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 $f : X \to X$ eine Lipschitz–stetige Abbildung eines vollständigen metrischen Raumes auf sich selber mit Lipschitzkonstante $L < 1$. Dann hat f genau einen Fixpunkt: $x \in X$ mit $f(x) = x$.*

Beweis: Sei $x_0 \in X$ beliebig und für alle $n \in \mathbb{N}$ x_n induktiv definiert durch $x_n = f(x_{n-1})$. Dann gilt

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

Mit der Dreiecksungleichung folgt dann für $n \leq m$

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

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

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

Hierbei ist $I \subset \mathbb{R}$ ein Intervall, V ein normierter Vektorraum und $U \subset V$ eine offene Teilmenge. Der Punkt bezeichnet die Ableitung nach t, die in vielen Anwendungen für die Zeit steht. Das Anfangswertproblem besteht aus der Suche nach einer differenzierbaren Funktion $u : I \to U$, 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 Lipschitz–stetig, wenn es für jedes $x_0 \in X$ eine Umgebung $U \subset X$ von x_0 gibt und eine Lipschitzkonstante $L > 0$, 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 $f : I \times U \to \mathbb{R}^n$ 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 $L > 0$, so dass für alle $(t, u), (t, \tilde{u}) \in [t_0 - \delta, t_0 + \delta] \times \overline{B(u_0, \delta)}$ auch $\|f(t, u) - f(t, \tilde{u})\| \leq L\|u - \tilde{u}\|$ gilt. Wegen der Stetigkeit von f ist die Abbildung

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

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

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

Wenn $\epsilon \leq \delta$ und $\epsilon\,(\|f(\cdot, u_0)\|_\infty + L\delta) \leq \delta$, dann gilt für alle $u \in C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ und alle $t \in [t_0 - \epsilon, t_0 + \epsilon]$

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

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

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

Sei also ϵ kleiner als $\quad \epsilon < \min\left\{\delta, \frac{\delta}{\|f(\cdot,u_0)\|_\infty + L\delta}, \frac{1}{L}\right\} = \min\left\{\delta, \frac{\delta}{\|f(\cdot,u_0)\|_\infty + L\delta}\right\}.$

Dann definiert die Abbildung F eine Lipschitzstetige Abbildung mit Lipschitzkonstante $\epsilon \cdot L < 1$ von dem vollständigen metrischen Raum $C([t_0 - \epsilon, t_0 + \epsilon], \overline{B(u_0, \delta)})$ auf sich selber. Jeder Fixpunkt ist wegen dem Hauptsatz der Differential– und Integralrechnung stetig differenzierbar und es gilt $\dot{u}(t) = f(t, u)$ für alle $t \in (t_0 - \epsilon, t_0 + \epsilon)$ mit $u(t_0) = u_0$. Also löst u dieses Anfangswertproblem auf $(t_0 - \epsilon, t_0 + \epsilon)$. Wenn u umgekehrt auf $(t_0 - \epsilon, t_0 + \epsilon)$ dieses Anfangswertproblem löst, dann ist die Ableitung von $F(u) - u$ gleich Null, und beide Funktionen $F(u)$ und u sind bei $t = t_0$ gleich u_0. Also stimmen beide Funktionen überein und jede Lösung des obigen Anfangswertproblems ist ein Fixpunkt von F. Also 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 $f : O \to \mathbb{R}^n$ eine stetige Abbildung, die wie bei der lokalen Existenz und Eindeutigkeit lokal Lipschitz–stetig 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*

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

genau eine Lösung u enthält. Das Intervall ist in dem Sinne maximal, dass an beiden Rändern, also bei a und b, 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: Zunächst bemerken wir, dass für jedes Intervall (a, b), das t_0 enthält, und auf dem das Anfangswertproblem

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

eine Lösung $\tilde{u}$ besitzt, so dass sich $\tilde{u}$ auf $[a, b)$ oder $(a, b]$ stetig fortsetzen läßt, und der Graph der Fortsetzung in O liegt, das neue Anfangswertproblem

