobei $r_2(h) \to 0$ für $h \to 0$, was aufgrund der Fréchet–Differenzierbarkeit von f in x gilt. Insgesamt erhalten wir dann

$$g(f(x+h)) = g(f(x)) + g'(f(x))f'(x)h + \|h\|\, r(h)\,,$$

wobei für $\|h\|\, r(h) := g'(f(x))\|h\|\, r_2(h) + \big\|f'(x)h + \|h\|\, r_2(h)\big\|\, r_1(k)$ gilt:

$$\big\|\, \|h\|\, r(h)\big\| \le \|h\| \big\|g'(f(x))\, r_2(h) + \|f'(x) + r_2(h)\|\, r_1(k)\big\|\,.$$

Unter Beachtung von $k(h) \to 0$ für $h \to 0$ und $r_1(k) \to 0$ für $k \to 0$ sowie $r_2(h) \to 0$ für $h \to 0$ erhalten wir $r(h) \to 0$ für $h \to 0$. Somit ist die Kettenregel bewiesen.

2. Sei $y \in X$. Dann gilt nach Voraussetzung:

$$\begin{aligned} f(x+y) &= f(x) + f'(x)y + \|y\| r_1(y)\,,\\ g(x+y) &= g(x) + g'(x)y + \|y\| r_2(y)\,, \end{aligned}$$

wobei $r_1(y) \to 0$, $r_2(y) \to 0$ für $y \to 0$. Wir betrachten nun

$$\begin{aligned} h(x+y) - h(x) &= B(f(x+y), g(x+y)) - B(f(x), g(x))\\ &= B\big(f(x) + f'(x)y + \|y\| r_1(y), g(x) + g'(x)y + \|y\| r_2(y)\big)\\ &\quad - B(f(x), g(x))\\ &= B(f(x), g'(x)y) + B(f(x), r_2(y))\|y\|\\ &\quad + B(f'(x)y, g(x)) + B\big(f'(x)y, g'(x)y + \|y\| r_2(y)\big)\\ &\quad + \|y\| B\big(r_1(y), g(x) + g'(x)y + \|y\| r_2(y)\big)\\ &=: B\big(f(x), g'(x)y\big) + B\big(f'(x)y, g(x)\big) + \|y\|\, r(y)\,, \end{aligned}$$

wobei die Bilinearität von B benutzt wurde. Aufgrund der Stetigkeit von B erhalten wir $r(y) \to 0$ für $y \to 0$. Es gilt z.B.

$$\big\|B\big(f(x), r_2(y)\big)\big\| \le c\, \|f(x)\|\, \|r_2(y)\| \to 0 \qquad \text{für } y \to 0\,.$$

Somit ist auch die Produktregel bewiesen. ■

• **Ableitungen höherer Ordnung:** Für eine stetig differenzierbare Funktion $f\colon D(f) \subseteq X \to Y$ haben wir

$$f'\colon D(f') \subseteq X \to L(X,Y)\,.$$

Höhere Ableitungen erhalten wir indem wir die Definition 2.4 iterieren, z.B. ist f'' dann eine Abbildung

$$f''\colon D(f'') \subseteq X \to L\big(X, L(X,Y)\big)\,.$$

Wir sehen, dass die Bildräume der Ableitungen eine immer kompliziertere Struktur annehmen. Im weiteren benutzen wir für die ***n*–te Ableitung** folgende Notation:

$$f^{(n)}\colon X \times \underbrace{X \times \cdots \times X}_{n\text{–mal}} \to Y : (x, h_1, \ldots, h_n) \mapsto f^{(n)}(x, h_1, \ldots, h_n).$$

Wenn $h_i = h$ für alle $i = 1, \ldots, n$ ist, dann schreiben wir

$$f^{(n)}(x)h^n := f^{(n)}(x, \underbrace{h, \ldots, h}_{n\text{–mal}})\,.$$

Eine Funktion $f\colon X \to Y$ heißt ***n*–mal stetig differenzierbar** im Punkt x_0, falls die n–te Ableitung $f^{(n)}$ im Punkt x_0 stetig ist. Wir schreiben $f \in C^n(U;Y)$, falls f in jedem Punkt der *offenen* Menge U n–mal stetig differenzierbar ist.

- **Partielle Ableitungen:** Seien X, Y und Z Banachräume. Wir betrachten eine Funktion

$$f\colon X \times Y \to Z : (x,y) \mapsto f(x,y)\,.$$

Wenn wir $y_0 \in Y$ festhalten, ist $F(\cdot) := f(\cdot, y_0)\colon X \to Z$ eine Funktion in einer Variablen $x \in X$. Die **partielle Ableitung** von f nach x ist dann, analog zu den partiellen Ableitungen von Funktionen im $\mathbb{R}^n$, definiert durch:

$$\frac{\partial f}{\partial x}(x_0, y_0) := F'(x_0)\,.$$

Analog kann man $x_0 \in X$ festhalten und mithilfe von $G(\cdot) := f(x_0, \cdot)\colon Y \to Z$ die partielle Ableitung von f nach y definieren:

