2 Implizite Funktionen

Wir betrachten jetzt den Fall eines unterbestimmten Systems, wenn es also weniger Gleichungen gibt als Unbekannte. Wir können die Funktion dann wie folgt schreiben, indem wir die Variablen in zwei Gruppen einteilen:

$$f : \Omega \to \mathbb{R}^k, \quad f = f(x,y), \quad \text{wobei } (x,y) \in \Omega \subset \mathbb{R}^m \times \mathbb{R}^k = \mathbb{R}^n.$$

Gegeben sei eine Lösung (x_0, y_0) der Gleichung $f(x,y) = z_0$. Wir interessieren uns dafür, wie die Lösungsmenge dieser Gleichung nahe bei (x_0, y_0) aussieht. Können wir nach y auflösen, d. h. die Lösungsmenge als Graph einer Funktion $y = g(x)$ darstellen?

Beispiel 2.1 Betrachte die Gleichung

$$f(x,y) = x^2 + y^2 = 1 \quad \text{für } (x,y) \in \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$$

Sei eine Lösung $(x_0, y_0) \in \mathbb{R}^2$ gegeben, also ein Punkt auf dem Einheitskreis. Ist $y_0 > 0$, so ist die Lösungsmenge nahe bei (x_0, y_0) Graph der Funktion $y = \sqrt{1-x^2}$. Analog haben wir im Fall $y_0 < 0$ die lokale Graphendarstellung $y = -\sqrt{1-x^2}$. Dagegen ist die Lösungsmenge in keiner Umgebung von $(1,0)$ (und ebenso in keiner Umgebung von $(-1,0)$) als Graph über der x-Achse darstellbar, denn es gibt die zwei Lösungen $y = \pm\sqrt{1-x^2}$.

Es kann auch vorkommen, dass (x_0, y_0) der einzige Punkt mit $f(x_0, y_0) = z_0$ ist, z.B. löst nur der Nullpunkt die Gleichung $x^2 + y^2 = 0$ in $\mathbb{R}^2$.

Beispiel 2.2 *Sei $f : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$ linear. Wir unterteilen die $k \times (m+k)$-Matrix von f in eine $k \times m$-Matrix A und eine $k \times k$-Matrix B, d. h.*

$$f(x,y) = Ax + By \quad \text{mit } A \in L(\mathbb{R}^m, \mathbb{R}^k), \quad B \in L(\mathbb{R}^k, \mathbb{R}^k).$$

Die Gleichung $Ax+By = z_0$ hat zu festem $x \in \mathbb{R}^m$ eine eindeutige Auflösung nach y dann und nur dann, wenn B invertierbar ist. Ist das der Fall, so lautet die Auflösung $y = B^{-1}(z_0 - Ax)$.

Allgemein schreiben wir die Jacobimatrix von $f = f(x,y)$ in der Form

$$Df(x,y) = (D_x f,\, D_y f) \in (\mathbb{R}^{k\times m},\, \mathbb{R}^{k\times k}).$$

Wenn wir nach $y = g(x)$ auflösen wollen, so sollte nach Beispiel 2.2 die Ableitung $D_y f(x_0, y_0) \in \mathbb{R}^{k\times k}$ invertierbar sein. In den Anwendungen ist die Einteilung in die beiden Variablengruppen nicht immer vorgegeben, das heißt es könnte nach verschiedenen Gruppen von je k Variablen aufgelöst werden. So kann der Einheitskreis in einer Umgebung von $(1,0)$ zwar nicht als Graph $y = g(x)$ geschrieben werden, wohl aber als Graph $x = g(y)$, und außer in den vier Punkten $\pm e_1, \pm e_2$ könnte sowohl nach x als auch nach y aufgelöst werden.