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

wegen dem vorangehenden Satz eine Lösung in einer Umgebung von a bzw. b besitzt, die dann auf $[a, a + \epsilon)$ bzw. $(b - \epsilon, b]$ mit $\tilde{u}$ übereinstimmt. Also existiert ein maximales Intervall (a, b), auf dem das Anfangswertproblem eine eindeutige Lösung besitzt. Wenn am linken bzw. rechten Rand die Bedingungen (i) und (ii) nicht erfüllt sind, dann ist die Ableitung der Lösung auf einer offenen Menge $(a, a+\epsilon)$ bzw. $(b-\epsilon, b)$ beschränkt und deshalb ist die Lösung dort Lipschitz–stetig. Dann konvergiert für jede Folge $(t_n)_{n\in\mathbb{N}}$, die gegen a bzw. b konvergiert auch die Folge $((t_n, u(t_n))_{n\in\mathbb{N}}$ in $\mathbb{R}\times\mathbb{R}^n$. Der Grenzwert kann dann aber nicht in O liegen, weil sonst die Lösung eine Fortsetzung auf eine Umgebung von a bzw. b hätte. **q.e.d.**

Bemerkung 11.5. *Wenn (ii) erfüllt ist, kann $t\mapsto f(t,u(t))$ nicht stetig auf $[a, a+\epsilon)$ bzw. $(b-\epsilon, b]$ fortgesetzt werden. Also können u und f nicht so stetig auf größere Definitionsbereiche fortgesetzt werden, dass a (bzw. b) im Definitionsbereich von u und $(a, u(a))$ (bzw. $(b, u(b))$) im Definitionsbereich von f liegt.*

11.2 Das Lösen von nichtlinearen Gleichungen

Die Lösungen der Gleichungen von der Form

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

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

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

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 9.60 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 $\frac{1}{1-\|A^{-1}\|\cdot\|A-B\|}$. Also ist auch B invertierbar und es gilt

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

Für die Differenz $B^{-1} - A^{-1}$ gilt dann

$$\begin{aligned} B^{-1} - A^{-1} &= ((\mathbb{1}_V - A^{-1}(A-B))^{-1} - \mathbb{1})A^{-1} \\ &= A^{-1}(A-B)(\mathbb{1}_V - A^{-1}(A-B))^{-1}A^{-1} \\ &= (\mathbb{1}_V - A^{-1}(A-B))^{-1}A^{-1}(A-B)A^{-1} \quad \text{und deshalb} \\ \|B^{-1} - A^{-1}\| &\le \frac{\|A^{-1}\|^2\|A-B\|}{1 - \|A^{-1}\| \cdot \|A-B\|}. \end{aligned}$$

q.e.d.

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

Korollar 11.7. *Seien V und W Banachräume und A ein invertierbares Element von $\mathcal{L}(V,W)$. Dann ist die Abbildung*

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

eine analytische Abbildung. Also insbesondere unendlich oft stetig differenzierbar.

Beweis: Aus Lemma 11.6 und der Neumannschen Reihe folgt für alle $B \in \mathcal{L}(V,W)$

$$(A+tB)^{-1} = A^{-1}\sum_{n=0}^{\infty} t^n(-BA^{-1})^n = \left(\sum_{n=0}^{\infty} t^n(-A^{-1}B)^n\right)A^{-1} \text{ für } t \in B\left(0, \frac{\|B\|}{\|A^{-1}\|}\right).$$

Insbesondere ist diese Abbildung analytisch mit der Ableitung $B \mapsto -A^{-1}BA^{-1}$ an der Stelle A, und damit genauso oft differenzierbar, wie $A \mapsto A^{-1}$. **q.e.d.**

Satz 11.8. *(Satz über die inverse Funktion) Seien V, W Banachräume, $f : U \to W$ eine stetig differenzierbare Abbildung von einer offenen Teilmenge $U \subset V$ nach W. Wenn $f'(x_0)$ bei $x_0 \in U$ invertierbar ist in $\mathcal{L}(V,W)$, dann gibt es offene Umgebungen $U' \subset U$ und $O \subset W$ von x_0 bzw. $f(x_0)$, so dass die Einschränkung $f : U' \to O$ bijektiv ist mit stetig differenzierbarer Umkehrabbildung $f^{-1} : O \to U'$ mit Ableitung $y \to (f'(f^{-1}(y)))^{-1}$.*

