

# Kapitel 11

## Nichtlineare Analysis

### 11.1 Der Banachsche Fixpunktsatz

In diesem Kapitel bieten wir eine kurze Einführung in die nichtlineare Analysis. Der bei weitem wichtigste Satz der nichtlinearen Analysis ist der sogenannte Banachsche Fixpunktsatz. Wir werden gleich mehrere Anwendungen kennenlernen.

**Banachscher Fixpunktsatz 11.1.** *Sei  $f : X \rightarrow X$  eine Lipschitzstetige Abbildung eines vollständigen metrischen Raumes  $X$  auf sich selber mit Lipschitzkonstante  $L < 1$ . Dann hat  $f$  genau einen Fixpunkt:  $x \in X$  mit  $f(x) = x$ . Für jedes  $x_0 \in X$  konvergiert die induktiv durch  $x_{n+1} = f(x_n)$  definierte Folge  $(x_n)_{n \in \mathbb{N}_0}$  gegen den Fixpunkt.*

**Beweis:** Sei  $x_0 \in X$  und  $(x_n)_{n \in \mathbb{N}_0}$  induktiv definiert durch  $x_{n+1} = f(x_n)$ . Dann gilt

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq Ld(x_{n-1}, x_n) \leq \dots \leq L^n d(x_0, x_1).$$

Mit der Dreiecksungleichung folgt dann für  $n < m$

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + \dots + d(x_{m-1}, x_m) \\ &\leq (L^n + \dots + L^{m-1})d(x_0, x_1) \\ &= L^n \frac{1 - L^{m-n}}{1 - L} d(x_0, x_1) \leq \frac{L^n}{1 - L} d(x_0, x_1) \end{aligned}$$

Also ist  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge und es existiert  $x = \lim_{n \rightarrow \infty} x_n$ . Aus der Stetigkeit von  $f$  folgt dann  $f(x) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x$ . Also ist  $x$  ein Fixpunkt. Ist  $y \in X$  ein zweiter Fixpunkt, so gilt  $d(x, y) = d(f(x), f(y)) \leq L \cdot d(x, y)$ . Es folgt  $(1-L)d(x, y) \leq 0$  und wegen  $L < 1$  auch  $0 \leq d(x, y) \leq 0$ . Also gilt  $x = y$ . **q.e.d.**

Eine Anwendung ist z.B. der Satz von Picard Lindelöf über die Existenz und Eindeutigkeit von Anfangswertproblemen von gewöhnlichen Differentialgleichungen. Eine gewöhnliche Differentialgleichung ist eine Gleichung, von der Form

$$\dot{u}(t) = f(t, u(t)) \quad \text{mit } u : I \rightarrow X \quad \text{und } f : I \times U \rightarrow X.$$

Hierbei ist  $I \subset \mathbb{R}$  ein offenes Intervall,  $X$  ein normierter Vektorraum und  $U \subset X$  eine offene Teilmenge. Der Punkt bezeichnet die Ableitung nach  $t$ . Diese Variable steht in vielen Anwendungen für die Zeit. Das Anfangswertproblem besteht aus der Suche nach einer differenzierbaren Funktion  $u : I \rightarrow U$ , die die Differentialgleichung erfüllt und an einem Punkt  $t_0 \in I$  den Anfangswert  $u(t_0) = u_0 \in U$  annimmt.

**Definition 11.2.** Eine Funktion  $f$  von einem metrischen Raum  $X$  in den metrischen Raum  $Y$  heißt lokal Lipschitzstetig, wenn es für jedes  $x_0 \in X$  eine Umgebung  $U \subset X$  von  $x_0$  gibt und eine Lipschitzkonstante  $L \geq 0$ , so dass für alle  $x, x' \in U$  gilt

$$d(f(x), f(x')) \leq Ld(x, x').$$

**Satz 11.3.** (Lokale Existenz und Eindeutigkeit) Sei  $I$  ein offenes Intervall,  $U \subset \mathbb{R}^n$  eine offene Teilmenge und  $f : I \times U \rightarrow \mathbb{R}^n$  eine stetige Abbildung, die bezüglich der zweiten Variablen lokal Lipschitzstetig ist, d.h. für jedes  $(t_0, u_0) \in I \times U$  gibt es ein  $\delta > 0$  und ein  $L > 0$ , so dass für alle  $(t, u), (t, \tilde{u}) \in (t_0 - \delta, t_0 + \delta) \times B(u_0, \delta)$  gilt

$$\|f(t, u) - f(t, \tilde{u})\| \leq L\|u - \tilde{u}\|.$$

Dann gibt es für jedes  $(t_0, u_0) \in I \times U$  ein  $\epsilon > 0$ , so dass das Anfangswertproblem  $\dot{u}(t) = f(t, u(t))$  mit  $u(t_0) = u_0$  auf  $(t_0 - \epsilon, t_0 + \epsilon)$  genau eine Lösung besitzt.

**Beweis:** Wegen der lokalen Lipschitzstetigkeit gibt es  $\delta > 0$  und  $L \geq 0$ , so dass  $\|f(t, u) - f(t, \tilde{u})\| \leq L\|u - \tilde{u}\|$  für alle  $(t, u), (t, \tilde{u}) \in [t_0 - \delta, t_0 + \delta] \times \overline{B(u_0, \delta)} \subset I \times U$  gilt. Wegen der Stetigkeit von  $f$  definiert die Abbildungsvorschrift

$$F : u \mapsto F(u) \text{ mit } F(u)(t) = u_0 + \int_{t_0}^t f(s, u(s)) ds$$

eine stetige Abbildung von  $C((t_0 - \delta, t_0 + \delta), \overline{B(u_0, \delta)})$  nach  $C((t_0 - \delta, t_0 + \delta), \mathbb{R}^n)$ . Sei

$$\|f(\cdot, u_0)\|_\infty = \sup\{\|f(s, u_0)\| \mid s \in [t_0 - \delta, t_0 + \delta]\}.$$

Für  $\epsilon \leq \delta$  und  $\epsilon(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$  gilt dann

$$\|F(u)(t) - u_0\| = \left\| \int_{t_0}^t (f(s, u_0) + f(s, u(s)) - f(s, u_0)) ds \right\| \leq \epsilon(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$$

für alle  $u \in C((t_0 - \epsilon, t_0 + \epsilon), \overline{B(u_0, \delta)})$  und alle  $t \in (t_0 - \epsilon, t_0 + \epsilon)$ . Also bildet  $F$  den vollständigen metrischen Raum  $C((t_0 - \epsilon, t_0 + \epsilon), \overline{B(u_0, \delta)})$  auf sich selber ab. Für  $u, \tilde{u} \in C((t_0 - \epsilon, t_0 + \epsilon), \overline{B(u_0, \delta)})$  und  $t \in (t_0 - \epsilon, t_0 + \epsilon)$  gilt dann

$$\|F(u)(t) - F(\tilde{u})(t)\| \leq \int_{t_0}^t \|f(s, u(s)) - f(s, \tilde{u}(s))\| ds \leq \int_{t_0}^t L\|u - \tilde{u}\| ds \leq \epsilon L\|u - \tilde{u}\|_\infty.$$

Für  $\|f(\cdot, u_0)\|_\infty > 0$  sei  $\epsilon = \min\left\{\delta, \frac{\delta}{\|f(\cdot, u_0)\|_\infty + L\delta}\right\}$  mit  $\epsilon L \leq \frac{L\delta}{\|f(\cdot, u_0)\|_\infty + L\delta} < 1$ , für  $\|f(\cdot, u_0)\|_\infty = 0$  und  $\delta L < 1$  sei  $\epsilon = \delta$ , und für  $\|f(\cdot, u_0)\|_\infty = 0$  und  $\delta L \geq 1$  sei  $\epsilon < \frac{1}{L} \leq \delta$ . Dann bildet  $F$  den vollständigen metrischen Raum  $C((t_0 - \epsilon, t_0 + \epsilon), \overline{B}(u_0, \delta))$  lipschitzstetig mit Lipschitzkonstante  $\epsilon \cdot L < 1$  auf sich selber ab. Jeder Fixpunkt von  $F$  ist wegen dem Hauptsatz der Differential- und Integralrechnung auf  $t \in (t_0 - \epsilon, t_0 + \epsilon)$  stetig differenzierbar und löst das Anfangswertproblem  $\dot{u}(t) = f(t, u)$  mit  $u(t_0) = u_0$ .