Merkregel. Die Ableitung nach den Variablen, nach denen aufgelöst werden soll, muss invertierbar sein. Im Spezialfall $k = 1$ bedeutet das $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

Satz 2.1 (über implizite Funktionen) *Sei $\Omega \subset \mathbb{R}^m \times \mathbb{R}^k$ offen und $f \in C^1(\Omega, \mathbb{R}^k)$. Ist $f(x_0, y_0) = z_0$ und $D_y f(x_0, y_0) \in L(\mathbb{R}^k, \mathbb{R}^k)$ invertierbar, so gibt es offene Umgebungen U von x_0 bzw. V von y_0 sowie eine Funktion $g \in C^1(U, V)$ mit*

$$\{(x,y) \in U \times V : f(x,y) = z_0\} = \{\big(x, g(x)\big) : x \in U\}, \tag{2.12}$$

insbesondere ist $g(x_0) = y_0$. Die Funktion g hat die Ableitung

$$Dg(x_0) = -\big(D_y f(x_0, y_0)\big)^{-1} D_x f(x_0, y_0). \tag{2.13}$$

Zusatz. Für jedes $r \in \mathbb{N} \cup \{\infty\}$ gilt die Implikation

$$f \in C^r(\Omega, \mathbb{R}^k) \quad \Rightarrow \quad g \in C^r(U, \mathbb{R}^k).$$

Beweis: Wir verwenden einen Trick, um den Satz über inverse Funktionen anwenden zu können, und zwar betrachten wir $F : \Omega \to \mathbb{R}^m \times \mathbb{R}^k$, $F(x,y) = \big(x, f(x,y)\big)$. Es gilt

$$DF = \begin{pmatrix} \mathrm{E}_m & 0 \\ D_x f & D_y f \end{pmatrix} \in \begin{pmatrix} \mathbb{R}^{m\times m} & \mathbb{R}^{m\times k} \\ \mathbb{R}^{k\times m} & \mathbb{R}^{k\times k} \end{pmatrix}.$$

Es folgt $\det DF(x_0,y_0) = \det D_y f(x_0,y_0) \neq 0$ nach Voraussetzung. Nach dem Umkehrsatz existieren offene Umgebungen $U_0 \times V$ von (x_0,y_0) sowie W von (x_0,z_0), so dass $F : U_0 \times V \to W$ diffeomorph ist. Wir bezeichnen die zugehörige Umkehrabbildung mit $G \in C^1(W, U_0 \times V)$. Ist $(x,z) \in W$, also $(x,z) = (x, f(x,y))$ mit $(x,y) \in U_0 \times V$ nach Konstruktion, so folgt

$$G(x,z) = G(x, f(x,y) = G(F(x,y)) = (x,y).$$

Also ist G von der Form $G(x,z) = (x, g_0(x,z))$ mit $g_0 \in C^1(W, \mathbb{R}^k)$. Sei nun $U = \{x \in U_0 : (x,z_0) \in W\}$. Da W offen in $\mathbb{R}^m \times \mathbb{R}^k$, ist U offen in $\mathbb{R}^m$ und für $(x,y) \in U \times V$ gilt

$$\begin{aligned} f(x,y) = z_0 \quad \Leftrightarrow \quad F(x,y) &= (x,z_0) \\ \Leftrightarrow \quad (x,y) &= G(x,z_0) \quad (\text{da } (x,z_0) \in W) \\ \Leftrightarrow \quad y &= g_0(x,z_0). \end{aligned}$$

Also gilt die Aussage des Satzes mit $g(x) = g(x,z_0)$. Die Formel für die Ableitung folgt aus der Kettenregel:

$$f\big(x, g(x)\big) = z_0 \quad \Rightarrow \quad D_x f(x_0,y_0) + D_y f(x_0,y_0)\, Dg(x_0) = 0.$$

□

Beispiel 2.3 Wir wollen als triviales Beispiel untersuchen, wie die Nullstelle eines quadratischen Polynoms von seinen Koeffizienten abhängt. Sei

$$f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \quad f(p,q,\lambda) = \lambda^2 + 2p\,\lambda + q = (\lambda + p)^2 - (p^2 - q).$$