Beweis Indem wir zu der Abbildung $x \to (f'(x_0))^{-1} \circ (f(x_0+x) - f(x_0))$ übergehen, können wir annehmen, dass $W = V$ ist und x_0 und $f(x_0)$ gleich Null sind und $f'(x_0) = \mathbb{1}_V$ ist. Weil f stetig differenzierbar ist, gibt es ein $\delta > 0$, so dass $\|f'(x) - \mathbb{1}_V\| \le \frac{1}{2}$ für $x \in \overline{B(0,\delta)} \subset U$ gilt. Wegen dem Schrankensatz ist für jedes $y \in V$ die Abbildung

$$F_y : x \mapsto y + x - f(x)$$

eine Lipschitzstetige Abbildung von $\overline{B(0,\delta)}$ nach $\overline{B(y, \frac{\delta}{2})}$ mit Lipschitzkonstante $\frac{1}{2}$.Außerdem gilt auch wegen dem Schrankensatz für alle $x \in \overline{B(0,\delta)}$

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

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 dann eine Abbildung von $\overline{B(0,\delta)}$ auf sich selbst. Aus dem Banachschen Fixpunktsatz folgt, dass für jedes $y \in \overline{B(y,\frac{\delta}{2})}$ die Abbildung F_y auf $\overline{B(0,\delta)}$ genau einen Fixpunkt hat und der Fixpunkt in $\overline{B(y,\frac{\delta}{2})}$ liegt. Weil aber x genau dann ein Fixpunkt von F_y ist, wenn $f(x) = y$ ist, gibt es für alle $y \in \overline{B(y,\frac{\delta}{2})}$ auf $\overline{B(0,\delta)}$ genau eine Lösung der Gleichung $f(x) = y$. Sei also

$$O = B\left(0, \frac{\delta}{2}\right) \qquad \text{und} \qquad U' = \left\{x \in B\left(0,\delta\right) | f(x) \in B\left(0, \frac{\delta}{2}\right)\right\}.$$

Dann ist O offen und U' als Urbild einer offenen Menge unter einer stetigen Abbildung auch offen und die Abbildung $f : U' \to O$ bijektiv. Weil aber die Abbildung F_0 auf $\overline{B(0,\delta)}$ Lipschitzstetig ist mit Lipschitzkonstante $\frac{1}{2}$, gilt für alle $x, x' \in \overline{B(0,\delta)}$ auch

$$\|x - x'\| = \|f(x) - f(x') - F_0(x) - F_0(x'))\| \leq \|f(x) - f(x')\| + \frac{1}{2}\|x - x'\|$$
$$\text{oder auch} \qquad \|x - x'\| \leq 2\|f(x) - f(x')\|.$$

Also ist f^{-1} sogar Lipschitz-stetig mit Lipschitzkonstante 2. Wegen der Neumann'schen Reihe ist für alle $x \in \overline{B(0,\delta)}$ dann $f'(x)$ in $\mathcal{L}(V)$ invertierbar, und wegen dem Lemma 11.6 die Abbildung

$$\overline{B(0,\delta)} \to \mathcal{L}(V), \quad x \mapsto (f'(x))^{-1}$$

stetig. Die Komposition von $f^{-1} : O \to U'$ mit dieser Abbildung ist dann auch stetig. Für $x, x_0 \in O$ folgt aus $\|f(x) - f(x_0) - f'(x_0)(x - x_0)\| < \epsilon\|x - x_0\|$

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

Also ist die Komposition von $f^{-1} : O \to U'$ mit $x \mapsto (f'(x))^{-1}$ die Ableitung von f^{-1}. Dann ist also f^{-1} auch stetig differenzierbar. **q.e.d.**

Beispiel 11.9. *Die Voraussetzung der stetigen Differenzierbarkeit kann nicht abgeschwächt werden zu einfacher Differenzierbarkeit. Die Funktion*

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) = \begin{cases} x + x^2 \sin(\frac{1}{x}) & \textit{für } x \neq 0 \\ 0 & \textit{für } x = 0 \end{cases}$$

ist differenzierbar und bei 0 *gilt* $f'(0) = 1$. *Aber* f *ist in keiner Umgebung der* 0 *injektiv.*

Korollar 11.10. *Die Umkehrabbildung einer bijektiven n-mal stetig differenzierbaren Abbildung ist bei allen Punkten x mit invertierbarer Ableitung $f'(x)$ auch n-mal stetig differenzierbar.* **q.e.d.**