$$\frac{\partial f}{\partial y}(x_0, y_0) := G'(y_0)\,.$$

2.8 Satz (Taylor). *Sei $U \subseteq X$ offen, $f \in C^n(U;Y)$ und $x \in U$. Ferner existiere $f^{(n+1)}(x+th)$ für festes $h \in X$ und für alle $t \in [0,1]$. Dann gilt:*

$$f(x+h) = f(x) + f'(x)h + \ldots + \frac{1}{n!} f^{(n)}(x)h^n + R_{n+1}(x,h),$$

wobei $\|R_{n+1}(x,h)\| \le \sup\limits_{\tau \in [0,1]} \|\frac{1}{(n+1)!} f^{(n+1)}(x+\tau h)h^{n+1}\|$.

Beweis. Für $y^* \in Y^*$ setzen wir $g(t) = \langle y^*, f(x+th)\rangle \colon [0,1] \to \mathbb{R}$. Der Satz von Taylor in $\mathbb{R}$ liefert für ein $t_1 \in [0,1]$

$$\langle y^*, f(x+th)\rangle = \left\langle y^*, f(x) + f'(x)h + \ldots + \frac{1}{n!} f^{(n)}(x)h^n + R_{n+1} \right\rangle,$$

wobei $R_{n+1}(x,h) = \frac{1}{(n+1)!} f^{(n+1)}(x + t_1 h)h^{n+1}$. Die Behauptung folgt nun mithilfe des Satzes von Hahn–Banach. ■

2.2.1 Satz über implizite Funktionen

Unser Ziel ist es, eine Verallgemeinerung des Satzes über implizite Funktionen, den wir für reellwertige Funktionen kennen, zu beweisen.

Für eine Funktion $F \colon X \times Y \to Z$ wollen wir die Gleichung

$$F(x,y) = 0$$

in einer Umgebung $U(x_0,y_0)$ lösen, wobei für (x_0,y_0) gilt: $F(x_0,y_0) = 0$. Wir suchen also eine Abbildung $y \colon x \mapsto y(x)$, die auf einer Umgebung von x_0 definiert ist, mit

$$\begin{aligned} F(x,y(x)) &= 0\,, \\ y(x_0) &= y_0\,. \end{aligned}$$

2.9 Satz (Hildebrandt, Graves 1927). *Seien X, Y und Z Banachräume, $U(x_0,y_0) \subseteq X \times Y$ eine offene Umgebung von (x_0,y_0), $F \colon X \times Y \to Z$ sei in $U(x_0,y_0)$ definiert, und es sei $F(x_0,y_0) = 0$. Ferner existiere $\frac{\partial F}{\partial y}$ als Fréchet–Ableitung in $U(x_0,y_0)$, und es sei $\frac{\partial F}{\partial y}(x_0,y_0) \colon Y \to Z$ bijektiv. Außerdem seien F und $\frac{\partial F}{\partial y}$ in (x_0,y_0) stetig. Dann existieren $r_0, r > 0$ so, dass für alle $x \in X$, mit $\|x - x_0\| \le r_0$, genau ein $y(x) \in Y$ existiert mit*

$$\|y(x) - y_0\| \le r \quad \text{und} \quad F(x,y(x)) = 0\,.$$

Die Funktion $y(\cdot)$ ist in einer Umgebung von x_0 stetig.

Beweis. O.B.d.A. seien $x_0 = y_0 = 0$. Wir setzen

$$g(x,y) := \frac{\partial F}{\partial y}(0,0)y - F(x,y)\,.$$

Die Abbildung $g(\cdot,\cdot)$ ist stetig in $(0,0)$ und für $\|x\|$ klein genug ist $g(x,\cdot)$ stetig in einer Umgebung von 0. Aufgrund der Bijektivität von $\frac{\partial F}{\partial y}(0,0)$ ist dann die Gleichung $F(x,y(x)) = 0$ äquivalent zu folgender Gleichung:

$$y = \left(\frac{\partial F}{\partial y}(0,0)\right)^{-1} g(x,y) =: T_x y\,. \tag{2.10}$$

Für alle $\|x\| \le r_0$ und $\|y\|, \|z\| \le r$ gilt:

$$\frac{\partial g}{\partial y}(x,y) = \frac{\partial F}{\partial y}(0,0) - \frac{\partial F}{\partial y}(x,y)\,,$$

insbesondere gilt also:

$$\frac{\partial g}{\partial y}(0,0) = 0\,. \tag{2.11}$$

Da F und $\frac{\partial F}{\partial y}$ stetig in $(0,0)$ sind, folgt mit der Taylor–Formel für $n=0$ und $h = y - z$ und (2.11):

$$\begin{aligned} \|g(x,y) - g(x,z)\| &\le \sup_{0<\tau<1} \left\| \frac{\partial g}{\partial y}(x, z + \tau(y-z)) \right\| \, \|y - z\| \\ &= o(1)\|y-z\|\,, \qquad\qquad r \to 0\,, r_0 \to 0\,. \end{aligned} \tag{2.12}$$

