XI. Der Satz über die Umkehrfunktion

In diesem Kapitel werden wir sehen, wie man die differenzierbare Version des Satzes über die Umkehrfunktion auf Funktionen in mehreren Veränderlichen verallgemeinert. In der eindimensionalen Situation konnten wir die Ordnungsstruktur von $\mathbb{R}$ ausnutzen, da stetige Funktionen auf Intervallen genau dann injektiv sind, wenn sie monoton sind. Im Mehrdimensionalen wird die Situation komplizierter. Der entsprechende Satz über die Umkehrfunktion ist ein zentrales Ergebnis der Differentialrechnung mehrerer Veränderlicher. Er hat wichtige Anwendungen auf die Beschreibung von Lösungsmengen von Gleichungen, der wir uns im nächsten Kapitel zuwenden.

XI.1. Der Banachsche Fixpunktsatz

Wir betrachten auf dem metrischen Raum $\mathbb{R}^n$ immer die euklidische Norm

$$\|x\| := \|x\|_2 := \Big(\sum_{j=1}^n x_j^2\Big)^{\frac{1}{2}}.$$

Auf dem Raum $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ betrachten wir die zugehörige Operatornorm $\|A\| := \sup\{\|Ax\| : x \in \mathbb{R}^n, \|x\| \leq 1\}$. Wir werden oft den linearen Abbildungen $A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ ihre Matrix bzgl. der kanonischen Basen in $\mathbb{R}^n$ und $\mathbb{R}^m$ zuordnen. Wir erhalten so eine bijektive lineare Abbildung

$$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to M_{m,n}(\mathbb{R}), \quad A \mapsto (a_{ij}), \quad \text{wobei} \quad A(e_j) = \sum_{i=1}^m a_{ij} e_i$$

gilt.

Definition XI.1.1. Es sei $n \in \mathbb{N}$. Dann heißt die Menge

$$\mathrm{GL}(\mathbb{R}^n) := \mathrm{Aut}(\mathbb{R}^n) = \{A \in \mathrm{End}(\mathbb{R}^n) : A \text{ invertierbar }\}$$

die *allgemeine lineare Gruppe* (*general linear group*). Identifizieren wir lineare Endomorphismen von $\mathbb{R}^n$ mit ihren Matrizen bzgl. der kanonischen Basis, so führt diese zu einer Identifikation von $\mathrm{End}(\mathbb{R}^n)$ mit dem Raum $M_n(\mathbb{R})$ der $(n \times n)$-Matrizen. In diesem Sinn schreiben wir auch $\mathrm{GL}_n(\mathbb{R})$ für die Menge der invertierbaren $(n \times n)$-Matrizen. ■

Bemerkung XI.1.2. (a) Für $A \in \mathrm{End}(\mathbb{R}^n)$ ist

$$A \in \mathrm{GL}(\mathbb{R}^n) \iff \det A \neq 0.$$

(b) Die Abbildung

$$\det : M_n(\mathbb{R}) \cong \mathbb{R}^{n^2} \to \mathbb{R}, \qquad A = (a_{ij}) \mapsto \det A = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

ist eine stetige Funktion. Sie ist ein Polynom des Grades n. Hierbei ist S_n die $n!$-elementige Gruppe der Permutationen der Menge $\{1,\ldots,n\}$ (vgl. Bemerkung I.4.5). ■

Aufgabe XI.1.1. Sei $U \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N}$. Ist $f \in C^k(U,\mathbb{R}^m)$ und $g \in C^k(U)$ mit $g(x) \neq 0$ für alle $x \in U$, so ist $\frac{1}{g} f \in C^k(U,\mathbb{R}^m)$. Die entsprechende Aussage gilt auch für $k = \infty$. ■

Satz XI.1.3. (a) *Die Gruppe* $\mathrm{GL}(\mathbb{R}^n)$ *ist eine offene Teilmenge von* $\mathrm{End}(\mathbb{R}^n)$.
(b) *Die Inversion* $\mathrm{GL}(\mathbb{R}^n) \to \mathrm{GL}(\mathbb{R}^n) \subseteq \mathrm{End}(\mathbb{R}^n)$ *ist beliebig oft differenzierbar.*

Beweis. Für den Beweis identifizieren wir lineare Abbildungen mit Matrizen.

(a) Nach Teil (a) von Bemerkung XI.1.2 ist

$$\mathrm{GL}_n(\mathbb{R}) = \det{}^{-1}(\mathbb{R} \setminus \{0\}).$$

Da $\mathbb{R} \setminus \{0\}$ in $\mathbb{R}$ offen ist und det stetig, ist $\mathrm{GL}_n(\mathbb{R})$ offen (vgl. Satz IV.1.4).
(b) Dies folgt aus der expliziten Formel zur Berechnung der Inversen. Die Cramersche Regel besagt

$$A^{-1} = \frac{1}{\det A}\big((-1)^{i+j} \det A_{ji}\big),$$

wobei A_{ji} die Matrix ist, die durch Streichen der j-ten Zeile und der i-ten Spalte entsteht. Insbesondere ist jeder Eintrag der inversen Matrix A^{-1} eine beliebig oft differenzierbare Funktion der Einträge der Matrix A, denn wegen $\det(A) \neq 0$ ist der Nenner immer $\neq 0$ (siehe Aufgabe XI.1.1). ■

Die folgende Aufgabe enthält eine wichtige Verallgemeinerung des Weierstraßschen Konvergenzkriteriums, das wir im Kontext der gleichmäßigen Konvergenz von Funktionenreihen kennengelernt hatten.

Aufgabe XI.1.2. (Allgemeines Majorantenkriterium) Sei $(V, \|\cdot\|)$ ein Banachraum. Ist $(v_n)_{n\in\mathbb{N}}$ eine Folge in V mit $\sum_{n=1}^\infty \|v_n\| < \infty$, so ist die Reihe

$$\sum_{k=1}^\infty v_k := \lim_{n\to\infty} \sum_{k=1}^n v_k$$

konvergent. Hinweis: Die Folge der Teilsummen ist eine Cauchy-Folge. ■

Satz XI.1.4. *Ist* $A \in \mathrm{End}(\mathbb{R}^n)$ *mit* $\|A\| < 1$ *und* $\mathbf{1} := \mathrm{id}_{\mathbb{R}^n}$*, so ist* $\mathbf{1} - A$ *invertierbar und die Neumannsche Reihe*

$$(\mathbf{1} - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

konvergiert.

Beweis. Wegen der Submultiplikativität der Norm ist $\|A^k\| \leq \|A\|^k$ für alle $k \in \mathbb{N}_0$. Es folgt

$$\sum_{k=0}^{\infty} \|A^k\| \leq \sum_{k=0}^{\infty} \|A\|^k = \frac{1}{1 - \|A\|}.$$