Die Menge $N = \{(p,q,\lambda) \in \mathbb{R}^2 \times \mathbb{R} : f(p,q,\lambda) = 0\}$ ist die Vereinigung der beiden disjunkten Graphen $G^{\pm} = \{(p,q,\lambda^{\pm}(p,q)) : p^2 > q\}$, wobei $\lambda^{\pm}(p,q) = -p \pm \sqrt{p^2 - q}$, mit der Menge $\overline{G}^{+} \cap \overline{G}^{-} = \{(p,q,\lambda) : p^2 = q,\ \lambda = -p\}$. Im Fall

$$\frac{\partial f}{\partial \lambda}(p_0,q_0,\lambda_0) = 2(\lambda_0 + p_0) \neq 0 \quad \Leftrightarrow \quad p_0^2 - q_0 > 0$$

liegt (p_0,q_0,λ_0) in einem der beiden Graphen G^+ oder G^-, und N ist in einer Umgebung $U \times V$ als Graph von λ^+ oder λ^- darstellbar. Ist dagegen

$$\frac{\partial f}{\partial \lambda}(p_0,q_0,\lambda_0) = 2(\lambda_0 + p_0) = 0 \quad \Leftrightarrow \quad p_0^2 - q_0 = 0,$$

so macht der implizite Funktionensatz keine Aussage. Tatsächlich lässt sich die Menge N in keiner Umgebung von (p_0,q_0,λ_0) als Graph $\lambda = \lambda(p,q)$ darstellen: für $p^2 < q$ hat die Gleichung überhaupt keine Lösung, für $p^2 = q$ genau eine und für $p^2 > q$ die zwei verschiedenen Lösungen $\lambda^{\pm}(p,q)$.

Beispiel 2.4 Betrachte jetzt $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $f(b,\lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_0$. Sei λ_0 eine einfache Nullstelle von $f(a,\lambda)$ für $a \in \mathbb{R}^n$ fest, das heißt es gilt

$$f(a,\lambda) = (\lambda - \lambda_0)\, q(\lambda) \quad \text{für ein Polynom } q(\lambda) \text{ mit } q(\lambda_0) \neq 0.$$

Es folgt $\frac{\partial f}{\partial \lambda}(a,\lambda_0) = q(\lambda_0) \neq 0$. Nach dem Satz über implizite Funktionen existiert eine Umgebung $U \times V$ von (a,λ_0), so dass zu jedem $b \in U$ genau eine Nullstelle $\lambda(b) \in V$ von $f(b,\cdot\,)$ existiert. Diese hängt unendlich oft differenzierbar von b ab, und es gilt für $0 \leq i \leq n-1$

$$\frac{\partial \lambda}{\partial b_i}(a) = -\Big(\frac{\partial f}{\partial \lambda}(a,\lambda_0)\Big)^{-1} \frac{\partial f}{\partial b_i}(a,\lambda_0) = -\frac{\lambda_0^i}{n\,\lambda_0^{n-1} + (n-1)\,a_{n-1}\,\lambda_0^{n-2} + \ldots + a_1}.$$

Wir kommen nun zu einer geometrischen Anwendung des Satzes über implizite Funktionen. Ist $f \in C^1(\mathbb{R}^2)$ und $z_0 \in \mathbb{R}$, so kann im allgemeinen die Niveaumenge

$$M = \{(x,y) \in \mathbb{R}^2 : f(x,y) = z_0\}$$

Punkte enthalten, in denen M nicht lokal wie eine Linie aussieht, z.B. Kreuzungspunkte von Linien oder isolierte Punkte. Ist aber $Df(x,y) \neq 0$ für alle $(x,y) \in M$, so ist M nach dem impliziten Funktionensatz in der Nähe jedes Punkts als C^1-Graph über der x-Achse oder der y-Achse darstellbar und damit wirklich eine Höhenlinie im strengen Sinn des Worts. Teilmengen des $\mathbb{R}^n$, die lokal aussehen wie ein Unterraum, heißen Untermannigfaltigkeiten; vergleiche hierzu auch die Diskussion des Tangentialraums von differenzierbaren Graphen mittels Blowup, siehe Kapitel 6 Abschnitt 3.