Korollar 11.11. *(Satz über die implizite Funktion) Seien V und W Banachräume, U eine offene Teilmenge von $V \times W$ und $f : U \to V$ eine stetig differenzierbare Funktion. Wenn die partielle Ableitung $\frac{\partial f}{\partial v}$ in $(v_0, w_0) \in U$ als Element von $\mathcal{L}(V)$ invertierbar ist, dann gibt es offene Umgebungen O von $f(v_0, w_0)$ in V, O' von w_0 in W und U' von (v_0, w_0) in U und eine stetig differenzierbare Funktion $g : O \times O' \to V$, so dass für alle $(u, w) \in O \times O'$ gilt $f(g(u, w), w) = u$. Außerdem sind für alle $u \in O$ alle Lösungen $(v, w) \in U'$ der Gleichungen $f(v, w) = u$ im Graphen der Abbildung $O' \to V, \quad w \mapsto g(u, w)$ enthalten.*

Beweis: Die Ableitung der Abbildung $F : U \to V \times W, \quad (v, w) \mapsto (f(v, w), w)$ ist gegeben durch

$$F'(v, w) : V \times W \to V \times W, \qquad (x, y) \mapsto \left(\frac{\partial f(v, w)}{\partial v} x + \frac{\partial f(v, w)}{\partial w} y, y \right)$$

Wenn $\frac{\partial f(v,w)}{\partial v}$ invertierbar ist, dann ist der inverse Operator gegeben durch

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

Also erfüllt sie die Voraussetzungen des Satzes über die inverse Funktion. Deshalb gibt es Umgebungen O von $f(v_0, w_0)$ in V, O' von w_0 in W und U' von (v_0, w_0) in U, so dass die Abbildung $U \to O \times O', \quad (v, w) \mapsto (f(f(v, w), w)$ bijektiv ist und eine Umkehrabbildung besitzt. Diese Umfehrabbildung muss aber wegen der Gestalt von F von der Form $O \times O' \to U', \quad (u, w) \mapsto (g(u, w), w)$ sein, mit einer stetig differenzierbaren Funktion g. Insbesondere sind für alle $u \in O$ alle Lösungen $(v, w) \in U'$ von $f(v, w) = u$ im Graphen von $O \to V, \quad w \mapsto g(u, w)$ enthalten. **q.e.d.**

Beispiel 11.12. *(i) Höhenlinien: Sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine stetig differenzierbare Funktion, die z.B. in Abhängigkeit von Längen- und Breitengraden die Höhe über dem Meeresspiegel beschreibt. Wenn die partielle Ableitung $\frac{\partial f}{\partial x}$ (oder eine andere partielle Ableitung) in einem Punkt $(x_0, y_0) \in \mathbb{R}$ nicht verschwindet, dann gibt es eine stetig differenzierbare Funktion*

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

so dass für alle $z \in (f(x_0, y_0) - \epsilon, f(x_0, y_0) + \epsilon)$ die Höhenlinien zur Höhe z von f gerade durch die Graphen der Funktionen

$$(y_0 - \epsilon, y_0 + \epsilon) \to \mathbb{R}, \quad y \mapsto g(z, y)$$

beschrieben werden mit Für festes z zeigen also die Richtungen

$$\left(\frac{\partial g(z,y)}{\partial y}, 1\right)$$

in Richtung der Höhenlinien und stehen senkrecht auf dem Gradienten von f.
(ii) Hyperflächen: *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion, deren Ableitung in eimem Punkt x_0 nicht verschwindet. Dann verschwindet auch mindestens eine partielle Ableitung nicht. Nach einer geegneten Permutation der Varibaeln, können wir $\frac{\partial f}{\partial x_1}(x_0) \neq 0$ annehmen. Sei $y_0 \in \mathbb{R}^{n-1}$ der Vektor der letzten $n-1$ Koordinaten von x_0 Dann lassen sich lokal die Niveaumengen: $\{(x \in \mathbb{R} \mid f(x) = z\}$ mit $z \in (f(x_0) - \epsilon, f(x_0 + \epsilon))$ durch den Graphen einer Funktion $g : (f(x_0 - \epsilon, f(x_0 + \epsilon \times B(y_0, \delta) \to \mathbb{R}$ beschreiben als $(g(z,y), y) \in \mathbb{R} \times B(y_0, \delta)$ mit $(z, y) \in B(f(x_0), \epsilon) \times B(y_0, \delta)$. Lokal werden die Niveauflächen also von $y \in \mathbb{R}^{n-1}$ parametrisiert. Für alle (z, y) in dieser Umgebung von $(f(x_0), y_0)$ ist das Bild der partiellen Ableitungen*

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

$$\textit{dann der Kern von } f'(g(z,y), y) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}).$$

Dieser Kern wird auch Tangentialraum an die Niveauflächen genannt.