Also ist die Reihe $\sum_{k=0}^{\infty} A^k$ in $(\mathrm{End}(\mathbb{R}^n), \|\cdot\|)$ nach dem Majorantenkriterium konvergent (Aufgabe XI.1.2). Weiter ist

$$(\mathbf{1} - A) \sum_{k=0}^{\infty} A^k = \sum_{k=0}^{\infty} A^k - \sum_{k=1}^{\infty} A^k = \mathbf{1},$$

also $\sum_{k=0}^{\infty} A^k = (\mathbf{1} - A)^{-1}$. ■

Der Banachsche Fixpunktsatz

Definition XI.1.5. Sei (X, d) ein metrischer Raum. Eine Abbildung $f : X \to X$ heißt *Kontraktion* mit *Kontraktionskonstante* $\lambda < 1$, wenn für alle $x, y \in X$ die Beziehung

$$d\big(f(x), f(y)\big) \leq \lambda \cdot d(x, y)$$

gilt. Da Kontraktionen Lipschitz-stetig sind, sind sie insbesondere stetig. ■

$\boxed{\text{Banachscher Fixpunktsatz}}$

Satz XI.1.6. *Ist* (X, d) *ein nichtleerer vollständiger metrischen Raum, so hat jede Kontraktion* $f : X \to X$ *genau einen Fixpunkt, d.h. es existiert genau ein* $x \in X$ *mit* $f(x) = x$*. Es gilt sogar*

$$x = \lim_{n \to \infty} f^n(y)$$

für jeden beliebigen Punkt $y \in X$*.*

Beweis. Sei $\lambda < 1$ die Kontraktionskonstante.
Eindeutigkeit: Seien x und y Fixpunkte. Dann gilt

$$d(x, y) = d\big(f(x), f(y)\big) \leq \lambda \cdot d(x, y),$$

woraus $d(x,y)=0$ folgt, d.h. $x=y$.
Existenz: Wir zeigen zuerst, dass für beliebiges $y \in X$ die Folge $\big(f^n(y)\big)_{n\in\mathbb{N}}$ konvergiert. Da X nach Voraussetzung vollständig ist, haben wir zu zeigen, dass dies eine Cauchy-Folge ist. Durch vollständige Induktion erhalten wir zunächst, dass für alle $n \in \mathbb{N}$ und $a,b \in X$ die Beziehung

$$d\big(f^n(a), f^n(b)\big) \leq \lambda^n \cdot d(a,b)$$

gilt. Für $n,k \in \mathbb{N}$ folgt hieraus durch wiederholtes Anwenden der Dreiecksungleichung

$$\begin{aligned}
&d\big(f^{n+k}(y), f^n(y)\big) \leq \lambda^n \cdot d\big(f^k(y), y\big) \\
&\leq \lambda^n \Big(d\big(f^k(y), f^{k-1}(y)\big) + d\big(f^{k-1}(y), f^{k-2}(y)\big) + \ldots + d\big(f(y), y\big) \Big) \\
&\leq \lambda^n \Big(\lambda^{k-1} d\big(f(y), y\big) + \ldots + \lambda d\big(f(y), y\big) + d\big(f(y), y\big) \Big) \\
&= \lambda^n (\lambda^{k-1} + \ldots + \lambda + 1) \cdot d\big(f(y), y\big) \\
&= \lambda^n \frac{1-\lambda^k}{1-\lambda} \cdot d\big(f(y), y\big) \leq \frac{\lambda^n}{1-\lambda} \cdot d\big(f(y), y\big).
\end{aligned}$$

Wegen $\lambda^n \to 0$ für $n \to \infty$ ist die Folge $\big(f^n(y)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge und konvergiert daher gegen ein $x \in X$. Dann gilt wegen der Stetigkeit von f

$$f(x) = f\big(\lim_{n\to\infty} f^n(y)\big) = \lim_{n\to\infty} f^{n+1}(y) = x.$$

Also ist x ein Fixpunkt von f. ■

Der Banachsche Fixpunktsatz ist ein wichtiges Werkzeug der Analysis, das uns auch in anderen Situationen wieder begegnen wird. Des öfteren wird dabei X eine Kugel im $\mathbb{R}^n$ sein, oder auch eine Teilmenge eines Funktionenraums (z.B. in der Vorlesung über Differentialgleichungen). Insbesondere in der numerischen Mathematik ist der Banachsche Fixpunktsatz ein Instrument, mit dem sich die Konvergenz von Approximationsverfahren bzw. von numerischen Lösungsmethoden in vielen Fällen beweisen lässt.

Kriterium für Vollständigkeit

Lemma XI.1.7. *Ist (X,d) ein vollständiger metrischer Raum, so ist eine Teilmenge $A \subseteq X$ genau dann abgeschlossen, wenn der metrische Raum $(A, d|_{A\times A})$ vollständig ist.*

Beweis. "$\Rightarrow$" Ist A abgeschlossen und $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in A, so ist sie auch eine Cauchy-Folge in X. Daher existiert ein $x \in X$ mit $x = \lim_{n\to\infty} x_n$. Dann ist $x \in \overline{A} = A$, da A abgeschlossen ist. Folglich konvergiert $(x_n)_{n\in\mathbb{N}}$ in A.
"$\Leftarrow$" Sei andererseits A nicht abgeschlossen und $x \in \overline{A} \setminus A$. Dann existiert eine Folge in A, die gegen x konvergiert. Dies ist dann eine Cauchy-Folge in A, die in A nicht konvergiert. Somit ist A nicht vollständig. ■

Bemerkung XI.1.8. Der Banachsche Fixpunktsatz hat eine anschauliche Konsequenz, die man leicht selbst beobachten kann. Hierzu sei X das Stadtgebiet Darmstadts, das wir als abgeschlossene Teilmenge des $\mathbb{R}^2$ auffassen. Dann ist X ein vollständiger metrischer Raum (Lemma IX.1.7). Wir breiten nun einen Stadtplan von Darmstadt an irgendeiner Stelle auf dem Boden in Darmstadt aus und betrachten die Abbildung $f\colon X \to X$, die jedem Punkt $x \in X$ denjenigen Punkt $f(x)$ zuordnet, über dem das Bild des Punktes x auf dem Stadtplan liegt. Ist der Stadtplan eine Karte im Maßstab $1 : n$, so ist f eine Kontraktion des metrischen Raumes X mit der Kontraktionskonstante $\lambda = \frac{1}{n} < 1$. Nach dem Banachschen Fixpunktsatz gibt es also genau einen Punkt $x \in X$ im Darmstädter Stadtgebiet für den $f(x) = x$ gilt, d.h. x liegt genau unter demjenigen Punkt des Stadtplans, der diesen Punkt abbildet. ∎

Bemerkung XI.1.9. (a) Die Bedingung der Vollständigkeit ist wesentlich. So ist beispielsweise der metrische Unterraum $X =]0,1]$ von $\mathbb{R}$ bzgl. der üblichen Metrik nicht vollständig, da er nicht abgeschlossen ist. Die Abbildung $f : X \to X$, $f(x) = \frac{1}{2}x$ ist eine Kontraktion, die in X *keinen* Fixpunkt besitzt.
(b) Sei $A \in \mathrm{End}(\mathbb{R}^n)$, $y \in \mathbb{R}^n$ und $\|A\| < 1$. Wir wollen die Gleichung

$$(\dagger) \qquad (\mathbf{1} - A)x = y$$

lösen. Hierzu betrachten wir die Abbildung

$$f : \mathbb{R}^n \to \mathbb{R}^n, \; z \mapsto Az + y.$$

Die Fixpunkte dieser Abbildung sind dann genau die Lösungen der Gleichung $(\dagger)$. Es gilt

$$\|f(z_1) - f(z_2)\| = \|A(z_1 - z_2)\| \le \|A\| \cdot \|z_1 - z_2\|$$

mit $\|A\| < 1$, das heißt, f ist eine Kontraktion. Damit hat f genau einen Fixpunkt in $\mathbb{R}^n$, d.h., die Gleichung $(\dagger)$ hat genau eine Lösung. Wir erhalten aus dem Beweis des Banachschen Fixpunktsatzes sogar ein Verfahren, mit dem wir die Lösung näherungsweise berechnen können. Die Folge $\big(f^n(y)\big)_{n\in\mathbb{N}}$ ist gegeben durch

$$f^0(y) = y, \quad f^1(y) = Ay + y, \quad f^2(y) = A(Ay + y) + y = A^2y + Ay + y,$$

und für allgemeine n ist

$$f^n(y) = \sum_{k=0}^{n} A^k y,$$

was man leicht per Induktion beweist. Es folgt

$$(\mathbf{1} - A)^{-1}y = x = \lim_{n\to\infty} f^n(y) = \sum_{k=0}^{\infty} A^k y.$$

Insbesondere erhalten wir so einen neuen Beweis für Satz XI.1.4. ∎

XI.2. Der Satz über die Umkehrfunktion

Definition XI.2.1. Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen. Eine Abbildung $f : U \to V$ heißt C^k-*Diffeomorphismus*, wenn sie eine C^k-Abbildung ist, bijektiv und ihre Umkehrabbildung $f^{-1} : V \to U$ ebenfalls C^k ist. ■

Lemma XI.2.2. *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen und $f : U \to V$ ein C^1-Diffeomorphismus mit $f \in C^k(U, \mathbb{R}^m)$. Dann gelten folgende Aussagen:*

(a) *Für jedes $u \in U$ ist das Differential $\mathtt{d}f(u) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ invertierbar mit*
$$\mathtt{d}f(u)^{-1} = \mathtt{d}(f^{-1})(f(u)).$$

(b) $m = n$.