Für jede Lösung  $u$  dieses Anfangswertproblems auf  $(t_0 - \epsilon, t_0 + \epsilon)$  verschwindet umgekehrt die Differenz  $F(u) - u$  bei  $t = t_0$  und ihre Ableitung auf  $(t_0 - \epsilon, t_0 + \epsilon)$ . Also stimmen  $F(u)$  und  $u$  überein und jede Lösung des obigen Anfangswertproblems ist ein Fixpunkt von  $F$ . Dann folgt die Existenz und Eindeutigkeit dieses Anfangswertproblems auf  $(t_0 - \epsilon, t_0 + \epsilon)$  aus dem Banachschen Fixpunktsatz. **q.e.d.**

**Satz 11.4\*** (Globale Existenz und Eindeutigkeit) Sei  $O \subset \mathbb{R} \times \mathbb{R}^n$  eine offene Teilmenge und  $f : O \rightarrow \mathbb{R}^n$  eine stetige Abbildung, die wie bei der lokalen Existenz und Eindeutigkeit lokal lipschitzstetig ist. Dann gibt es für jedes  $(t_0, u_0) \in O$  genau ein maximales Intervall  $(a, b) \subset \mathbb{R}$ , das  $t_0$  enthält, und auf dem das Anfangswertproblem

$$\dot{u}(t) = f(t, u) \quad \text{mit} \quad u(t_0) = u_0$$

genau eine Lösung  $u$  enthält. Das Intervall ist in dem Sinne maximal, dass an beiden Rändern, also bei  $a$  und  $b$ , jeweils eine der folgenden Bedingungen erfüllt ist:

- (i)  $a = -\infty$  bzw.  $b = \infty$ .
- (ii)  $t \mapsto \|f(t, u(t))\|$  ist für alle  $\epsilon > 0$  auf  $(a, a + \epsilon)$  bzw.  $(b - \epsilon, b)$  unbeschränkt.
- (iii) Die Lösung  $u$  läßt sich stetig auf  $[a, b]$  bzw.  $(a, b]$  fortsetzen, der Graph der Fortsetzung liegt aber nicht in  $O$ , d.h.  $\lim_{t \downarrow a} (t, u(t)) \notin O$  bzw.  $\lim_{t \uparrow b} (t, u(t)) \notin O$ .

**Beweis\*:** Wir betrachten zuerst zwei Lösungen  $u_1 : (a_1, b_1) \rightarrow \mathbb{R}^n$  und  $u_2 : (a_2, b_2) \rightarrow \mathbb{R}^n$  von  $\dot{u}(t) = f(t, u(t))$  mit  $u_1(t_1) = u_2(t_1)$  für ein  $t_1 \in (a_1, b_1) \cap (a_2, b_2)$ . Die Menge

$$A = \{t \in (a_1, b_1) \cap (a_2, b_2) \mid u_1(t) = u_2(t)\}$$

ist in  $(a_1, b_1) \cap (a_2, b_2)$  wegen der Stetigkeit von  $u_1 - u_2$  abgeschlossen und wegen der Eindeutigkeit in Satz (11.3) offen. Wenn es ein  $b \in (a_1, t_1] \cap (a_2, t_1] \setminus A$  gäbe, dann wäre  $a = \inf(A \cap [b, t_1]) \in A$ , weil  $A$  abgeschlossen ist, also  $a > b$  und  $[b, a) \cap A = \emptyset$ , im Widerspruch zu der Offenheit von  $A$ . Es gilt sogar  $A = (a_1, b_1) \cap (a_2, b_2)$ , weil  $d = \sup(A \cap [t_1, c]) \in A$  und  $(d, c] \cap A = \emptyset$  für  $c \in [t_1, b_1) \cap [t_1, b_2) \setminus A$  gelten müsste. Insbesondere definieren alle Lösungen  $u : I \ni t_0 \rightarrow \mathbb{R}^n$  des Anfangswertproblems

$$\dot{u}(t) = f(t, u) \quad \text{mit} \quad u(t_0) = u_0$$

auf der Vereinigung ihrer Definitionsbereiche auch eine Lösung. Wir zeigen, dass diese Vereinigung der Definitionsbereiche an beiden Rändern die Bedingung (iii) erfüllt, wenn

jeweils (i) und (ii) nicht gelten. Dann ist die Ableitung der Lösung auf einer Menge  $(a, a + \epsilon)$  bzw.  $(b - \epsilon, b)$  beschränkt und die Lösung ist dort lipschitzstetig. Für jede Folge  $(t_n)_{n \in \mathbb{N}}$ , die gegen  $a$  bzw.  $b$  konvergiert, konvergiert  $((t_n, u(t_n))_{n \in \mathbb{N}}$  in  $\mathbb{R} \times \mathbb{R}^n$ . Läge der Grenzwert in  $O$ , dann sei  $\tilde{u}$  eine Lösung nach Satz (11.3) des Anfangswertproblems

$$\dot{\tilde{u}}(t) = f(t, \tilde{u}(t)) \quad \text{mit} \quad \tilde{u}(a) = \lim_{t \rightarrow a^+} u(t) \quad \text{bzw.} \quad \tilde{u}(b) = \lim_{t \rightarrow b^-} u(t)$$

auf einer Umgebung von  $a$  bzw.  $b$ . Dann wäre  $u(t)$  für  $t \in (a, b)$  zusammen mit  $\tilde{u}(t)$  für  $t \notin (a, b)$  auf einem größeren Intervall als  $(a, b)$  stetig und eine Lösung, weil es für  $t \neq a, b$  eine sich stetig nach  $a$  bzw.  $b$  fortsetzende Ableitung hätte. Also gilt (iii). **q.e.d.**

**Bemerkung 11.5\*:** Im Fall (ii) setzt sich  $t \mapsto f(t, u(t))$  nicht stetig nach  $a$  bzw.  $b$  fort. Also entweder setzt sich  $u$  nicht stetig nach  $a$  bzw.  $b$  fort, oder  $f$  nicht nach  $(a, u(a))$  bzw.  $(b, u(b))$ . Im Fall (iii) können wir eventuell  $f$  und die Lösung stetig fortsetzen.

## 11.2 Das Lösen von nichtlinearen Gleichungen

Die Lösungen der Gleichungen von der Form

$$Ax = y, \quad A \in \mathcal{L}(V, W), \quad x \in V \quad \text{und} \quad y \in W$$

in einem (endlichdimensionalen) Vektorraum  $V$  sind in der linearen Algebra untersucht worden. Wenn  $A$  invertierbar ist, dann ist  $x = A^{-1}y$  die eindeutige Lösung. In diesem Abschnitt nutzen wir das Verständnis dieser Gleichungen für Gleichungen von der Form

$$f(x) = y, \quad f : V \rightarrow W, \quad x \in V \quad \text{und} \quad y \in W$$

mit nichtlinearen Abbildungen  $f$ . Dabei nehmen wir an, dass  $f$  differenzierbar ist, und durch lineare Abbildungen angenähert werden kann. Ausgangspunkt ist die Beobachtung, dass kleine Störungen von invertierbaren linearen Abbildungen invertierbar sind.