Definition 11.13. *Eine unendlich oft (stetig) differenzierbare bijektive Abbildung mit unendlich oft (stetig) differenzierbarer Umkehrabbildung heißt Diffeomorphismus.*

Beispiel 11.14. *Polarkoordinaten Die Abbildung*

$$\mathbb{R}^+ \times \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}^2 \setminus \{0\}, \quad (r, \phi) \mapsto (r\cos\phi, \sin\phi)$$

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

$$\begin{aligned} (x,y) \mapsto (r,\phi) \qquad & \textit{mit} \qquad r = \sqrt{x^2+y^2} \\ & \textit{und} \qquad \phi = \begin{cases} \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \textit{für } y \geq 0 \\ 2\pi - \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \textit{für } y < 0 \end{cases} \end{aligned}$$

Also ist diese Abbildung ein Diffeomorphismus von $\mathbb{R}^+ \times \mathbb{R}/2\pi\mathbb{Z}$ nach $\mathbb{R}^2 \setminus \{0\}$. 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 Unterlänge kleiner ist als 2π, verschiedene Elemente verschiedene Äquivalenzklassen repräsentieren. Deshalb sind die Einschränkungen der Abbildung $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$ auf beliebige offene Intervalle mit Längen nich t gößer als 2π, die jedes Element auf die entsprechende Äquivalenzklasse abbilden, Diffeomorphismen.

11.3 Lagrangemultiplikatoren

Ziel dieses Abschnittes ist es ein Verfahren vorzustellen, mit dem man die lokalen Extremwerte von Funktionen auf solchen Teilmengen eines Banachraumes bestimmen kann, die die Nullstellen von endlich vielen reellen diffenzierbaren 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 sehr vielen Anwendungen gerade 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 garantiert.

Definition 11.15. *Sei $U \subset X$ eine offene Teilmenge eines Banchraumes X und $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$ ein Funktion von U nach $\mathbb{R}^m$. Dann definiert für jedes $x_0 \in U$*

$$A = x \in U \mid g(x) = g(x_0)$$

eine abgeschlossene Menge A von U auf der die reellen Funktionen $g_1, \ldots, g_m(x_0)$ konstant sind. Wir nennen solche Mengen Niveaumengen zu den Zwangsbedingungen $g_1, \ldots, g_n$,

Typischerweise werden die Zwangsbedingungen glatte Funktionen sein. Aber selbst in diesem Fall sind die Niveaumengen nicht immer glatt. Sie können Singularitäten besitzen. Wenn allerdings der Rang der linearen Abbildung

$$g'(x_0) = (g'_1(x_0), \ldots, g'_m(x_0)) : X \to \mathbb{R}^m$$

auf einer offenen Menge konstant ist, dann wissen wir wegen dem Satz der implizierten Funktion, dass dort die Niveaumengen glatt sind. In der folgenden Diskussion wollen wir uns zunächst auf solche Punkte der Niveaumenge beschränken, an denen der Rang von g' gleich m ist. Wenn g stetig differenzierbar ist, und der Rang der Jacobimatrix $g'(x_0)$ an einem Punkt $x_0 \in U$ gleich m ist, dann gibt es auch eine Umgebung von x_0, auf dem der Rang der Jacobimatrix gleich m ist. Um in diesem Fall den Satz der impliziten Funktion anzuwenden müssen wir X so in ein kartesisches Produkt von $\mathbb{R}^m$ mit einem normierten Vektorraum zerlegen, dass die erste partielle Ableitung als Element von $\mathcal{L}(\mathbb{R}^m)$ invertierbar ist. In diesem Fall folgt aus dem Satz der implizierten Funktion, dass wir die Niveaumengen lokal in einer Umgebubg von x_0 durch offene Teilmengen eines normierten Vektorraumes parametrisieren können. Wir nennen alle Funktionen f auf A, die Einschränkungen von glatten Funktionen f auf offenen Umgebungen von A sind, glatte Funktionen auf A.