Definition 2.1 *Sei $1 \leq m \leq n$. Eine Menge $M \subset \mathbb{R}^n$ heisst m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ der Klasse C^r, wobei $r \in \mathbb{N} \cup \{\infty\}$, falls gilt: zu jedem $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und einen C^r-Diffeomorphismus $\phi : \Omega \to \phi(\Omega)$ mit*

$$\phi(M \cap \Omega) = (\mathbb{R}^m \times \{0\}) \cap \phi(\Omega).$$

Wir nennen den Diffeomorphismus ϕ eine (lokale) Plättung von M. Im Einzelfall kann der Nachweis, dass eine gegebene Menge $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit ist, anhand der Definition mühevoll sein. Für Mengen, die als Niveaumengen einer Funktion gegeben sind, liefert jedoch der Satz über implizite Funktionen folgendes Kriterium.

Satz 2.2 (Untermannigfaltigkeitskriterien) *Sei $M \subset \mathbb{R}^n$ und $m + k = n$. Dann sind folgende Aussagen äquivalent:*

(1) *M ist eine m-dimensionale Untermannigfaltigkeit der Klasse C^r.*

(2) Niveaumengenkriterium: *Zu $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und eine Funktion $f \in C^r(\Omega,\mathbb{R}^k)$, so dass $M \cap \Omega = f^{-1}(0)$ und $\operatorname{rang} Df = k$ auf Ω.*

(3) Graphenkriterium: *Zu $p \in M$ gibt es eine offene Umgebung $U \times V \subset \mathbb{R}^m \times \mathbb{R}^k$ und $g \in C^r(U,V)$, so dass nach geeigneter Permutation der Koordinaten gilt:*

$$M \cap (U \times V) = \{(x,g(x)) : x \in U\}.$$

Beweis: Wir zeigen (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1). Es gelte (1), das heißt zu jedem $p \in M$ gibt es eine lokale Plättung $\phi : \Omega \to \phi(\Omega)$ der Klasse C^r mit $p \in \Omega$. Sei $\pi^\perp : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$ die Projektion auf den zweiten Faktor. Dann folgt mit $f = \pi^\perp \circ \phi$ für $q \in \Omega$

$$f(q) = 0 \quad \Leftrightarrow \quad \phi(q) \in \mathbb{R}^m \times \{0\} \quad \Leftrightarrow \quad q \in \phi^{-1}(\mathbb{R}^m \times \{0\}) = M \cap \Omega.$$

Außerdem gilt $\operatorname{rang} Df = \operatorname{rang} \big(\pi^\perp \circ D\phi\big) = k$ und $f \in C^r(\Omega, \mathbb{R}^k)$.

Ist (2) erfüllt, so ist nach evtl. Permutation der Koordinaten $D_y f(p)$ invertierbar, wobei $(x, y) \in \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^k$, und (3) folgt aus dem Satz über implizite Funktionen.

Für (3) $\Rightarrow$ (1) können wir annehmen, dass die Graphendarstellung ohne Permuation der Koordinaten gilt. Wir setzen $\Omega = U \times V$ und

$$\phi : \Omega \to \phi(\Omega),\ \phi(x, y) = \big(x, y - g(x)\big).$$

Dann ist $\phi \in C^r(\Omega, \mathbb{R}^n)$ injektiv und es gilt

$$D\phi(x, y) = \begin{pmatrix} E_m & 0 \\ -Dg(x) & E_k \end{pmatrix} \quad \Rightarrow \quad \det D\phi(x, y) = 1 \quad \text{für alle } (x, y) \in \Omega.$$

Also ist $\phi : \Omega \to \phi(\Omega)$ ein Diffeomorphismus der Klasse C^r nach dem Umkehrsatz. Da $(x, y) \in M \cap \Omega$ genau wenn $y = g(x)$, also $\phi(x, y) \in \mathbb{R}^k \times \{0\}$, ist die in (1) verlangte lokale Plättung gefunden. $\square$

Beispiel 2.5 Die Sphäre $\mathbb{S}^m = \{x \in \mathbb{R}^{m+1} : |x| = 1\}$ ist eine m-dimensionale Untermannigfaltigkeit im $\mathbb{R}^{m+1}$ der Klasse C^∞. Denn es gilt

$$\mathbb{S}^m = f^{-1}(0) \ \text{ für } \ f : \mathbb{R}^{m+1}\backslash\{0\} \to \mathbb{R},\ f(x) = |x|^2 - 1,$$

und $Df(x) \neq 0$ für alle $x \in \mathbb{R}^{m+1}\backslash\{0\}$. Die Behauptung folgt also aus Satz 2.2.