**Lemma 11.6.** Seien  $V$  und  $W$  Banachräume und  $A$  ein invertierbares Element von  $\mathcal{L}(V, W)$ . D.h. es gibt ein Element  $A^{-1} \in \mathcal{L}(W, V)$  mit  $AA^{-1} = \mathbb{1}_W$  und  $A^{-1}A = \mathbb{1}_V$ . Dann sind alle Elemente des folgenden Balles um  $A$  invertierbar:

$$B \in B\left(A, \frac{1}{\|A^{-1}\|}\right) \subset \mathcal{L}(V, W) \quad \text{mit} \quad \|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2 \|A - B\|}{1 - \|A^{-1}\| \|A - B\|}$$

**Beweis:** Offenbar ist  $B = A - (A - B) = A(\mathbb{1}_V - A^{-1}(A - B))$ . Wegen Satz 9.61 gilt  $\|A^{-1}(A - B)\| \leq \|A^{-1}\| \cdot \|A - B\| < 1$  für  $B \in B\left(A, \frac{1}{\|A^{-1}\|}\right)$ . Dann folgt aus der Neumannschen Reihe, dass  $\mathbb{1}_V - A^{-1}(A - B)$  invertierbar ist in  $\mathcal{L}(V)$  und der inverse

Operator beschränkt ist durch  $\frac{1}{1-\|A^{-1}\|\|A-B\|}$ . Also ist auch  $B$  invertierbar und es gilt

$$B^{-1} = (\mathbb{1}_V - A^{-1}(A - B))^{-1} A^{-1} \quad \text{mit} \quad \|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|A - B\|},$$

$$B^{-1} - A^{-1} = ((\mathbb{1}_V - A^{-1}(A - B))^{-1} - \mathbb{1}_V)A^{-1} = A^{-1}(A - B)(\mathbb{1}_V - A^{-1}(A - B))^{-1}A^{-1}$$

mit  $\|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|A - B\|}{1 - \|A^{-1}\| \cdot \|A - B\|}$ . **q.e.d.**

Damit bilden die invertierbaren Elemente von  $\mathcal{L}(V, W)$  eine offene Teilmenge.

**Korollar 11.7.** Für Banachräume  $V$  und  $W$  und invertierbare  $A \in \mathcal{L}(V, W)$  ist

$$B \left( A, \frac{1}{\|A^{-1}\|} \right) \rightarrow \mathcal{L}(W, V), \quad B \mapsto B^{-1}$$

eine analytische Abbildung, also insbesondere unendlich oft stetig differenzierbar.

**Beweis:** Aus Lemma 11.6 und der Neumannschen Reihe folgt für alle  $B \in B(0, \frac{1}{\|A^{-1}\|})$

$$\|(A + B)^{-1} - A^{-1} + A^{-1}BA^{-1}\| \leq \left\| A^{-1} \sum_{n=2}^{\infty} (-BA^{-1})^n \right\| \leq \frac{\|B\|^2\|A^{-1}\|^3}{1 - \|B\|\|A^{-1}\|}.$$

Insbesondere ist  $A \mapsto A^{-1}$  differenzierbar mit der Ableitung  $A \mapsto (B \mapsto -A^{-1}BA^{-1})$ , und damit einmal mehr differenzierbar als  $A \mapsto A^{-1}$ , also unendlich oft. Für obige  $B$  und alle  $t \in \mathbb{K}$  mit  $|t| \leq 1$  gilt  $(A + tB)^{-1} = \sum_{n \in \mathbb{N}_0} A^{-1}(-BA^{-1})^n t^n$ . **q.e.d.**

**Lemma 11.8.** Seien  $X$  und  $Y$  normierte Vektorräume und  $f : U \rightarrow V$  eine bijektive Abbildung zwischen den offenen Teilmengen  $U \subset X$  und  $V \subset Y$ . Wenn  $f$  bei  $x_0 \in U$  differenzierbar,  $f'(x_0)$  in  $\mathcal{L}(X, Y)$  invertierbar, und  $f^{-1}$  bei  $y_0 = f(x_0)$  stetig ist, dann ist  $f^{-1}$  bei  $y_0$  differenzierbar mit  $(f^{-1})'(y_0) = (f'(x_0))^{-1} = (f'(f^{-1}(y_0)))^{-1}$ .

**Beweis:** Sei  $x_0 \in U$  wie im Lemma,  $0 < \epsilon \leq 1$  und  $\delta > 0$  so gewählt, dass

$$\|f(x) - f(x_0) - f'(x_0)(x - x_0)\| \leq \frac{\epsilon}{2} \|(f'(x_0))^{-1}\|^{-1} \cdot \|x - x_0\| \quad \text{für alle } x \in B(x_0, \delta) \subset U$$

gilt. Es folgt  $\|f'(x_0)(x - x_0)\| \leq \|f(x) - f(x_0)\| + \frac{\epsilon}{2} \|(f'(x_0))^{-1}\|^{-1} \cdot \|x - x_0\|$  und damit

$$\|(f'(x_0))^{-1}\| \cdot \|f(x) - f(x_0)\| \geq \|(f'(x_0))^{-1}\| \cdot \|f'(x_0)(x_0 - x)\| - \frac{\epsilon}{2} \|x - x_0\| \geq \frac{2-\epsilon}{2} \cdot \|x - x_0\|.$$

Bei  $y_0$  ist  $f^{-1}$  stetig, und  $f[B(x_0, \delta)]$  eine Umgebung von  $f(x_0)$ . Für  $x \in B(x_0, \delta)$  gilt

$$\begin{aligned} \|x - x_0 - (f'(x_0))^{-1}(f(x) - f(x_0))\| &\leq \|(f'(x_0))^{-1}\| \cdot \|f'(x_0)(x - x_0) - f(x) + f(x_0)\| \\ &\leq \frac{\epsilon}{2} \|x - x_0\| \leq \frac{\epsilon}{2-\epsilon} \|(f'(x_0))^{-1}\| \cdot \|f(x) - f(x_0)\|. \end{aligned}$$

Wegen  $\frac{\epsilon}{2-\epsilon} \leq \epsilon$  ist  $f^{-1}$  bei  $f(x_0)$  differenzierbar mit  $(f^{-1})'(y_0) = (f'(x_0))^{-1}$ . **q.e.d.**

**Satz 11.9.** (Satz über die inverse Funktion) Seien  $X, Y$  Banachräume,  $f : U \rightarrow Y$  eine differenzierbare Abbildung von einer offenen Teilmenge  $U \subset X$  nach  $Y$ . Wenn  $f'$  bei  $x_0 \in U$  invertierbar und stetig ist, dann gibt es offene Umgebungen  $V \subset U$  und  $W \subset Y$  von  $x_0$  bzw.  $f(x_0)$ , so dass die Einschränkung  $f|_V : V \rightarrow W$  bijektiv ist mit differenzierbarer Umkehrabbildung  $(f|_V)^{-1} : W \rightarrow V$  mit Ableitung  $y \mapsto (f'((f|_V)^{-1}(y)))^{-1}$ .

Dabei ist  $(f|_V)^{-1}$  auf  $V$  genauso oft (stetig) differenzierbar wie  $f$ .

**Beweis:** Zuerst ersetzen wir  $f$  durch  $h \circ f \circ g$  mit folgenden Abbildungen

$$g : X \rightarrow X, \quad x \mapsto x + x_0, \quad h : Y \rightarrow X, \quad y \mapsto (f'(x_0))^{-1}(y - f(x_0)).$$

Die Abbildungen  $g$  und  $h$  sind bijektiv und glatt und haben glatte Umkehrabbildungen. Dadurch wird  $Y = X$ ,  $x_0 = 0 = f(x_0)$ , und  $f'(x_0) = \mathbb{1}_X$ . Weil  $f'$  stetig bei 0 ist, gibt es ein  $\delta > 0$ , so dass  $\|f'(x) - \mathbb{1}_X\| \leq \frac{1}{2}$  für  $x \in \overline{B(0, \delta)} \subset U$  gilt. Wegen Lemma 11.6 sind diese  $f'(x)$  invertierbar. Wegen dem Schrankensatz ist für jedes  $y \in X$  die Abbildung

$$F_y : \overline{B(0, \delta)} \rightarrow \overline{B(y, \frac{\delta}{2})}, \quad x \mapsto y + x - f(x)$$

lipschitzstetige mit Lipschitzkonstante  $\frac{1}{2}$ . Dann gilt auch für alle  $x \in \overline{B(0, \delta)}$

$$\|F_y(x) - y\| = \|F_y(x) - F_y(0)\| = \frac{1}{2}\|x - 0\| \leq \frac{\delta}{2}.$$