(c) *Die Abbildung $f^{-1} : V \to \mathbb{R}^n$ ist k-mal stetig differenzierbar.*

Beweis. (a) Sei $u \in U$ und $v = f(u)$. Dann folgt aus $f^{-1} \circ f = \mathrm{id}_U$ und $f \circ f^{-1} = \mathrm{id}_V$ mit der Kettenregel
$$\mathtt{d}\big(f^{-1}\big)\big(f(u)\big) \circ \mathtt{d}f(u) = \mathtt{d}(f^{-1} \circ f)(u) = \mathrm{id}_{\mathbb{R}^n}$$
$$\mathtt{d}f\big(f^{-1}(v)\big) \circ \mathtt{d}\big(f^{-1}\big)(v) = \mathtt{d}(f \circ f^{-1})(v) = \mathrm{id}_{\mathbb{R}^m},$$
also $\mathtt{d}\big(f^{-1}\big)\big(f(u)\big) = \big(\mathtt{d}f(u)\big)^{-1}$, und $\mathtt{d}f(u)$ ist eine invertierbare lineare Abbildung.

(b) Dies folgt sofort aus der Invertierbarkeit von $\mathtt{d}f(u)$ (Lineare Algebra).

(c) Wir erinnern uns, dass die Inversion $\mathrm{GL}(\mathbb{R}^n) \to \mathrm{GL}(\mathbb{R}^n), g \mapsto g^{-1}$, beliebig oft differenzierbar ist (Satz XI.1.3(b)). Ist $f \in C^k$ und $f^{-1} \in C^\ell$, wobei $0 \leq \ell \leq k-1$, so ist die Funktion
$$\mathtt{d}(f^{-1}) : V \to \mathrm{GL}(\mathbb{R}^n), \qquad v \mapsto \big(\mathtt{d}f\big(f^{-1}(v)\big)\big)^{-1}$$
ℓ-mal stetig differenzierbar, also f^{-1} mindestens $\ell + 1$ mal stetig differenzierbar. Induktiv ergibt sich also $f^{-1} \in C^k(V, \mathbb{R}^n)$. ■

Bemerkung XI.2.3. Sei $f : U \to V$ eine stetig differenzierbare, surjektive Funktion, deren Differential $\mathtt{d}f(x)$ in allen Punkten $x \in U$ invertierbar ist. Kann man hieraus schon schließen, dass f ein Diffeomorphismus ist? Sind $U, V \subseteq \mathbb{R}$ offene Intervalle, so ist dies richtig, denn die Invertierbarkeit des Differentials $\mathtt{d}f(x)$ ist im Eindimensionalen genau die Bedingung, dass $f'(x) \neq 0$ ist. Ist dies für alle $x \in U$ der Fall, so ist f streng monoton, insbesondere also injektiv. Im Mehrdimensionalen trifft dies jedoch nicht mehr zu. Hierzu betrachten wir die Funktion
$$f : \begin{array}{ccc}]0, \infty[\, \times \mathbb{R} & \to & \mathbb{R}^2 \setminus \{(0,0)\} \\ (r, \varphi) & \mapsto & (r\cos\varphi, r\sin\varphi). \end{array}$$

Die Jacobimatrix
$$J_{(r,\varphi)}(f) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$
ist wegen $\det\big(J_{(r,\varphi)}(f)\big) = r \neq 0$ immer invertierbar. Die Funktion f ist auch surjektiv. Wegen $f(r, 2\pi+\varphi) = f(r,\varphi)$ für alle (r,φ) ist sie nicht injektiv, also kein Diffeomorphismus.

Trotzdem hat f „lokale Umkehrfunktionen": Ist beispielsweise
$$U :=]0,\infty[\times]\varphi_0 - \pi, \varphi_0 + \pi[$$
für ein $\varphi_0 \in \mathbb{R}$, so ist $f|_U : U \to f(U)$ injektiv, und
$$V := f(U) = \mathbb{R}^2 \setminus \big(\mathbb{R}^+ \cdot (\cos(\varphi_0+\pi), \sin(\varphi_0+\pi))\big)$$
ist offen in $\mathbb{R}^2$. Man kann also eine Umkehrfunktion $f^{-1} : V \to U$ definieren. Im folgenden werden wir uns der Frage zuwenden, wie es mit den Differenzierbarkeitseigenschaften solcher Umkehrfunktionen steht. ■

Definition XI.2.4. Eine C^k-Abbildung $f : U \to \mathbb{R}^n$ heißt *lokal um $u \in U$ invertierbar*, wenn es offene Umgebungen U_1 von u und V_1 von $f(u)$ derart gibt, dass $f|_{U_1} : U_1 \to V_1$ ein Diffeomorphismus ist. Die Abbildung
$$\big(f|_{U_1}\big)^{-1} \;:\; V_1 \to U_1$$
heißt dann *lokale Umkehrfunktion* von f. Ist f in jedem Punkt $u \in U$ lokal invertierbar, so heißt f *lokaler Diffeomorphismus.* ■

Satz über die Umkehrfunktion

Theorem XI.2.5. *Sei $U \subseteq \mathbb{R}^n$ offen, $u \in U$ und $f : U \to \mathbb{R}^n$ eine C^k-Abbildung. Genau dann ist f um u lokal invertierbar, wenn $\mathtt{d}f(u)$ invertierbar ist. Die lokale Umkehrfunktion ist dann ebenfalls eine C^k-Abbildung.*

Beweis. Ist f um u lokal invertierbar, so ist $\mathtt{d}f(u) = \mathtt{d}(f \mid_{U_1})(u)$ nach Lemma XI.2.2 invertierbar. Sei nun umgekehrt $\mathtt{d}f(u)$ invertierbar. Wir zeigen, dass f lokal um u invertierbar ist. Dazu reduzieren wir die Situation zunächst auf eine einfachere. Zunächst können wir wegen Lemma XI.2.2(c) annehmen, dass $k = 1$ ist. Ferner können wir
$$u = 0, \qquad f(u) = 0 \quad \text{und} \quad \mathtt{d}f(0) = \mathbf{1} = \mathrm{id}_{\mathbb{R}^n}$$
annehmen, indem wir die Funktion f durch $\widetilde{f}(x) := \mathtt{d}f(u)^{-1}\big(f(u+x) - f(u)\big)$ ersetzen. Dann ist
$$f(x) = \mathtt{d}f(u)\big(\widetilde{f}(x-u)\big) + f(u) \quad \text{für} \quad x \in U$$

und falls $\widetilde{f}$ auf einer Teilmenge von $\widetilde{U} := U - u$ invertierbar sein sollte, erhalten wir direkt eine Umkehrfunktion der entsprechenden Einschränkung von f durch

$$f^{-1}(y) = \widetilde{f}^{-1}\big(\mathtt{d}f(u)^{-1}(y - f(u))\big) + u$$

indem wir die Formel für f nach $y = f(x)$ auflösen.