Beispiel 11.16. **(i)** $$g : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto g(x, y) = x^2 + y^2.$$

Der Gradient $\nabla g(x,y) = (2x, 2y)$ von g verschwindet nur bei $(x,y) = 0$. Also sind alle Niveaumengen $g(x,g) = g_0$ 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 allerdings nur aus dem Nullpunkt. Er ist eine Singularität der Niveaumenge.

(ii)
$$g : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto g(x,y) = x^2 - y^2.$$
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 jewils zwei Hyperebenen. Die Teilmenge $g(x,y) = x^2 - y^2 = (x-y)(x+y) = 0$ 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)
$$g : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto g(x,y) = y^2 - x^3.$$
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 $g = \pm\sqrt{x^3}$ 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 11.17. *(Kritische Punkte auf Niveaumengen) Sei $U \subset X$ eine offene Teilmenge eines Banachraumes und $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktionen auf U. Für ein $x_0 \in U$ sei A die entsprechende Niveaumenge $A = \{x \in U \mid g(x) = g(x_0)\}$. Sei f eine glatte Funktion auf einer Umgebung von x_0 in X. Wenn die Niveaumenge A bei x_0 keine Singularität hat, also die Ableitungen $g_1'(x_0), \ldots, g_m'(x_0)$ linear unabhängig sind in $\mathcal{L}(X, \mathbb{R})$, dann ist x_0 genau dann ein kritischer Punkt von der Einschränkung $f|_A$ von f auf die Niveaumenge A, wenn es Zahlen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ gibt, so dass in $\mathcal{L}(X, \mathbb{R})$*

$$f'(x_0) = \lambda_1 g_1'(x_0) + \ldots + \lambda_m g_m'(x_0)$$

gilt. Die reellen Zahlen $\lambda_1, \ldots, \lambda_m$ heißen Lagrangemultiplikatoren.

Beweis: Weil $g_1'(x_0), \ldots, g_m'(x_0)$ in $\mathcal{L}(X, \mathbb{R})$ linear unabhängig sind, gibt es einen m-dimensionalen Unterraum $Y \subset X$ von X, so dass die lineare Abbildung

$$g'(x_0) = (g_1'(x_0), \ldots, g_m'(x_0)) : Y \to \mathbb{R}^m$$

eine bijektive lineare Abbildung ist, und damit wegen Satz 9.56 eine stetige lineare Abbildung mit stetiger linearer Umkehrabbildung von Y nach $\mathbb{R}^m$ ist. Sei $Z \subset X$ der Unterraum

$$Z = \{x \in X \mid g'(x_0)x = 0\}.$$

Er ist das Urbild von $\{0\} \subset \mathbb{R}$ unter $g'(x_0)$ und damit abgeschlossen. Offenbar besitzt Y eine Basis von Elementen $x_1, \ldots, x_m$, so dass

$$g'_j(x_0)x_i = \delta_{ij} = \begin{cases} 0 \text{ für } j \neq i \\ 1 \text{ für } j = 1 \end{cases} \qquad \text{mit } i, j \in \{1, \ldots, m\}$$

gilt. Dann sind die Abbildungen

$$X \to Y, \quad x \mapsto (g'_1(x_0)x)x_1 + \ldots + (g'_m(x_0)x)x_m$$

und

$$X \to Z, \quad x \mapsto x - (g'_1(x_0)x)x_1 - \ldots - (g'_m(x_0)x)x_m$$

stetige lineare Abbildungen. Die Verkettung von der kannonischen Einbettung $Y \hookrightarrow X$ mit der ersten ergibt offenbar die identische Abbildung von Y. Deshalb setzen sich beide Abbildungen zu der Umkehrabbildung von