Wenn  $y$  in  $\overline{B(0, \frac{\delta}{2})}$  liegt, dann liegt  $\overline{B(y, \frac{\delta}{2})}$  in  $\overline{B(0, \delta)}$ . Also definiert  $F_y$  eine Abbildung von  $\overline{B(0, \delta)}$  auf sich selbst. Wegen dem Banachschen Fixpunktsatz hat  $F_y$  auf  $\overline{B(0, \delta)}$  für jedes  $y \in \overline{B(0, \frac{\delta}{2})}$  genau einen Fixpunkt in  $\overline{B(y, \frac{\delta}{2})}$ . Weil  $x$  genau dann ein Fixpunkt von  $F_y$  ist, wenn  $f(x) = y$  ist, gibt es für alle  $y \in \overline{B(0, \frac{\delta}{2})}$  auf  $\overline{B(0, \delta)}$  genau eine Lösung von  $f(x) = y$ . Sei also  $W = \overline{B(0, \frac{\delta}{2})}$  und  $V = f^{-1}[W] \cap \overline{B(0, \delta)}$ . Dann sind  $W$ , sein Urbild  $f^{-1}[W]$  unter der stetigen Abbildung  $f$  und  $V$  offen und  $f|_V : V \rightarrow W$  bijektiv. Weil  $F_0$  auf  $\overline{B(0, \delta)}$  lipschitzstetig ist mit Lipschitzkonstante  $\frac{1}{2}$ , gilt für alle  $x, x' \in \overline{B(0, \delta)}$

$$\begin{aligned} \|x - x'\| &= \|F_0(x) - F_0(x') + f(x) - f(x')\| \leq \frac{1}{2}\|x - x'\| + \|f(x) - f(x')\|, \\ &\text{oder auch} \quad \|x - x'\| \leq 2\|f(x) - f(x')\|. \end{aligned}$$

Also ist  $(f|_V)^{-1}$  lipschitzstetig und mit Lemma 11.8 differenzierbar, weil  $(f'(x))^{-1}$  für  $x \in V$  existieren. Die letzte Aussage folgt aus Korollar 11.7 und Satz 10.4 (iii). **q.e.d.**

**Beispiel 11.10.** Die Voraussetzung der stetigen Differenzierbarkeit kann nicht abgeschwächt werden zu einfacher Differenzierbarkeit. Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto f(x)$

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin(\frac{1}{x}) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0, \end{cases} \quad f'(x) = \begin{cases} \frac{1}{2} + 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}) & \text{für } x \neq 0, \\ \frac{1}{2} & \text{für } x = 0. \end{cases}$$

ist differenzierbar mit  $f'(0) = \frac{1}{2}$ . Aber  $f$  ist in keiner Umgebung der 0 injektiv.

**Korollar 11.11.** (Satz über die implizite Funktion) Seien  $X$  und  $Y$  Banachräume,  $U$  eine offene Teilmenge von  $X \times Y$  und  $f : U \rightarrow Y$  differenzierbar. Wenn  $f'$  in  $(x_0, y_0) \in U$  stetig und  $\frac{\partial f}{\partial y}(x_0, y_0)$  in  $\mathcal{L}(Y)$  invertierbar ist, dann gibt es offene Umgebungen  $V$  von  $x_0$  in  $X$ ,  $W$  von  $f(x_0, y_0)$  in  $Y$  und  $O$  von  $(x_0, y_0)$  in  $U$  und eine differenzierbare Funktion  $g : V \times W \rightarrow Y$ , so dass  $f^{-1}[\{z\}] \cap O = \text{Graph}(g(\cdot, z))$  für alle  $z \in W$  gilt:

$$\{(x, y) \in O \mid f(x, y) = z\} = \{(x, g(x, z)) \mid x \in V\} \quad \text{für alle } z \in W.$$

**Beweis:** Die Ableitung der Abbildung  $F : U \rightarrow X \times Y$ ,  $(x, y) \mapsto (x, f(x, y))$  ist

$$F'(x, y) : X \times Y \rightarrow X \times Y, \quad (v, w) \mapsto \left( v, \frac{\partial f(x, y)}{\partial x} v + \frac{\partial f(x, y)}{\partial y} w \right)$$

Wenn  $\frac{\partial f(x, y)}{\partial y}$  in  $\mathcal{L}(Y)$  invertierbar ist, dann ist der inverse Operator gegeben durch

$$(F'(x, y))^{-1} : X \times Y \rightarrow X \times Y, \quad (v, w) \mapsto \left( v, \left( \frac{\partial f(x, y)}{\partial y} \right)^{-1} \left( w - \frac{\partial f(x, y)}{\partial x} v \right) \right)$$

Wegen Satz 10.11 erfüllt  $F$  in  $(x_0, y_0)$  die Voraussetzungen des Satzes über die inverse Funktion. Deshalb gibt es Umgebungen  $V$  von  $x_0$  in  $X$ ,  $W$  von  $f(x_0, y_0)$  in  $W$  und  $O$  von  $(x_0, y_0)$  in  $U$ , so dass die Abbildung  $O \rightarrow V \times W$ ,  $(x, y) \mapsto (x, f(x, y))$  bijektiv ist und eine Umkehrabbildung besitzt. Diese Umkehrabbildung muss aber wegen der Gestalt von  $F$  von der Form  $V \times W \rightarrow O$ ,  $(x, y) \mapsto (x, g(x, y))$  sein, mit einer differenzierbaren Funktion  $g : V \times W \rightarrow Y$ . Insbesondere sind für alle  $z \in W$  alle Lösungen  $(x, y) \in O$  von  $f(x, y) = z$  im Graphen von  $g : V \rightarrow Y$ ,  $x \mapsto g(x, z)$  enthalten. **q.e.d.**

**Definition 11.12** (Untermannigfaltigkeit). Eine Teilmenge  $A \subset \mathbb{R}^n$  heißt *Untermannigfaltigkeit*, wenn bei jedem  $x \in A$  eine stetig differenzierbare Funktion  $f : U \rightarrow \mathbb{R}^m$  auf einer in  $\mathbb{R}^n$  offenen Menge existiert, so dass  $f'(x)$  surjektiv und  $f^{-1}[\{0\}] = A \cap U$  ist.

Nach einer geeigneten Permutation der Koordinaten erfüllt  $f$  bei  $x \in A$  als Funktion auf  $U \subset \mathbb{R}^{n-m} \times \mathbb{R}^m$  die Voraussetzungen des Satzes der impliziten Funktion. Also gilt lokal  $A = \text{Graph}(g)$  für eine Funktion  $g : V \rightarrow \mathbb{R}^m$  auf einer offenen Menge  $V \subset \mathbb{R}^{n-m}$ .

**Beispiel 11.13. (i) Höhenlinien:** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion, die z.B. in Abhängigkeit von Längen- und Breitengraden die Höhe über dem Meeresspiegel beschreibt. Wenn die partielle Ableitung  $\frac{\partial f}{\partial y}$  in einem Punkt  $(x_0, y_0) \in \mathbb{R}$  nicht verschwindet, dann gibt es eine stetig differenzierbare Funktion

$$g : (x_0 - \epsilon, x_0 + \epsilon) \times (f(x_0, y_0) - \delta, f(x_0, y_0) + \delta) \rightarrow \mathbb{R},$$

so dass für  $z \in (f(x_0, y_0) - \delta, f(x_0, y_0) + \delta)$  die Höhenlinien zur Höhe  $z$  der Graph

$$\{x, g(x, z) \mid x \in (x_0 - \epsilon, x_0 + \epsilon)\} \quad \text{von} \quad (x_0 - \epsilon, x_0 + \epsilon) \rightarrow \mathbb{R}, \quad x \mapsto g(x, z)$$

ist. Für festes  $z$  zeigt also der Vektor  $(1, \frac{\partial g}{\partial x}(x, z))$  in Richtung der Höhenlinie durch  $(x, g(x, z))$  und steht senkrecht auf dem Gradienten von  $f$  bei  $(x, g(x, z))$ .