Nun wollen wir also für „kleine" y die Gleichung $f(x) = y$ nach x auflösen. Definieren wir

$$g_y : U \to \mathbb{R}^n, \quad x \mapsto y + \big(x - f(x)\big),$$

so entspricht dies der Fixpunktgleichung

$$x = g_y(x) = y + \big(x - f(x)\big).$$

Wir suchen also Fixpunkte der Abbildungen g_y. Wir müssen eine offene Menge finden, so dass für alle y aus dieser Menge die Abbildung g_y eine Kontraktion ist, denn dann wissen wir nach dem Banachschen Fixpunktsatz, dass g_y in jeder abgeschlossenen Teilmenge von U, die unter g_y invariant ist, genau einen Fixpunkt besitzt. Zunächst betrachten wir $y = 0$.

Für $g(x) := g_0(x) = x - f(x)$ gilt $g(0) = 0$ und $\mathtt{d}g(0) = \mathrm{id} - \mathrm{id} = 0$. Wegen $g \in C^1(U, \mathbb{R}^n)$ existiert ein $r > 0$ mit $U_{2r}(0) \subseteq U$ und $\|\mathtt{d}g(x)\| \leq \frac{1}{2}$ für alle x mit $\|x\| \leq r$ (ε-δ-Stetigkeit von $\mathtt{d}f$). Aus dem Satz vom endlichen Zuwachs X.2.16 folgt damit

$$\|g(x) - g(x')\| \leq \tfrac{1}{2}\|x - x'\|$$

für alle x und x' mit $\|x\|, \|x'\| \leq r$. Insbesondere erhalten wir für $x' = 0$ die Beziehung $\|g(x)\| \leq \frac{r}{2}$ für $\|x\| \leq r$. Sei $X := \{x \in \mathbb{R}^n : \|x\| \leq r\}$. Dann ist

$$g_y : X \to X, \quad g_y(x) := g(x) + y$$

für alle y mit $\|y\| \leq \frac{r}{2}$ ebenfalls eine Kontraktion, denn wir haben für $x \in X$:

$$\|g_y(x)\| = \|y + g(x)\| \leq \|y\| + \|g(x)\| \leq \frac{r}{2} + \frac{r}{2} = r.$$

Da X in dem vollständigen metrischen Raum $(\mathbb{R}^n, d_2)$ abgeschlossen ist, also nach Lemma XI.1.7 ein vollständiger metrischer Raum, folgt die Existenz genau eines $x \in X$ mit $g_y(x) = x$ aus dem Banachschen Fixpunktsatz XI.1.6. Wir haben also gezeigt, dass zu jedem y mit $\|y\| \leq \frac{r}{2}$ genau ein x mit $\|x\| \leq r$ und $f(x) = y$ existiert.

Sei

$$U_1 := \{x \in \mathbb{R}^n : \|x\| < r, \|f(x)\| < \frac{r}{2}\} = U_r(0) \cap f^{-1}(U_{\frac{r}{2}}(0)).$$

Da f stetig ist, ist diese Menge offen (Urbilder offener Mengen sind offen). Sei weiter $V_1 := f(U_1)$. Für zwei Punkte $x, x' \in \overline{U_r(0)}$ erhalten wir

$$\begin{aligned}
\|x - x'\| &= \|g(x) + f(x) - \big(g(x') + f(x')\big)\| \\
&\leq \|g(x) - g(x')\| + \|f(x) - f(x')\| \\
&\leq \tfrac{1}{2}\|x - x'\| + \|f(x) - f(x')\|,
\end{aligned}$$

woraus sich $\|x-x'\| \leq 2\|f(x)-f(x')\|$ ergibt. Für $x'=0$ folgt speziell $\|x\| < r$, wenn $\|f(x)\| < \frac{r}{2}$ ist. Damit ist $V_1 = f(U_1) = \{y \in \mathbb{R}^n : \|y\| < \frac{r}{2}\}$; diese Menge ist offen. Ferner sehen wir, dass $f\,|_{U_1} : U_1 \to V_1$ bijektiv ist. Also existiert eine Umkehrabbildung $\varphi := \big(f\,|_{U_1}\big)^{-1} : V_1 \to U_1$, und mit obiger Abschätzung erhalten wir

$$\|\varphi(y) - \varphi(y')\| \leq 2\|y - y'\|,$$

indem man $x = \varphi(y)$ und $x' = \varphi(y')$ setzt. Die Funktion φ ist also stetig.

Nun zeigen wir, dass $\mathtt{d}f(x)$ für alle $x \in U_1$ invertierbar ist. Wir wissen, dass für alle $x \in U_1$ die Beziehung

$$\|\mathtt{d}f(x) - \mathbf{1}\| = \|\mathtt{d}g(x)\| \leq \tfrac{1}{2}$$

gilt, und nach Satz XI.1.4 ist daher $\mathtt{d}f(x)$ invertierbar.

φ **ist stetig differenzierbar:** Sei $v \in V_1$; dann ist $v = f(u)$ für ein $u \in U_1$. Da f in u differenzierbar ist, existiert eine in u stetige Funktion $\Phi : U_1 \to \mathrm{End}(\mathbb{R}^n)$ mit

$$f(u+h) - f(u) = \Phi(u+h)(h)$$

für alle $u + h \in U_1$. Wegen der Invertierbarkeit von $\Phi(u) = \mathtt{d}f(u)$ und der Stetigkeit von Φ existiert ein $\delta > 0$ mit $\Phi(u+h) \in \mathrm{GL}(\mathbb{R}^n)$ für alle $h \in \mathbb{R}^n$ mit $\|h\| < \delta$, denn $\mathrm{GL}(\mathbb{R}^n)$ ist in $\mathrm{End}(\mathbb{R}^n)$ offen und für stetige Abbildungen sind Urbilder offener Mengen offen. Für solche h ist dann

$$\varphi\big(f(u+h)\big) - \varphi\big(f(u)\big) = u + h - u = h = \Phi(u+h)^{-1}\big(f(u+h) - f(u)\big).$$

Da φ stetig ist, ist die Menge

$$\{y \in V_1 : \|\varphi(y) - u\| < \delta\} = \varphi^{-1}(U_\delta(0))$$

in $\mathbb{R}^n$ offen. Für $y \in V_1$ setzen wir $h := \varphi(y) - u$. Dann ist $y = f(u+h)$, also

$$\varphi(y) - \varphi(v) = \Phi\big(\varphi(y)\big)^{-1}(y - v),$$

und die Abbildung $y \mapsto \Phi\big(\varphi(y)\big)^{-1}$ ist in v stetig, da φ stetig ist und die Inversion in $\mathrm{GL}(\mathbb{R}^n)$ ebenfalls. Damit haben wir bewiesen, dass φ in v differenzierbar ist mit

$$\mathtt{d}\varphi(v) = \Phi\big(\varphi(v)\big)^{-1} = \Phi(u)^{-1} = \mathtt{d}f(u)^{-1} = \mathtt{d}f\big(\varphi(v)\big)^{-1}.$$

Da $\mathtt{d}\varphi$ auch stetig ist, erhalten wir $\varphi \in C^1(V_1, \mathbb{R}^n)$. ■

Folgerung XI.2.6. *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ eine stetig differenzierbare Funktion sowie $p \in U$ und $r > 0$ mit $\overline{U_r(p)} \subseteq U$ und*

$$\|\mathtt{d}f(x) - \mathbf{1}\| \leq \tfrac{1}{2} \quad \text{für } x \in \overline{U_r(p)}.$$

Dann gilt:

(a) *Zu jedem* y *mit* $\|y - f(p)\| \leq \frac{r}{2}$ *existiert genau ein* $x \in \overline{U_r(p)}$ *mit* $f(x) = y$, *und es gilt für alle* $x, x' \in \overline{U_r(p)}$:

$$\|x - x'\| \leq 2\|f(x) - f(x')\|.$$

Insbesondere ist f *auf* $U_r(p)$ *injektiv.*

(b) *Ist* $\rho < \frac{r}{3}$, *so gilt*

$$U_{\frac{\rho}{2}}\big(f(p)\big) \subseteq f\big(U_\rho(p)\big) \subseteq U_{\frac{r}{2}}(f(p))$$

und $f|_{U_\rho(p)}$ *ist ein Diffeomorphismus auf eine offene Bildmenge.*

Beweis. (a) folgt sofort aus dem ersten Teil des Beweises.