Dies zusammen mit der Stetigkeit von g in $(0,0)$ und $g(0,0) = 0$ liefert:

$$\begin{aligned} \|g(x,y)\| &\le \|g(x,y) - g(x,0)\| + \|g(x,0)\| \\ &= o(1)\|y\| + o(1)\,, \qquad\qquad r \to 0, r_0 \to 0\,. \end{aligned} \tag{2.13}$$

Wir setzen $M := \{y \in Y \,|\, \|y\| \le r\}$. Aus (2.13) erhalten wir für alle $y \in M$ und alle $x \in \overline{B_{r_0}(0)} = \{x \in X \,|\, \|x\| \le r_0\}$

$$\begin{aligned} \|T_x y\| &\le \left\| \Big(\frac{\partial F}{\partial y}(0,0) \Big)^{-1} \right\| \, \|g(x,y)\| \\ &\le K\big(o(1)\|y\| + o(1)\big) \\ &\le r\,, \end{aligned}$$

falls r_0 und r klein genug gewählt wurden. Deshalb bildet T_x die Menge M in sich selbst ab, d.h. es gilt für alle $x \in \overline{B_{r_0}(0)}$:

$$T_x \colon M \to M\,.$$

Mithilfe von (2.12) erhalten wir, dass T_x eine k–Kontraktion ist. In der Tat gilt:

$$\begin{aligned} \|T_x y - T_x z\| &\le \left\| \Big(\frac{\partial F}{\partial y}(0,0) \Big)^{-1} \right\| \, \|g(x,y) - g(x,z)\| \\ &\le K o(1)\|y - z\| \\ &\le \frac{1}{2}\|y - z\|\,, \end{aligned}$$

wobei wir r eventuell noch kleiner als oben wählen. Nach dem Banachschen Fixpunktsatz gilt dann für alle $x \in X$ mit $\|x\| \le r_0$

$$\exists\,!\, y \in Y \text{ mit } \|y\| \le r: \quad y = T_x y\,.$$

Dies ist äquivalent dazu, dass die Gleichung

$$F(x, y(x)) = 0$$

eine eindeutige Lösung $y(x) \in Y$ besitzt. Die Stetigkeit von $y(\cdot)$ in 0 folgt sofort aus der Stetigkeit der Abbildung $(x, y) \mapsto T_x y$ (cf. (2.10)) in $(0, 0)$ mithilfe von Folgerung 1.1.12. ∎

2.14 Folgerung. *Zusätzlich zu den Voraussetzungen von Satz 2.9 sei F im Punkt (x_0, y_0) Fréchet–differenzierbar. Dann ist $y = y(x)$ in x_0 Fréchet–differenzierbar, und es gilt:*

$$y'(x_0) = -\left(\frac{\partial F}{\partial y}(x_0, y_0)\right)^{-1} \frac{\partial F}{\partial x}(x_0, y_0)\,.$$

Beweis. Nach Voraussetzung ist F in (x_0, y_0) Fréchet–differenzierbar, d.h.

$$\begin{aligned} F(x, y) = F(x_0, y_0) &+ \frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y - y_0) \\ &+ o(\|x - x_0\| + \|y - y_0\|)\,. \end{aligned}$$

Für $y = y(x)$ ergibt sich wegen $F(x_0, y_0) = 0$ und $F(x, y(x)) = 0$

$$\begin{aligned} 0 = &\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y(x) - y_0) \\ &+ r(x - x_0, y - y_0)\big(\|x - x_0\| + \|y - y_0\|\big)\,, \end{aligned}$$

mit $r(x - x_0, y - y_0) \to 0$ für $(x, y) \to (0, 0)$. Wenn wir diese Gleichung nach $y(x)$ auflösen, erhalten wir

$$\begin{aligned} y(x) - \underbrace{y(x_0)}_{= y_0} = -\left(\frac{\partial F}{\partial y}(x_0, y_0)\right)^{-1} &\left(\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0)\right. \\ &\left. + r(x - x_0, y - y_0)\big(\|x - x_0\| + \|y - y_0\|\big)\right)\,. \end{aligned} \tag{2.15}$$

Da y stetig in x_0 ist, haben wir $y(x) \to y(x_0)$ für $x \to x_0$. Somit folgt aus (2.15)

$$\|y(x) - y(x_0)\| \le K\,\|x - x_0\| + \tfrac{1}{2}\big(\|x - x_0\| + \|y - y_0\|\big)$$

für genügend kleine $\|x - x_0\|$. Dies impliziert

$$o(\|x - x_0\| + \|y(x) - y(x_0)\|) = o(\|x - x_0\|)$$

und wir erhalten insgesamt:

$$y(x) = y(x_0) - \left(\frac{\partial F}{\partial y}(x_0, y_0)\right)^{-1} \left(\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0)\right) + o(\|x - x_0\|)\,.$$