**(ii) Hyperflächen:** Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion, deren Ableitung in einem Punkt  $x_0$  nicht verschwindet. Dann verschwindet auch mindestens eine partielle Ableitung nicht. Nach einer geeigneten Permutation der Koordinaten, können wir  $\frac{\partial f}{\partial x_n}(x_0) \neq 0$  annehmen. Sei  $y_0 \in \mathbb{R}^{n-1}$  der Vektor der ersten  $n-1$  Koordinaten von  $x_0$ . Dann lassen sich für  $z \in (f(x_0) - \delta, f(x_0) + \delta)$  lokal die Niveaumengen  $f^{-1}[\{z\}]$  durch den Graphen einer Funktion  $g : B(y_0, \epsilon) \times (f(x_0) - \delta, f(x_0) + \delta) \rightarrow \mathbb{R}$  beschreiben als  $\{(y, g(y, z)) \mid y \in B(y_0, \epsilon)\}$ , also durch  $y \in B(y_0, \epsilon) \subset \mathbb{R}^{n-1}$  parametrisiert. Für alle  $(y, z)$  in dieser Umgebung von  $(y_0, f(x_0))$  ist das Bild der partiellen Ableitungen

$$\mathbb{1}_{\mathbb{R}^n} \times \frac{\partial g(y, z)}{\partial y} \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^{n-1}) \times \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}) \simeq \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^{n-1} \times \mathbb{R})$$

der Kern von  $f'(y, g(y, z)) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ , der Tangentialraum an die Niveauflächen heißt.

**Definition 11.14.** Eine unendlich oft (stetig) differenzierbare bijektive Abbildung mit unendlich oft (stetig) differenzierbarer Umkehrabbildung heißt Diffeomorphismus.

**Beispiel 11.15.** Polarkoordinaten Die Abbildung

$$(0, \infty) \times \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}, \quad (r, \varphi) \mapsto (r \cos \varphi, \sin \varphi)$$

heißt Polarkoordinaten von  $\mathbb{R}^2$ . Offenbar ist diese Abbildung unendlich oft stetig differenzierbar. Die Umkehrabbildung ist dann gegeben durch

$$(x, y) \mapsto (r, \varphi) \quad \text{mit} \quad r = \sqrt{x^2 + y^2} \quad \text{und} \\ \varphi = \begin{cases} \arccos\left(\frac{x}{r}\right) & \text{für } y \geq 0 \\ -\arccos\left(\frac{x}{r}\right) & \text{für } y < 0 \end{cases} \quad \text{mod } 2\pi \quad \begin{cases} \arcsin\left(\frac{y}{r}\right) & \text{für } x \geq 0 \\ \pi - \arcsin\left(\frac{y}{r}\right) & \text{für } x < 0. \end{cases}$$

Also ist diese Abbildung ein Diffeomorphismus von  $(0, \infty) \times \mathbb{R}/2\pi\mathbb{Z}$  nach  $\mathbb{R}^2 \setminus \{(0, 0)\}$ . Hier beschreibt  $\mathbb{R}/2\pi\mathbb{Z}$  den Raum aller Äquivalenzklassen von  $\mathbb{R}$ , wobei

$$x \sim y \Leftrightarrow x - y \in 2\pi\mathbb{Z}.$$

Dieser Raum ist offenbar lokal diffeomorph zu  $\mathbb{R}$ , weil in jedem Intervall dessen Länge kleiner ist als  $2\pi$ , verschiedene Elemente verschiedene Äquivalenzklassen repräsentieren. Deshalb sind die Einschränkungen der Abbildung  $\mathbb{R} \rightarrow \mathbb{R}/2\pi\mathbb{Z}$  auf beliebige offene Intervalle mit Längen nicht größer als  $2\pi$ , die jedes Element auf die entsprechende Äquivalenzklasse abbilden, Diffeomorphismen.

## 11.3 Lagrangemultiplikatoren

Ziel dieses Abschnittes ist es die lokalen Extremwerte von Funktionen auf solchen Teilmengen eines normierten Vektorraumes  $X$  zu bestimmen, die die Nullstellen von endlich vielen reellen differenzierbaren Funktionen bilden. Wir sprechen dann von Zwangsbedingungen, wegen denen nur die Punkte in diesen Nullstellenmengen in Betracht kommen. Diese Situation ist recht allgemein und kommt in vielen Anwendungen der Wirtschaftswissenschaften vor. Dieses Verfahren ist die Grundlage für die nichtlineare Optimierung, in der man nach Extremwerten auf Teilmengen eines Banachraumes sucht. Darauf aufbauend wird in der konvexen Analysis nach Bedingungen gesucht, die die Existenz und Eindeutigkeit von solchen Extremwertproblemen garantieren.

**Definition 11.16.** Sei  $U \subset X$  eine offene Teilmenge eines normierten Vektorraumes  $X$  und  $g = (g_1, \dots, g_m) : U \rightarrow \mathbb{R}^m$  eine stetig differenzierbare Funktion von  $U$  nach  $\mathbb{R}^m$ . Für jedes  $x_0 \in U$  definiert  $A = g^{-1}[\{g(x_0)\}] = \{x \in U \mid g(x) = g(x_0)\}$  eine abgeschlossene Teilmenge  $A$  von  $U$  auf der die reellen Funktionen  $g_1, \dots, g_m$  konstant sind. Wir nennen solche Mengen Niveaumengen zu den Zwangsbedingungen  $g_1, \dots, g_m$ .

**Definition 11.17.** Die Niveaumenge  $A$  heißt in einem Punkt  $x_0 \in A$  stetig differenzierbar (glatt), wenn es eine offene Umgebung  $U \subset X$  von  $x_0$  und einen stetig differenzierbaren (glatten) Homöomorphismus  $\Phi$  von einer offenen Teilmenge  $O$  des Kerns  $Y = \{x \in X \mid g'(x_0)(x) = 0\}$  nach  $A \cap U$  gibt, so dass  $\Phi'(\Phi^{-1}(x_0))$  in  $\mathcal{L}(Y)$  invertierbar ist. Ein solcher Punkt  $x_0$  heißt kritischer Punkt der Einschränkung  $f|_{(A \cap U)}$  einer differenzierbaren Funktion  $f : U \rightarrow \mathbb{R}$  auf die Niveaumenge, falls  $(f \circ \Phi)'(\Phi^{-1}(x_0)) = 0$ .

Aufgrund der Definition ist ein Punkt  $x_0 \in U$ , an dem die Niveaumenge  $A$  stetig differenzierbar ist, höchstens dann ein lokaler Extremwert der Einschränkung  $f|_A$  einer differenzierbaren Funktion  $f : U \rightarrow \mathbb{R}$ , wenn er ein kritischer Punkt im Sinne dieser Definition ist. Deshalb kommen als die lokalen Extremwerte von  $f|_A$  neben diesen kritischen Punkten nur solche Punkte in Betracht, an denen die Niveaumenge nicht stetig differenzierbar ist. Sie werden Singularitäten genannt. Im folgenden werden also einerseits diese kritischen Punkte und andererseits die Singularitäten charakterisiert.

**Lemma 11.18.** Sei  $X$  ein Banachraum,  $U \subset X$  offen und seien  $g : U \rightarrow \mathbb{R}^m$  und in  $x_0 \in U$  die Niveaumenge  $A = g^{-1}[\{g(x_0)\}]$  stetig differenzierbar. Für  $r = \dim g'(x_0)[X]$  gibt es  $1 \leq l_1 < l_2 < \dots < l_r \leq m$  und eine offene Umgebung  $V \subset U$  von  $x_0$  so dass die Ableitung von  $\tilde{g} = (g_{l_1}, \dots, g_{l_r})$  bei allen  $x \in V$  surjektiv ist und  $\tilde{g}^{-1}[\{\tilde{g}(x_0)\}] \cap V = A \cap V$  gilt. Insbesondere ist  $A \cap V$  eine Untermannigfaltigkeit von  $V$ .