(b) Ist $y \in U_{\frac{\rho}{2}}(f(p))$, so erhalten wir mit (a) sofort ein $x \in \overline{U_r(p)}$ mit $f(x) = y$. Weiter ist

$$\|x - p\| \leq 2\|f(x) - f(p)\| = 2\|y - f(p)\| < \rho,$$

also $U_{\frac{\rho}{2}}\big(f(p)\big) \subseteq f\big(U_\rho(p)\big)$. Um die zweite Inklusion einzusehen, beachten wir, dass für alle $x \in U_r(p)$ gilt

$$\|\mathtt{d}f(x)\| \leq \|\mathtt{d}f(x) - \mathbf{1}\| + \|\mathbf{1}\| \leq \frac{3}{2},$$

also nach dem Satz vom endlichen Zuwachs X.2.16 die Beziehung $\|f(x) - f(p)\| \leq \frac{3}{2}\|x - p\|$. Daher ist $f(U_\rho(p)) \subseteq U_{r/2}\big(f(p)\big)$. Dass die Einschränkung von f auf $U_\rho(p)$ ein Diffeomorphismus auf eine offene Bildmenge ist, folgt aus dem Beweis des Satz über die Umkehrfunktion (alternativ kann man Folgerung XI.2.7 verwenden, da $\mathtt{d}f(x)$ für $x \in U_\rho(p)$ invertierbar ist und f auf dieser Menge injektiv). ■

Der Vorteil der Folgerung XI.2.6 ist, dass man nur wissen muss, dass

$$\|\mathtt{d}f(x) - \mathbf{1}\| \leq \tfrac{1}{2} \quad \text{für alle } x \in U_{3\rho}(u)$$

gilt. Dann kann man direkt eine Umgebung von u angeben, auf der f injektiv ist, und mit Folgerung XI.2.6 eine Umgebung von $f(u)$, die ganz im Bild von f liegt.

Der Satz über die Umkehrfunktion ist ein wichtiges Werkzeug der Analysis. Er dient beispielsweise dazu, in vielen Situationen geeignete „Koordinaten“ einzuführen. Hierbei denken wir uns lokal invertierbare Abbildungen als „Koordinatenwechsel“, genau wie die Basistransformationen in der Linearen Algebra.

Die folgende Konsequenz aus dem Satz über die Umkehrfunktion findet häufige Anwendung, da sie es erlaubt, sich den Nachweis der Differenzierbarkeit der Umkehrfunktion zu ersparen. Man hat nur Daten zu betrachten, die unmittelbar durch die Funktion f selbst gegeben sind.

Folgerung XI.2.7. *Sei $k \in \mathbb{N} \cup \{\infty\}$, $U \subseteq \mathbb{R}^n$ offen und $f \in C^k(U, \mathbb{R}^n)$ injektiv, so dass $\mathtt{d}f(x)$ für alle $x \in U$ invertierbar ist. Dann ist $f(U) \subseteq \mathbb{R}^n$ offen und $f^{-1} \in C^k(f(U), \mathbb{R}^n)$, d.h. die Umkehrfunktion ist automatisch eine C^k-Funktion.*

Beweis. Aus dem Satz über die Umkehrfunktion folgt sofort, dass für jedes $x \in U$ die Menge $f(U)$ eine Umgebung von $f(x)$ ist. Also ist $V := f(U) \subseteq \mathbb{R}^n$ offen. Da $f\colon U \to V$ nach Voraussetzung bijektiv ist, existiert eine Umkehrfunktion $f^{-1}\colon V \to U$. Ist $v = f(u) \in V$, so finden wir mit dem Satz über die Umkehrfunktion offene Umgebungen U_1 von u in U und V_1 von v in V, so dass $f|_{U_1}\colon U_1 \to V_1$ ein C^k-Diffeomorphismus ist. Also ist $f^{-1}|_{V_1} = (f|_{U_1})^{-1}$ eine C^k-Abbildung und somit ist $f^{-1} \in C^k(V, \mathbb{R}^n)$. ∎

XII. Gleichungen und Mannigfaltigkeiten

Es ist eines der Grundanliegen der Mathematik, Gleichungen der Gestalt

$$F(x) = y$$

für eine gegebene rechte Seite y zu lösen, bzw. die Struktur ihrer Lösungsmengen zu beschreiben. In diesem Kapitel werden wir den Fall behandeln, wo $F: U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion und $U \subseteq \mathbb{R}^n$ offen ist. Der Satz über implizite Funktionen, den wir in Abschnitt XII.1 behandeln, ist eine wichtige Folgerung aus dem Satz über die Umkehrfunktion. Er gibt uns die Möglichkeit, die Lösungsmenge

$$\{x \in U: F(x) = y\}$$

geeignet zu parametrisieren. Insbesondere werden wir hierdurch auf den Begriff der Untermannigfaltigkeit des $\mathbb{R}^n$ geführt. In Abschnitt XII.2 wenden wir uns einer weiteren wichtigen Klasse von Problemen zu, die in vielen Anwendungen eine Rolle spielt: den Extremwertaufgaben mit Nebenbedingungen, die durch Gleichungen gegeben sind.

XII.1. Der Satz über implizite Funktionen

Um den folgenden Satz besser zu verstehen, betrachten wir zuerst das folgende lineare Problem. Sei hierzu

$$f: \mathbb{R}^{n+k} \to \mathbb{R}^k$$

eine lineare Abbildung. Wir schreiben die Elemente von $\mathbb{R}^{n+k}$ als Paare (x, y) mit $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^k$. Wir möchten nun die Gleichung $f(x, y) = 0$ nach y auflösen. Da f linear ist, existieren lineare Abbildungen $f_1: \mathbb{R}^n \to \mathbb{R}^k$ und $f_2: \mathbb{R}^k \to \mathbb{R}^k$ mit

$$f(x, y) = f_1(x) + f_2(y) \quad \text{für} \quad (x, y) \in \mathbb{R}^{n+k}.$$

In dieser Darstellung von f sehen wir, dass sich die Gleichung $0 = f(x, y) = f_1(x) + f_2(y)$ eindeutig nach y auflösen lässt, wenn die lineare Abbildung f_2 invertierbar ist. In diesem Fall erhalten wir

$$y = -f_2^{-1}(f_1(x)) \quad \Longleftrightarrow \quad f(x, y) = 0.$$

Der Satz über implizite Funktion ist eine Verallgemeinerung dieser Beobachtung auf nichtlineare Abbildungen. Wegen der Nichtlinearität erhält man allerdings nur eine lokale Aussage.