Definition 2.2 *Ein Vektor $v \in \mathbb{R}^n$ heisst Tangentialvektor von $M \subset \mathbb{R}^n$ m Punkt $p \in M$, falls es eine Abbildung $\gamma : (-\varepsilon, \varepsilon) \to M$ gibt mit $\gamma(0) = p$, $\gamma'(0) = v$. Die Menge der Tangentialvektoren von M im Punkt p wird mit T_pM bezeichnet.*

Folgerung 2.1 *Sei $M \subset \mathbb{R}^n$ eine m-dimensionale C^1-Untermannigfaltigkeit, und $n = m+k$. Ist $p \in M \cap \Omega = f^{-1}(0)$ für eine Funktion $f \in C^1(\Omega, \mathbb{R}^k)$ mit $\operatorname{rang} Df = k$ auf Ω, so gilt*

$$T_pM = \ker Df(p).$$

Insbesondere ist T_pM ein m-dimensionaler Unterraum des $\mathbb{R}^n$.

Beweis: Für $\gamma : (-\varepsilon, \varepsilon) \to M$ mit $\gamma(0) = p$ und $\gamma'(0) = w$ gilt

$$0 = \frac{d}{dt} f\big(\gamma(t)\big)|_{t=0} = Df\big(\gamma(0)\big)\gamma'(0) = Df(p)\, w \quad \Rightarrow \quad T_pM \subset \ker Df(p).$$

Nach Satz 2.2 gibt es andererseits, nach eventueller Permutation der Koordinaten, offene Mengen $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^k$ mit $p \in U \times V$, sowie $g \in C^1(U, V)$ mit

$$M \cap (U \times V) = \{(x, g(x)) : x \in U\}.$$

Die Graphenabbildung $G \in C^1(U, \mathbb{R}^n)$, $G(x) = (x, g(x))$, bildet nach M ab. Schreiben wir $p = (x_0, g(x_0))$ für geeignetes $x_0 \in U$, so folgt für alle $v \in \mathbb{R}^m$

$$DG(x_0)v = \frac{d}{dt} G(x_0 + tv)|_{t=0} \in T_pM \quad \Rightarrow \quad \mathrm{Bild}\, DG(x_0) \subset T_pM.$$

Zusammenfassend ist $\mathrm{Bild}\, DG(x_0) \subset T_pM \subset \ker Df(p)$. Aber $DG(x_0)$ ist injektiv, denn $DG(x_0)v = (v, Dg(x_0)v)$, und wegen $\mathrm{rang}\, Df(p) = k$ folgt mit der Dimensionsformel

$$\dim \mathrm{Bild}\, DG(x_0) = m = n - \mathrm{rang}\, Df(p) = \dim \ker Df(p).$$

Also gilt $\mathrm{Bild}\, DG(x_0) = \ker Df(p) = T_pM$. □

Die Folgerung zeigt, dass der Tangentialraum einer m-dimensionalen Untermannigfaltigkeit tatsächlich ein m-dimensionaler Vektorraum ist. Dies zeigt, dass die Dimension einer C^1-Untermannigfaltigkeit wohldefiniert ist, das heißt es kann nicht Plättungen mit verschiedenen Dimensionen von M geben. Auf die Idee wäre vermutlich auch kaum jemand gekommen. Wir kommen nun zur sogenannten *Lagrange-Multiplikatorenregel*.

Folgerung 2.2 (Extrema mit Nebenbedingungen) *Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega, \mathbb{R}^k)$ und $\varphi \in C^1(\Omega)$. Gilt dann für ein $p \in f^{-1}(0)$*

(1) *$\varphi(q) \geq \varphi(p)$ für alle $q \in \Omega$ mit $f(q) = 0$,*

(2) $\mathrm{rang}\, Df(p) = k$,

so gibt es $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ mit $\mathrm{grad}\, \varphi(p) = \sum_{i=1}^k \lambda_i \,\mathrm{grad}\, f_i(p)$.