**Beweis:** Wir wählen induktiv Indizes  $1 \leq l_1 < \dots < l_r$ , so dass  $\nabla g_{l_1}(x_0), \dots, \nabla g_{l_r}(x_0)$  linear unabhängig sind und  $\nabla g_1(x_0), \dots, \nabla g_m(x_0)$  aufspannen. Dann stimmt der Kern von  $\tilde{g}'(x_0)$  mit dem Kern  $Y$  von  $g'(x_0)$  überein, und  $\tilde{g}'(x_0)$  ist surjektiv und bildet

geeignete  $z_1, \dots, z_r \in X$  auf die Standardbasis von  $\mathbb{R}^r$  ab. Aufgrund der Wahl von  $z_1, \dots, z_r$  erfüllt für alle  $j = 1, \dots, r$  die erste Komponente der Abbildung

$$x \mapsto \left( x - \sum_{i=1}^r \tilde{g}'_i(x_0)(x)z_i, \sum_{i=1}^r \tilde{g}'_i(x_0)(x)z_i \right)$$

die Gleichung  $\tilde{g}'_j(x_0)(x - \sum_{i=1}^r \tilde{g}'_i(x_0)(x)z_i) = \tilde{g}'_j(x_0)(x) - \tilde{g}'_j(x_0)(x) = 0$ . Also liegt sie im Kern von  $\tilde{g}'_j(x_0)$ , der gleich  $Y$  ist. Dann ist diese Abbildung in  $\mathcal{L}(X, Y \times Z)$  die inverse von der Abbildung  $(y, z) \mapsto y + z$  in  $\mathcal{L}(Y \times Z, X)$ . Ihre Verkettung mit  $\tilde{g}$  erfüllt an dem Punkt  $(y_0, z_0) \in Y \times Z$ , der auf  $x_0$  abgebildet wird, die Voraussetzungen des Satzes der impliziten Funktion. Also gibt es eine Umgebung  $V \subset X$  von  $x_0$ , eine Umgebung  $W \subset Y \times Z$  von  $(y_0, z_0)$  und eine stetig differenzierbare Abbildung  $\Psi : W \rightarrow Z$ , so dass  $\tilde{g}^{-1}[\{\tilde{g}(x_0)\}] \cap V$  gleich dem Bild von  $W$  unter  $W \rightarrow X, y \mapsto y + \Psi(y)$  ist.

Weil  $A$  in  $x_0$  stetig differenzierbar ist, gibt es einen stetig differenzierbaren Homöomorphismus  $\Phi$  von einer offenen Teilmenge  $O$  von  $Y$  auf eine in  $A$  offene Umgebung  $\Phi[O]$  von  $x_0$ . Wegen  $A = g^{-1}[\{g(x_0)\}] \subset \tilde{g}^{-1}[\{\tilde{g}(x_0)\}]$  liegt  $\Phi[O] \cap V$  im Graphen von  $\Psi$ . Also ist  $\Phi$  auf  $\Phi^{-1}[\Phi[O] \cap V]$  gleich der Verkettung von  $\Phi$  mit der obigen ersten Komponente der inversen Abbildung von  $(y, z) \mapsto y + z$  und von  $y \mapsto y + \Psi(y)$ . Weil  $\Phi'(\Phi^{-1}(x_0))$  ein Isomorphismus nach  $Y$  ist, ist wegen dem Satz der inversen Funktion die stetig differenzierbare Verkettung von  $\Phi$  mit dieser ersten Komponente auf einer Umgebung von  $\Phi^{-1}(x_0)$  ein Homöomorphismus. Insbesondere bildet  $y \mapsto y + \Psi(y)$  eine Umgebung von  $y_0$  bijektiv auf eine Umgebung von  $x_0$  in  $A$  ab. Also können die offenen Umgebungen  $W$  von  $y_0$  und  $V$  von  $x_0$  so verkleinert werden, dass  $y \mapsto y + \Psi(y)$  eine Homöomorphismus von  $W$  auf die Untermannigfaltigkeit  $\tilde{g}^{-1}[\{\tilde{g}(x_0)\}] \cap V = A \cap V$  von  $V$  ist. Das zeigt die Aussage. **q.e.d.**

**Satz 11.19.** Sei  $f : U \rightarrow \mathbb{R}$  eine differenzierbare und  $g : U \rightarrow \mathbb{R}^m$  eine stetig differenzierbare Funktion auf einer offenen Teilmenge  $U$  eines normierten Vektorraumes  $X$  und die Niveaumenge  $A = g^{-1}[\{g(x_0)\}]$  bei  $x_0 \in U$  stetig differenzierbar. Dann ist  $x_0$  genau dann ein kritischer Punkt von der Einschränkung  $f|_A$  von  $f$  auf  $A$ , wenn es reelle Zahlen  $\lambda_1, \dots, \lambda_m$  gibt, so dass  $f'(x_0) = \lambda_1 g'_1(x_0) + \dots + \lambda_m g'_m(x_0)$  gilt.

**Definition 11.20.** Die Zahlen  $\lambda_1, \dots, \lambda_m$  heißen Lagrangemultiplikatoren.

**Beweis:** Ein Punkt  $x_0 \in U$  ist genau dann ein kritischer Punkt, wenn  $\Phi^{-1}(x_0)$  ein kritischer Punkt von  $f \circ \Phi : O \rightarrow \mathbb{R}$  ist. Hierbei ist  $\Phi : O \rightarrow U$  eine bijektive stetig differenzierbare Abbildung, und die Ableitung  $\Phi'(\Phi^{-1}(x_0))$  bei  $\Phi^{-1}(x_0)$  ist eine bijektive Abbildung auf den Kern von  $g'(x_0)$ . Das ist äquivalent dazu, dass  $f'(x_0)\Phi'(\Phi^{-1}(x_0))$  verschwindet, oder dazu, dass  $f'(x_0)$  auf dem Bild von  $\Phi'(\Phi^{-1}(x_0))$ , also auf dem Kern  $Y = \{x \in X \mid g'(x_0)(x) = 0\}$  von  $g'(x_0)$  verschwindet.

Sei  $R$  die Dimension vom Bild von  $g'(x_0)$  in  $\mathbb{R}^m$ . Für alle  $r = 1, \dots, R$  gibt es dann einen kleinsten Index  $l_r > l_{r-1}$ , so dass  $(g'_1(x_0), \dots, g'_{l_r}(x_0))$  den Rang  $r$  hat

und damit auch  $(g'_{l_1}(x_0), \dots, g'_{l_r}(x_0))$ . Dann ist  $\tilde{g}'(x_0) = (g'_{l_1}(x_0), \dots, g'_{l_r}(x_0))$  eine surjektive Abbildung nach  $\mathbb{R}^R$  deren Kern gleich  $Y$  ist. Sie bildet geeignet gewählte Vektoren  $z_1, \dots, z_R \in X$  auf die Standardbasis  $e_1, \dots, e_R$  von  $\mathbb{R}^R$  ab. Dann liegt  $x - \sum_{r=1}^R g'_{l_r}(x_0)(x)z_r$  für alle  $x \in X$  in  $Y$ . Wenn  $f'(x_0)$  auf  $Y$  verschwindet gilt also

$$f'(x_0)(x) = \sum_{r=1}^R \lambda_r g'_{l_r}(x_0)(x) \quad \text{für alle } x \in X \quad \text{mit} \quad \lambda_r = f'(x_0)(z_r).$$

Umgekehrt verschwindet  $\lambda_1 g'_{l_1}(x_0) + \dots + \lambda_m g'_{l_m}(x_0)$  für  $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$  auf  $Y$ . **q.e.d.**

**Lemma 11.21.** *Sei  $g : U \rightarrow \mathbb{R}^m$  eine stetig differenzierbare Abbildung auf einer offenen Menge  $U$  eines normierten Vektorraumes  $X$ . Dann ist für alle  $r \in \mathbb{N}_0$  die Menge  $\{x \in U \mid \dim(g'(x)[X]) \geq r\}$  entweder leer oder offen.*

**Beweis:** Sei  $x_0 \in \{x \in U \mid \dim(g'(x)[X]) \geq r\}$ . Dann gibt es Indizes  $1 \leq l_1 < \dots < l_r \leq m$ , so dass  $(g'_{l_1}(x_0), \dots, g'_{l_r}(x_0))$  eine surjektive Abbildung nach  $\mathbb{R}^r$  ist. Sie bildet geeignet gewählte Elemente  $z_1, \dots, z_r \in X$  auf die Standardbasis  $e_1, \dots, e_r$  von  $\mathbb{R}^r$  ab. Weil  $g$  stetig differenzierbar ist, sind die Funktionen  $x \mapsto g'_{l_i}(x)(z_j)$  auf  $U$  stetig. Deshalb ist die Teilmenge von  $\{x \in U \mid \dim(g'(x)[X]) \geq r\}$ , auf der die Determinante der  $r \times r$  Matrix  $(g'_{l_i}(x)(z_j))_{1 \leq i, j \leq r}$  nicht verschwindet, offen. **q.e.d.**

Typischerweise werden die Zwangsbedingungen glatte Funktionen sein. Aber selbst dann sind die Niveaumengen nicht immer glatt.