$\boxed{\text{Satz über implizite Funktionen}}$

Theorem XII.1.1. *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^k$ offen und $f : U \times V \to \mathbb{R}^k$ eine C^m-Abbildung. Für $(x,y) \in U \times V$ spalten wir das Differential $\mathtt{d}f$ von f in zwei Bestandteile auf:*

$$\mathtt{d}f(x,y) = \big(d_1 f(x,y), d_2 f(x,y)\big),$$

mit

$$\begin{aligned} d_1 f(x,y) &= \mathtt{d}f(x,y)|_{\mathbb{R}^n \times \{0\}} \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^k), \\ d_2 f(x,y) &= \mathtt{d}f(x,y)|_{\{0\} \times \mathbb{R}^k} \in \operatorname{End}(\mathbb{R}^k) \end{aligned} .$$

Ist $(x_0, y_0) \in U \times V$ mit $f(x_0, y_0) = 0$ und $d_2 f(x_0, y_0)$ invertierbar, so existieren offene Umgebungen U_1 von x_0 in U und V_1 von y_0 in V sowie eine C^m-Abbildung $\eta : U_1 \to V_1$ mit $\eta(x_0) = y_0$ und

$$\{(x,y) \in U_1 \times V_1 : f(x,y) = 0\} = \{\big(x, \eta(x)\big) : x \in U_1\}.$$

Insbesondere gilt $f\big(x, \eta(x)\big) = 0$ für alle $x \in U_1$.

Beweis. Die Abbildung

$$\varphi : U \times V \to \mathbb{R}^n \times \mathbb{R}^k, \quad (x,y) \mapsto \big(x, f(x,y)\big)$$

hat die Jacobimatrix

$$J_{(x,y)}(\varphi) = \begin{pmatrix} 1 & & 0 & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & 1 & 0 & \dots & 0 \\ \multicolumn{3}{c}{\left(\frac{\partial f_i}{\partial x_l}(x,y)\right)_{i,l}} & \multicolumn{3}{c}{\left(\frac{\partial f_i}{\partial y_j}(x,y)\right)_{i,j}} \end{pmatrix},$$

das heißt,

$$\det\big(J_{(x_0,y_0)}(\varphi)\big) = \det\left(\frac{\partial f_i}{\partial y_j}(x_0, y_0)\right) = \det((d_2 f)(x_0, y_0)) \neq 0.$$

Das Differential $d\varphi(x_0, y_0)$ ist also invertierbar. Nach dem Satz über die Umkehrfunktion existiert daher eine Umgebung W von (x_0, y_0), so dass

$$\varphi|_W : W \to \varphi(W) \subseteq \mathbb{R}^n \times \mathbb{R}^k$$

ein C^1-Diffeomorphismus ist. Die Umkehrfunktion $\psi := (\varphi|_W)^{-1} \colon \varphi(W) \to W$ hat dann die Gestalt

$$\psi(x,y) = \big(x, g(x,y)\big) \quad \text{mit einer } C^1\text{-Abb.} \quad g \colon \varphi(W) \to \mathbb{R}^k.$$

Wir definieren

$$\eta\colon \{x \in \mathbb{R}^n \colon (x,0) \in \varphi(W)\} \to \mathbb{R}^k, \quad \eta(x) := g(x,0).$$

Dann ist

$$\psi(x,0) = (x, g(x,0)) = (x, \eta(x))$$

und daher

$$(x,0) = \varphi(\psi(x,0)) = \varphi\big(x, \eta(x)\big) = \big(x, f(x, \eta(x))\big),$$

d.h. $f(x,\eta(x)) = 0$. Ist andererseits $f(x,y) = 0$ für $(x,y) \in W$, so ist $\varphi(x,y) = (x,0)$ und daher $(x,y) = \psi(x,0) = (x,\eta(x))$, also $y = \eta(x)$. Wir haben also gezeigt, dass

$$\{(x,y) \in W \colon f(x,y) = 0\} = \{(x,\eta(x)) \in W \colon (x,0) \in \varphi(W)\}. \tag{1.1}$$

Wir wählen jetzt offene Umgebungen U_1' von x_0 und V_1 von y_0 zunächst so klein, dass $U_1' \times V_1 \subseteq W$ gilt. Wegen der Stetigkeit von η finden wir eine offene Umgebung $U_1 \subseteq U_1'$ von x_0 mit $\eta(U_1) \subseteq V_1$. Da $\eta|_{U_1} \in C^m(U_1, \mathbb{R}^k)$ aus dem Satz über die Umkehrfunktion folgt, ist damit wegen (1.1) alles gezeigt. ∎

Bemerkung XII.1.2. (a) Die Voraussetzung von Theorem XII.1.1 lässt sich wie folgt nachprüfen. Wir schreiben hierzu $x = (x_1, \ldots, x_n, y_1, \ldots, y_k)$ für die Elemente von $\mathbb{R}^{n+k}$. Dann ist die Matrix der linearen Abbildung $d_2 f(x,y)$ gegeben durch

$$\frac{\partial f}{\partial y}(x,y) := \left(\frac{\partial f_i}{\partial y_j}(x,y)\right)_{i,j=1,\ldots,k}.$$

Man hat also die Invertierbarkeit dieser Matrix zu überprüfen.

Ist diese Bedingung an einer Stelle (x_0, y_0) erfüllt, so garantiert der Satz über implizite Funktionen lokal die Auflösbarkeit der Gleichung

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

auf $U_1 \times V_1$ nach y durch die Funktion η, denn für $(x,y) \in U_1 \times V_1$ mit $f(x,y) = 0$ gilt $y = \eta(x)$. Man kann dies auch so interpretieren, dass der Schnitt von $U_1 \times V_1$ mit der Nullstellenmenge von f der Graph der Funktion η ist. Die Bedingung

$$\det\Big(\frac{\partial f}{\partial y}(x_0, y_0)\Big) \neq 0$$

denkt man sich daher als eine hinreichende Bedingung für die lokale Auflösbarkeit der Gleichung $f(x,y) = 0$ nach y.

(b) Ist diese Bedingung in einem Punkt (x_0, y_0) mit $c := f(x_0, y_0) \neq 0$ erfüllt, so beschreibt sie eine hinreichende Bedinung für die Auflösbarkeit der Gleichung

$$f(x,y) = c$$

nach y, denn man kann statt f die Funktion $f - c$ betrachten.

(c) Eine notwendige Bedingung für die Anwendbarkeit von Theorem XII.1.1 ist, dass die lineare Abbildung $\mathtt{d}f(x_0, y_0) \colon \mathbb{R}^{n+k} \to \mathbb{R}^k$ den Rang k besitzt, d.h. surjektiv ist. Es kann allerdings durchaus passieren, dass dies der Fall ist, ohne dass $d_2 f(x_0, y_0)$ invertierbar ist.

Wir nehmen an, dass $\mathtt{d}f(x_0, y_0)$ den Rang k besitzt, also surjektiv ist und schreiben $x = (x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+k})$ für die Elemente von $\mathbb{R}^{n+k}$. Dann existieren verschiedene Indizes $r_1, \ldots, r_k \in \{1, \ldots, n+k\}$, so dass die Matrix

$$\left(\frac{\partial f_i}{\partial x_{r_j}}(x_0, y_0)\right)_{i,j=1,\ldots,k}$$

invertierbar ist. In diesem Fall schreiben wir

$$\mathbb{R}^{n+k} = E_1 \oplus E_2, \quad \text{mit} \quad E_1 = \operatorname{span}\{e_i \colon (\forall j) i \neq r_j\}, \quad E_2 = \operatorname{span}\{e_{r_j} \colon j = 1, \ldots, k\}.$$

Dann ist $E_1 \cong \mathbb{R}^n$ und $E_2 \cong \mathbb{R}^k$. Wir schreiben die Elemente von E_1 als $(z_1, \ldots, z_n)$ und die Elemente von E_2 als $(z_{n+1}, \ldots, z_{n+k})$ bzgl. einer Basis, die durch eine Permutation aus der kanonischen Basis entsteht, die die Menge der Indizes $\{r_1, \ldots, r_k\}$ auf $\{n+1, \ldots, n+k\}$ abbildet. So erhalten wir die Situation aus Theorem XII.1.1, denn nun ist

$$\det\left(\frac{\partial f_i}{\partial z_{n+j}}(x_0, y_0)\right)_{i,j=1,\ldots,k} \neq 0$$

und daher $\left(\frac{\partial f_i}{\partial z_{n+j}}(x_0, y_0)\right)_{i,j=1,\ldots,k}$ invertierbar.