$$Y \times Z \to X, \quad (y, z) \mapsto y + z$$

zusammen. Also sind die normierten Vektorräume X und $Y \times Z$ isomorph mit äquivalenten Normen. Wegen dem Satz der implizierten Funktion gibt es dann offene Umgebungen O von $(g_1(x_0), \ldots, g_m(x_0))$ in $\mathbb{R}^m$, O' von z_0 in Z und U' von x_0 in X, wobei $x_0 = (y_0, z_0) \in Y \times Z$ entspricht, und eine stetig differenzierbare Abbildung $h : O \times O' \to X$, so dass für alle $u \in O$ alle Lösungen der Gleichung $g(x) = u$ in U' zum Graphen der Funktion $z \mapsto h(u, z)$ gehören, also von der Form $h(u, z) + z$ sind mit $z \in O'$. Wir haben den Unterraum $Z \subset X$ gerade so definiert, dass die partielle Ableitung $\frac{\partial g}{\partial z}$ an der Stelle $x_0 = y_0 + z_0$ vreschwindet. Deshalb verschwindet auch die partielle Ableitung $\frac{\partial h}{\partial z}(g(x_0), z_0)$. Also ist das Bild von Z unter der Ableitung von $z \mapsto h(g(x_0), z_0) + z$ an der Stelle $z = z_0$ gleich Z. Dann ist x_0 genau dann ein kritischer Punkt, wenn $f'(x_0)$ auf Z verschwindet. Das ist wegen der Definition von $x_1, \ldots, x_m$ äquivalent zu

$$\begin{aligned} f'(x_0)x &= f'(x_0)\left((g'_1(x_0)x)x_1 + \ldots + (g'_m(x_0)x)x_m\right) \qquad \text{für alle } x \in X \\ &= (f'(x_0)x_1)g'_1(x_0)x + \ldots + (f'(x_0)x_m)g'_m(x_0)x \\ &= \lambda_1 g'_1(x_0)x + \ldots + \lambda_m g'_m(x_0)x \quad \text{mit} \\ \lambda_1 &= f'(x_0)x_1, \quad \ldots, \quad \lambda_m = f'(x_0)x_m. \end{aligned}$$

Daraus folgt sofort $\qquad f'(x_0) = \lambda_1 g'_1(x_0) + \ldots + \lambda_m g'_m(x_0).$ **q.e.d.**

Die Singularitäten von den Niveaumengen können allerdings auch lokale Extemwerte sein. Deshalb muss man im Allgemeinen erst alle Singularitäten von den Niveaumengen bestimmen und dann noch alle glatten kritischen Punkte, um alle kritischen Punkte von der Einschränkung von f auf die Niveaumenge zu bestimmen.

Korollar 11.18. *Sei $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion auf einer offenen Teilmenge eines Banachraumes X. Wenn bei allen $x \in U'$ aus einer offenen Teilmenge U' von U die Ableitungen $g'_1 \ldots, g'_{m-1}$ linear unabhängig sind und die Ableitung von g_m linear von $g_1, \ldots, g_{m-1}$ abhängt, dann ist für alle $x_0 \in U'$ g_m loakl auf folgender Niveaumenge A' konstant:*

$$A' = \{x \in U' \mid g_1(x) = g_1(x_0), \ldots, g_{m-1}(x) = g_{m-1}(x_0)\}.$$

Beweis: Mit der analogen Anwendung des Satzes der impliziten Funktion im vorangehenden Satz auf die Nivequmenge A' folgt, dass A' sich auf U' lokal mit stetig differenzierbaren Abbildungen bijektiv auf offene Teilmengen eines Banchraumes abbilden läßt, also glatt ist, und dass alle Punkte dieser Niveaumenge kritische Punkte von g_m sind. Dann ist g_m lokal konstant. **q.e.d.**

Durch mehrfaches Anwenden und Umordnen der Funktionen $g_1, \ldots, g_m$ können wir für jede nicht konstante stetig differenzierbare Funktion $g = (g_1, \ldots, g_m) : U \to \mathbb{R}^m$ erreichen, dass es einen Punkt $x_1 \in U$ und ein $1 \leq l \leq m$ gibt, so dass $g'_1(x_1), \ldots, g'_l(x_1)$ linear unabhängig sind, aber $g'_{l+1}, \ldots, g'_m$ an allen Punkten von U linear von $g'_1, \ldots, g'_l$ abhängen. Weil g' stetig ist folgt aus Lemma 11.6, dass die Menge aller x_1, bei denen $g'_1, \ldots, g'_l$ linear unabhängig sind, sogar eine offene Teilmenge von U ist. Danach bestimmen wir für ein gegebenes $x_0 \in U$ alle Singularitäten der entsprechenden Niveaumenge

$$\{x \in U \mid g_1(x) = g_1(x_0), \ldots, g(x) = g_l(x_0)\}.$$

Zuletzt bestimmen wir alle kritischen glatten Punkte von einer differenzierbaren Funktion f auf dieser Niveaumenge. Dadurch haben wir alle möglichen lokalen Extremwerte bestimmt.