**Satz 11.22. (Rangatz)** *Für eine stetig differenzierbare Abbildung  $g : U \rightarrow \mathbb{R}^m$  auf einer offenen Menge  $U$  eines Banachraumes  $X$  ist die Niveaumenge  $A = g^{-1}[\{g(x_0)\}]$  in allen lokalen Maxima  $x_0$  von der Funktion  $x \mapsto \dim(g'(x)[X])$  stetig differenzierbar.*

**Beweis:** Wenn  $g'(x_0) \in \mathcal{L}(X, \mathbb{R}^m)$  surjektiv ist, dann gibt es Elemente  $z_1, \dots, z_m \in X$ , die durch  $g'(x_0)$  auf die Standardbasis  $e_1, \dots, e_m$  von  $\mathbb{R}^m$  abgebildet werden. Sei  $Z$  der von  $z_1, \dots, z_m$  aufgespannte Unterraum von  $X$ , und  $Y$  der Kern von  $g'(x_0)$ . Für alle  $x \in X$  gilt  $g'(x_0)(x - g'_{l_1}(x_0)(x)z_1 - \dots - g'_{l_m}(x_0)(x)z_m) = 0$ . Deshalb ist die Abbildung  $I$  mit  $x \mapsto (x - g'_{l_1}(x_0)(x)z_1 - \dots - g'_{l_m}(x_0)(x)z_m, g'_{l_1}(x_0)(x)z_1 + \dots + g'_{l_m}(x_0)(x)z_m)$  ein linearer stetiger Isomorphismus von  $X$  nach  $Y \times Z$  der,  $U$  auf eine offene Teilmenge von  $O \subset Y \times Z$  abbildet. Dann ist  $g \circ I^{-1}$  stetig differenzierbare Abbildung von  $O$  nach  $\mathbb{R}^m$ , deren partielle Ableitung nach der zweiten Koordinate  $Z$  isomorph auf  $\mathbb{R}^m$  abbildet. Die Aussage folgt aus dem Satz über die implizite Funktion.

Wenn die Dimension  $\dim(g'(x)[X])$  bei  $x_0$  gleich  $r$  und auf einer Umgebung nicht größer als  $r$  ist, dann gibt es Indizes  $1 \leq l_1 < l_2 < \dots < l_r \leq m$ , so dass die linearen Abbildungen  $g'_{l_1}(x_0), \dots, g'_{l_r}(x_0)$  der Komponenten von  $\tilde{g}(x) = (g_{l_1}(x), \dots, g_{l_r}(x))$  bei  $x = x_0$  linear unabhängig sind. Dann ist die Niveaumenge  $\tilde{A} = \tilde{g}^{-1}[\{\tilde{g}(x_0)\}]$  bei  $x_0$  stetig differenzierbar. Wegen Lemma 11.21 sind  $g'_{l_1}(x), \dots, g'_{l_r}(x)$  auf einer Umgebung von  $x_0$  linear unabhängig. Weil  $\dim(g'(x)[X])$  auf einer Umgebung von  $x_0$  nicht größer

als  $r$  ist, stimmen auf einer Umgebung von  $x_0$  die Kerne von  $g'(x)$  und  $\tilde{g}'(x)$  überein. Wegen Satz 11.19 verschwindet in dieser Umgebung  $g'(x)$  auf der Niveaumenge  $\tilde{A}$ , und diese Niveaumenge  $\tilde{A}$  stimmt in dieser Umgebung mit der von  $g$  überein. **q.e.d.**

Weil die Funktion  $\dim(g'(x)[X])$  nur die endlich vielen Werte  $0, \dots, m$  annehmen kann, gibt es in jeder offenen Menge  $U$  ein lokales Maximum. Die Menge aller solcher lokalen Maxima ist wegen Lemma 11.21 sogar offen und dicht in  $U$ . Mit Hilfe von dem Lemma 11.21 und dem Rangsatz können wir die Singularitäten der Niveaumengen von stetig differenzierbaren Funktionen  $g : U \rightarrow \mathbb{R}^m$  dadurch bestimmen, dass wir

- (i) wie im Beweis des Rangsatzes maximal viele Komponenten  $\tilde{g}$  von  $g$  auswählen, deren Ableitungen  $\tilde{g}'(x)$  an möglichst vielen Punkten surjektiv sind,
- (ii) und dann die Punkte bestimmen, an denen diese Ableitung nicht surjektiv sind. Das sind die Nullstellen der Determinante aus dem Beweis von Lemma 11.21.

Zum Abschluss formulieren wir folgende Umkehrung vom Lemma 11.18: Wenn es  $1 \leq l_1 < \dots < l_r \leq m$  für  $r = \dim g'(x_0)[X]$  gibt, so dass die Ableitung von  $\tilde{g} = (g_{l_1}, \dots, g_{l_r})$  bei  $x_0$  surjektiv ist, und  $\tilde{g}^{-1}[\{\tilde{g}(x_0)\}] \cap V = g^{-1}[\{g(x_0)\}] \cap V$  gilt, dann folgt aus dem Rangsatz und dem Satz der impliziten Funktion, dass die Niveaumenge  $g^{-1}[\{g(x_0)\}]$  in  $x_0$  stetig differenzierbar ist. Allerdings gilt nicht, dass eine Niveaumenge in allen den Punkten stetig differenzierbar ist, in denen Sie eine Untermannigfaltigkeit ist. So besteht z.B. die Niveaumenge  $g^{-1}[\{0\}]$  von  $g : \mathbb{R}^d \rightarrow \mathbb{R}, x \mapsto x^2$  nur aus dem Punkt  $x = 0$  und ist deshalb eine Untermannigfaltigkeit. Weil aber der Kern von  $g'(0)$  gleich  $\mathbb{R}^d$  ist, ist  $x = 0$  eine Singularität. Wir können das so zusammenfassen, dass die Frage, ob eine Niveaumenge in einem Punkt  $x_0$  eine Singularität hat oder nicht, nicht nur von der Niveaumenge durch  $x_0$  abhängt, sondern auch von der entsprechenden Familie von Niveaumengen durch die Punkte in einer Umgebung von  $x_0$ . So stimmt z.B. die Niveaumenge  $g^{-1}[\{0\}]$  mit der Niveaumenge  $(g^2)^{-1}[\{0\}]$  überein. Wegen  $\nabla(g^2) = 2g\nabla g$ , besteht die zweite Niveaumenge nur aus Singularitäten, während die erste natürlich auch gar keine Singularitäten haben kann. Andererseits stimmen für  $g'(x_0) \neq 0$  lokal bei  $x = x_0$  die Niveaumengen  $g^{-1}[\{g(x)\}]$  und  $(g^2)^{-1}[\{g^2(x)\}]$  genau dann überein, wenn  $g(x_0) \neq 0$  gilt. Deshalb hängt die Frage, ob eine Niveaumenge an einer Stelle  $x_0$  eine Singularität hat oder nicht, davon ab, wie die Niveaumengen in einer Umgebung von  $x_0$  aussehen.