Beweis: Nach eventueller Verkleinerung von Ω ist $\mathrm{rang}\, Df = k$ auf Ω und $M := f^{-1}(0)$ ist eine m-dimensionale Untermannigfaltigkeit, wobei $m = n - k$. Ist $\gamma : (-\varepsilon, \varepsilon) \to M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$, so hat hat $\varphi \circ \gamma$ in $t = 0$ ein lokales Minimum und folglich

$$0 = \frac{d}{dt} \varphi\big(\gamma(t)\big)|_{t=0} = \langle \mathrm{grad}\, \varphi(p), v \rangle,$$

das heißt $\mathrm{grad}\, \varphi(p) \in (T_pM)^\perp$. Da $f_j|_M \equiv 0$, gilt analog $\mathrm{grad}\, f_j(p) \in (T_pM)^\perp$ für $1 \leq j \leq k$. Aber $\dim (T_pM)^\perp = k$ nach Folgerung 2.1, und die Vektoren $\mathrm{grad}\, f_j(p)$ sind die Zeilenvektoren der Matrix $Df(p)$ mit Rang k. Wegen der Gleichheit von Zeilenrang und Spaltenrang ist $\{\mathrm{grad}\, f_j(p) : 1 \leq j \leq k\}$ eine Basis von $(T_pM)^\perp$, und die Behauptung folgt. □

Beispiel 2.6 Für eine symmetrische Matrix $B \in \mathbb{R}^{n \times n}$ betrachten wir hier nochmals das Problem, die quadratische Form $\varphi(x) = \langle Bx, x \rangle$ zu minimieren unter der Nebenbedingung $f(x) = |x|^2 - 1 = 0$. Da $\mathbb{S}^{n-1} = f^{-1}(0)$ kompakt und φ stetig, wird das Infimum in einem Punkt $x_0 \in \mathbb{S}^{n-1}$ angenommen. Nach Folgerung 2.2 gibt es dann ein $\lambda \in \mathbb{R}$ mit $\mathrm{grad}\, \varphi(x_0) = \lambda \,\mathrm{grad}\, f(x_0)$, also $Bx_0 = \lambda x_0$. Somit hat jede symmetrische Matrix B mindestens einen Eigenvektor, vergleiche Kapitel 7, Satz 2.3.

Wir haben die Sätze über inverse und implizite Funktionen im Endlichdimensionalen formuliert, damit das Wesentliche nicht durch zuviel Abstraktion verdeckt wird. An der Verallgemeinerung auf Abbildungen zwischen Banachräumen besteht aber großes Interesse: in den Anwendungen ist die Gleichung $f(x) = y$ zum Beispiel eine Differential- oder Integralgleichung, die durch eine gesuchte Funktion x in einem geeigneten Funktionenraum X gelöst werden soll. Eine Inspektion des Beweises des Umkehrsatzes ergibt, dass die Konstruktion der inversen Abbildung einschließlich ihrer Differenzierbarkeit ohne wesentliche Änderungen auch dann richtig ist, wenn f eine offene Teilmenge des Banachraums X in den Banachraum Y abbildet. Dabei muss die euklidische Norm von A durch die Operatornorm $\|A\| = \sup_{\|x\|=1} \|Ax\|$ ersetzt werden, und der Raum $L(\mathbb{R}^n, \mathbb{R}^m)$ muss ersetzt werden durch

$$L(X,Y) = \{A : X \to Y | \, A \text{ linear }, \, \|A\| < \infty\}.$$

Es ist leicht zu sehen, dass die Bedingung $\|A\| < \infty$ äquivalent dazu ist, dass die lineare Abbildung A stetig ist. Dies ist eine *conditio sine qua non*, darum nehmen wir sie in die Definition von $L(X,Y)$ auf. Insbesondere wird in der Definition der Differenzierbarkeit verlangt, dass $Df(x_0) \in L(X,Y)$.

Expliziten Gebrauch von den Koordinaten des $\mathbb{R}^n$ haben wir nur gemacht, um die höhere Differenzierbarkeit der Inversen g zu etablieren. Im Unendlichdimenensionalen wird dazu alternativ die sogenannte Neumannsche Reihe benutzt.