Insbesondere spielt es keine Rolle, wie wir den $\mathbb{R}^{n+k}$ aufteilen. Wichtig ist, dass man in zwei Unterräume E_1 und E_2 aufteilt, so dass $df(x_0, y_0)|_{E_2} \colon E_2 \to \mathbb{R}^k$ invertierbar ist. ■

Beispiel XII.1.3. Wir betrachten die Funktion:

$$f \colon \mathbb{R}^3 \to \mathbb{R}^2, \quad f(x, y, z) = \begin{pmatrix} x^2 - y^2 \\ x^2 - z^2 \end{pmatrix}$$

mit der Jacobimatrix

$$J_{(x,y,z)}(f) = \begin{pmatrix} 2x & -2y & 0 \\ 2x & 0 & -2z \end{pmatrix}.$$

In diesem Fall ist $k = 2$ und $n = 1$. Die lokale Auflösbarkeitsbedingung nach (y, z) ist erfüllt, wenn

$$\det\left(\frac{\partial f}{\partial(y, z)}(x_0, y_0, z_0)\right) = \det\begin{pmatrix} -2y_0 & 0 \\ 0 & -2z_0 \end{pmatrix} = 4y_0 z_0 \neq 0$$

ist. In diesem Fall existieren offene Umgebungen U von x_0 und V von (y_0, z_0) in $\mathbb{R}^2$ sowie eine Funktion $\eta \colon U \to \mathbb{R}^2$ (eine Kurve), so dass für ein Tripel $(x, y, z) \in U \times V$ die Gleichung

$$f(x, y, z) = f(x_0, y_0, z_0)$$

genau dann erfüllt ist, wenn $(y,z) = \eta(x)$ ist, d.h., wenn der Punkt (x,y,z) auf dem Graphen der Kurve η liegt.

Ist obige Lösungsbedingung nicht erfüllt, d.h. $y_0 = 0$, $z_0 \neq 0$ und $x_0 \neq 0$, so ist

$$\det\Big(\frac{\partial f}{\partial(x,z)}(x_0,y_0,z_0)\Big) = \det\begin{pmatrix} 2x_0 & 0 \\ 2x_0 & -2z_0 \end{pmatrix} = -4x_0z_0 \neq 0$$

und wir erhalten entsprechend lokale Auflösbarkeit nach dem Variablenpaar (x,z). ■

Der Satz über implizite Funktionen ermöglicht uns, die Nullstellenmenge einer Funktion f lokal als Graph einer Funktion η zu beschreiben, wenn die Bedingung an das Differential erfüllt ist.

Betrachtet man beispielsweise den Einheitskreis, d.h. die Nullstellenmenge von

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = x^2 + y^2 - 1,$$

so ist die lokale Lösbarkeitsbedingung nach y für $U = V = \mathbb{R}$ gegeben durch

$$0 \neq \frac{\partial f}{\partial y}(x_0,y_0) = 2y_0,$$

d.h., in den Punkten mit $y_0 = 0$ kann man kein η finden. Für $y_0 > 0$ findet man $\eta(x) = \sqrt{1-x^2}$, und $\eta(x) = -\sqrt{1-x^2}$ für $y_0 < 0$. Um die Punkte $(1,0)$ und $(-1,0)$ ist der Kreisbogen nicht als Graph einer Funktion beschreibbar. Allerdings kann man hier die in Bemerkung XII.1.2(c) beschriebene Methode verwenden. Die Bedingung für die Auflösbarkeit der Gleichung nach x ist

$$0 \neq \frac{\partial f}{\partial x}(x_0,y_0) = 2x_0$$

und ist in den Punkten $(\pm 1, 0)$ erfüllt. Entsprechend erhalten wir Funktionen $\eta(y) = \sqrt{1-y^2}$ für $x_0 > 0$ und $\eta(y) = -\sqrt{1-y^2}$ für $x_0 < 0$. ■

Bemerkung XII.1.5. (a) Es ist instruktiv, sich klarzumachen, wie der Satz über implizite Funktionen im Kontext der Linearen Algebra aussieht. In diesem Fall ist $f\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$ eine lineare Abbildung. Gesucht ist eine Parametrisierung von

$$\ker f = \{v \in \mathbb{R}^{n+k} \colon f(v) = 0\}.$$

Die Bedingung aus Theorem XII.1.1 bedeutet, dass $f|_{\{0\}\times\mathbb{R}^k}$ surjektiv ist. Für die zugehörige Matrix bedeutet dies, dass die letzten k Spalten linear unabhängig sind. In diesem Fall ist $\eta\colon \mathbb{R}^n \to \mathbb{R}^k$ eine lineare Abbildung, deren Graph $\{(x,\eta(x)) \colon x \in \mathbb{R}^n\}$ der Kern von f ist. Die Vektoren $(e_j, \eta(e_j))$, $j = 1,\ldots,n$, bilden also eine Basis des Kerns. Im allgemeinen kann man nicht erwarten, dass die Lösungsbedingung erfüllt ist, wenn man nicht vorher die Koordinaten geeignet permutiert.

(b) Theorem XII.1.1 besagt letztendlich, dass die nichtlineare Gleichung

$$f(x,y) \;=\; 0$$

in einer Umgebung von (x_0,y_0) eindeutig und stetig differenzierbar nach y auflösbar ist, falls ihre lineare Approximation

$$0 = \mathtt{d}f(x_0,y_0)(u,v) = d_1f(x_0,y_0)(u) + d_2f(x_0,y_0)(v)$$

eindeutig nach v auflösbar ist (siehe die Diskussion vor Theorem XII.1.1). Wir beachten hierbei, dass die Bedingung der Invertierbarkeit von $d_2f(x_0,y_0)$ im allgemeinen nicht notwendig ist. So hat zum Beispiel für

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = (x-y)^2$$

die Gleichung $f(x,y)=0$ die eindeutige Lösung $y=\eta(x)=x$; aber für $x_0=y_0$ gilt

$$J_{(x_0,y_0)}(f) = \Big(2(x_0-y_0), 2(y_0-x_0)\Big) = (0,0). \qquad \blacksquare$$

Anwendung XII.1.6. (Algebraische Funktionen) Ein traditionelles Problem der Algebra ist das Auflösen von Polynomgleichungen. Wir schauen uns jetzt an, was uns der Satz über implizite Funktionen hierüber sagt.

Sei $f:\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ das Polynom

$$f(x,t) = t^n + \sum_{k=1}^n x_kt^{n-k} = t^n + x_1t^{n-1} + \ldots + x_{n-1}t + x_n.$$

Dann ist $f(x,t)=0$ genau dann, wenn t eine Nullstelle des Polynoms $f_x(t) := f(x,t) = t^n + \sum_{k=1}^n x_kt^{n-k}$ ist. Sei $t_0\in\mathbb{R}$ eine *einfache Nullstelle* von f_{x_0}, d.h. $f'_{x_0}(t)\neq 0$. Dann ist

$$\frac{\partial f}{\partial t}(x_0,t_0) = f'_{x_0}(t_0) \neq 0.$$

Dies entspricht der Lösbarkeitsbedingung nach t im Satz über implizite Funktionen XII.1.1. Es existieren also eine Umgebung U von $x_0=(x_{0,1},\ldots,x_{0,n})$ und eine beliebig oft differenzierbare Funktion $\eta\in C^\infty(U)$ mit $f\big(x,\eta(x)\big)=0$ für alle $x\in U$ und $\eta(x_0)=t_0$. Damit haben wir folgende bemerkenswerte Aussage bewiesen: Die einfachen Nullstellen eines Polynoms hängen lokal beliebig oft differenzierbar von seinen Koeffizienten ab.