**Beispiel 11.23. (i)**  $g : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto g(x, y) = x^2 + y^2.$

Der Gradient  $\nabla g(x, y) = (2x, 2y)$  von  $g$  verschwindet nur bei  $(x, y) = 0$ . Also sind alle Niveaumengen  $g(x, y) = g_0$  mit  $g_0 \neq g(0, 0) = 0$  glatte 1-dimensionale Teilmengen von  $\mathbb{R}^2$ . Es sind jeweils die Kreise mit Radius  $\sqrt{g_0}$  um den Nullpunkt. Für  $g_0 = 0$  besteht die Niveaumenge nur aus  $\{0\}$ . Der Nullpunkt ist eine Singularität der Niveaumenge.

**(ii)**  $g : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto g(x, y) = x^2 - y^2.$

Der Gradient  $\nabla g(x, y) = (2x, -2y)$  verschwindet wieder nur bei  $(x, y) = (0, 0)$ . Also sind alle Niveaumengen  $g(x, y) = g_0$  mit  $g_0 \neq g(0, 0) = 0$  glatte eindimensionale

Teilmengen von  $\mathbb{R}^2$ . Es sind jeweils zwei Hyperebenen. Die Niveaumenge  $g(x, y) = x^2 - y^2 = (x - y)(x + y) = 0$  besteht aber aus zwei Geraden  $y = x$  und  $y = -x$ , die sich im Nullpunkt schneiden. Diese Niveaumenge hat also im Nullpunkt eine Singularität, weil sich dort zwei glatte Teilmengen schneiden. Man spricht von einem Doppelpunkt.

(iii)  $g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto g(x, y) = y^2 - x^3.$

Der Gradient  $\nabla g(x, y) = (-3x^2, 2y)$  verschwindet wieder nur im Nullpunkt. Also sind wieder alle Niveaumengen  $g(x, y) = g_0$  mit  $g_0 \neq g(0, 0) = 0$  glatte eindimensionale Teilmengen von  $\mathbb{R}^2$ . Die Niveaumenge  $g(x, y) = y^2 - x^3 = 0$  besteht aus zwei Lösungen  $y = \pm\sqrt{x^3}$  mit  $x \geq 0$ , die sich bei  $(x, y) = 0$  einer gemeinsamen Halbgeraden parallel zu der  $x$ -Achse annähern. Man nennt deshalb die Singularität im Nullpunkt eine Spitze.

**Satz 11.24\*** (Whitney) *Jede nicht leere abgeschlossene Teilmenge  $A \subset \mathbb{R}^n$  ist die Nullstellenmenge einer glatten Funktion auf  $\mathbb{R}^n$ .*

**Beweis\***: Sei  $A \subset \mathbb{R}^n$  eine nicht leere abgeschlossene Menge. Dann besitzt  $\mathbb{Q}^n \cap O$  mit  $O = \mathbb{R}^n \setminus A$  eine Abzählung durch eine Folge  $(x_n)_{n \in \mathbb{N}}$ . Für jedes  $n \in \mathbb{N}$  ist  $B(x_n, r_n) = \bigcup_{\{r > 0 | B(x, r) \subset O\}} B(x, r)$  eine Teilmenge von  $O$ . Also ist die Vereinigung aller Bälle  $(B(x_n, r_n))_{n \in \mathbb{N}}$  eine offene Teilmenge von  $O$ . Jeder Punkt  $x \in O$  ist in einem Ball  $B(x, r) \subset O$  enthalten, und in  $B(x, \frac{r}{2})$  ist ein  $x_n$  enthalten mit  $r_n > \frac{r}{2}$ . Daraus folgt  $x \in B(x_n, r_n)$  und die Vereinigung der Bälle  $(B(x_n, r_n))_{n \in \mathbb{N}}$  ist  $O$ . Die Funktion

$$\psi : \mathbb{R}^n \rightarrow [0, \infty), \quad x \mapsto \begin{cases} \exp(\frac{1}{x^2-1}) & \text{für } \|x\| < 1 \\ 0 & \text{für } \|x\| > 1 \end{cases}$$

ist unendlich oft stetig differenzierbar, und für alle  $n \in \mathbb{N}_0$  sind alle partiellen Ableitungen höchstens  $n$ -ter Ordnung beschränkt durch ein  $C_n > 0$ . Alle partiellen Ableitungen höchstens  $n$ -ter Ordnung von  $\psi_n(x) = \frac{1}{C_n} (\frac{\min\{1, r_n\}}{2})^n \psi(\frac{x-x_n}{r_n})$  sind beschränkt durch  $2^{-n}$ . Für jedes Monom  $D$  in  $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$  folgt aus Satz 8.38 induktiv im Grad von  $D$ , dass die Summe der partiellen Ableitungen  $(\sum D\psi_n)_{n \in \mathbb{N}}$  auf ganz  $\mathbb{R}^n$  gleichmäßig gegen die entsprechende partielle Ableitung  $Df$  von  $f = \sum_{n \in \mathbb{N}} \psi_n$  konvergiert. Deshalb ist  $f$  glatt und die Nullstellenmenge von  $f$  ist gleich  $A$ . **q.e.d.**

**Beispiel 11.25. (i)** *Die sogenannte Cantormenge ist definiert als das Komplement  $A = [0, 1] \setminus I$  folgender offenen Teilmenge von  $[0, 1]$ :*

$$I = \bigcup_{n \in \mathbb{N}_0} \bigcup_{(z_1, \dots, z_n) \in \{0, 1\}^n} \left( \frac{1}{3^{n+1}} + \sum_{l=1}^n \frac{2z_l}{3^l}, \frac{2}{3^{n+1}} + \sum_{l=1}^n \frac{2z_l}{3^l} \right) = (\frac{1}{3}, \frac{2}{3}) \cup (\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9}) \dots$$

Sie ist offenbar eine abgeschlossene Teilmenge von  $[0, 1]$  und wegen dem Satz von Whitney die Niveaumenge einer glatten reellen Funktion  $f$  auf  $\mathbb{R}$ . Weil in jeder Umgebung von jedem Punkt von  $A$  sowohl Elemente von  $A$  als auch Elemente von  $I$  enthalten sind, verschwinden alle Ableitungen von  $f$  auf  $A$ . Alle Punkte von  $A$  sind Singularitäten.

(ii) Der Sierpinski Teppich ist definiert als das Komplement  $A = [0, 1]^2 \setminus I$  folgender offenen Teilmenge von  $[0, 1]^2$ :

$$I = \bigcup_{n \in \mathbb{N}_0} \left( \bigcup_{(z_1, \dots, z_n) \in \{0, 1, 2\}^n} \left( \frac{1}{3^{n+1}} + \sum_{l=1}^n \frac{z_l}{3^l}, \frac{2}{3^{n+1}} + \sum_{l=1}^n \frac{z_l}{3^l} \right) \right)^2$$

$$= \left( \frac{1}{3}, \frac{2}{3} \right)^2 \cup \left( \left( \frac{1}{9}, \frac{2}{9} \right) \cup \left( \frac{4}{9}, \frac{5}{9} \right) \cup \left( \frac{7}{9}, \frac{8}{9} \right) \right)^2 \cup \dots$$

Die Menge  $I$  ist eine Teilmenge von dem kartesischen Produkt des Komplementes der Cantormenge in (i) mit sich selber. Wenn also  $x$  oder  $y$  zu der Cantormenge gehören, dann sind  $\{x\} \times [0, 1]$  und  $[0, 1] \times \{y\}$  Teilmengen von  $A$ . Deshalb ist  $A$  zusammenhängend. Wegen dem Satz von Whitney ist  $A$  die Nullstellenmenge einer glatten Funktion of  $\mathbb{R}^2$ . Wieder enthält jede Umgebung von jedem Punkt von  $A$  sowohl Elemente von  $A$  als auch Elemente von  $I$ , so dass alle Ableitungen von  $f$  auf  $A$  verschwinden. Alle Punkte von  $A$  sind Singularitäten.