Funktionen wie das oben beschriebene η nennt man *algebraische Funktionen*, weil sie Lösungen von polynomialen Gleichungen sind. Als Beispiel sei für $n=2$ das Polynom $f(x,y,t)=t^2+xt+y$ angeführt. Die Bedingung für die lokale Auflösbarkeit der Gleichung $f(x,y,t)=0$ nach t an der Stelle (x_0,y_0,t_0) mit $f(x_0,y_0,t_0)$ ist $2t_0+x_0\neq 0$. Wegen

$$f(x,y,t) = \big(t+\frac{x}{2}\big)^2 + y - \frac{x^2}{4}$$

bedeutet dies, dass t_0 keine zweifache Nullstelle des Polynoms $t^2 + x_0 t + y_0$ ist, d.h. $y_0 < \frac{x_0^2}{4}$. Wir nehmen zuerst

$$t_0 = -\frac{x_0}{2} + \sqrt{\frac{x_0^2}{4} - y_0}$$

an. Eine hierzu passende Funktion η ist dann gegeben durch

$$\eta(x, y) = -\frac{x}{2} + \sqrt{\frac{x^2}{4} - y}$$

auf der Menge $U_1 := \{(x, y)\colon \frac{x^2}{4} > y\}$. Gilt

$$t_0 = -\frac{x_0}{2} - \sqrt{\frac{x_0^2}{4} - y_0},$$

so erhalten wir

$$\eta(x, y) = -\frac{x}{2} - \sqrt{\frac{x^2}{4} - y}$$

auf der Menge $U_1 := \{(x, y)\colon \frac{x^2}{4} > y\}$. Man kann dies so interpretieren, dass über jedem Punkt (x, y) der offenen Menge U_1 genau zwei Lösungen der Gleichung $f(x, y, t) = 0$ liegen. Über den Randpunkten $(x, y) \in \partial U_1$ gilt $\frac{x^2}{4} = y$, und über diesen liegt nur eine Lösung. Die Lösungsmenge sieht also aus wie eine Fläche, die man an der Kurve ∂U_1 (eine Parabel) über sich selbst gefaltet hat. ■

Implizites Differenzieren

Bemerkung XII.1.7. (Die Ableitung von η) Die Voraussetzungen seien wie im Satz über implizite Funktionen. Wir schreiben wieder

$$\mathtt{d}f(x, y) = \Big(d_1 f(x, y), d_2 f(x, y)\Big),$$

mit $d_1 f(x, y) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ und $d_2 f(x, y) \in \mathrm{End}(\mathbb{R}^k)$. Aus $f\big(x, \eta(x)\big) = 0$ für $x \in U_1$ folgt dann mit der Kettenregel

$$0 = \mathtt{d}f\big(x, \eta(x)\big) \circ \big(\mathrm{id}, \mathtt{d}\eta(x)\big) = d_1 f\big(x, \eta(x)\big) + d_2 f\big(x, \eta(x)\big) \circ \mathtt{d}\eta(x)$$

Ist $d_2 f\big(x, \eta(x)\big)$ invertierbar, so ergibt sich hieraus

$$\mathtt{d}\eta(x) = -d_2 f\big(x, \eta(x)\big)^{-1} \circ d_1 f\big(x, \eta(x)\big).$$

Speziell ergibt sich für $n = k = 1$:

$$\eta'(x) = -\frac{\frac{\partial f}{\partial x}\big(x, \eta(x)\big)}{\frac{\partial f}{\partial y}\big(x, \eta(x)\big)}$$

für $f\big(x, \eta(x)\big) = 0$, falls $\frac{\partial f}{\partial y}\big(x, \eta(x)\big) \neq 0$ ist. ■

Wir betrachten die Funktion $\eta\colon U_1 \to V_1$ als eine lokale Parametrisierung der Lösungsmenge der Gleichung $f(x, y) = 0$. Letztere wird somit lokal als Graph der Funktion η dargestellt. Der folgende Begriff beschreibt allgemein Teilmengen des $\mathbb{R}^n$, die sich so beschreiben lassen, unabhängig davon, ob sie Lösungsmenge einer Gleichung sind oder nicht.

Untermannigfaltigkeiten des $\mathbb{R}^n$

Definition XII.1.8. Eine Teilmenge $M \subseteq \mathbb{R}^n$ heißt k-dimensionale C^m-*Untermannigfaltigkeit*, wenn folgende Bedingung erfüllt ist. Für jeden Punkt $p \in M$ existiert eine offene Umgebung $U \subseteq \mathbb{R}^n$, eine offene Teilmenge $U' \subseteq \mathbb{R}^n$ und ein C^m-Diffeomorphismus

$$\varphi : U \to U' \subseteq \mathbb{R}^n \quad \text{mit} \quad \varphi(U \cap M) = U' \cap \big(\mathbb{R}^k \times \{0\}\big).$$

Eine solche Abbildung heißt *Umgebungskarte von M*. Eine Familie $(\varphi_j)_{j\in J}$ von Umgebungskarten $\varphi_j : U_j \to U_j' \cap \big(\mathbb{R}^k \times \{0\}\big)$ von M heißt *Umgebungsatlas von M*, wenn $M \subseteq \bigcup_{j\in J} U_j$ ist. ■

Eine Untermannigfaltigkeit ist also eine Menge, die in geeigneten krummlinigen Koordinaten (beschrieben durch φ) lokal wie $\mathbb{R}^k$ in $\mathbb{R}^n$ aussieht.

Beispiel XII.1.9. (a) (Funktionsgraphen) Sei $n = m + k$, $V \subseteq \mathbb{R}^k$ offen und $f\colon V \to \mathbb{R}^m$ eine stetig differenzierbare Funktion. Wir zeigen, dass

$$M := \Gamma(f) = \{(x, f(x))\colon x \in V\}$$

eine k-dimensionale C^1-Untermannigfaltigkeit von $\mathbb{R}^{m+k}$ ist.

Hierzu sei $U := V \times \mathbb{R}^m$. Dies ist eine offene Menge, die Umgebung aller Punkte in M ist. Wir betrachten die Abbildung

$$\varphi\colon U \to U \subseteq \mathbb{R}^n, \qquad \varphi(x, y) = \big(x, y - f(x)\big).$$

Dann ist φ ein C^1-Diffeomorphismus der Menge U mit der inversen Abbildung $\varphi^{-1}(x, y) = \big(x, y + f(x)\big)$. Weiter gilt

$$\varphi(U \cap M) = \varphi(M) = V \times \{0\} = U \cap (\mathbb{R}^k \times \{0\}).$$

Die Abbildunge φ liefert also einen einelementigen Umgebungsatlas von M.
(b) (Die n-Sphäre) Wir betrachten die n-*Sphäre*

$$\mathbb{S}^n := \big\{x \in \mathbb{R}^{n+1}\colon \|x\|_2 = 1\big\} = \Big\{x \in \mathbb{R}^{n+1}\colon \sum_{j=1}^n x_j^2 = 1\Big\}.$$

Wir zeigen, dass $\mathbb{S}^n$ eine n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^{n+1}$ ist. Für jeden Index $j \in \{1, \ldots, n+1\}$ betrachten wir die offenen Mengen

$$U_j^{\pm} = \Big\{x \in \mathbb{R}^{n+1}\colon \sum_{i\neq j} x_i^2 < 1, \pm x_j > 0\Big\}.$$

Jede der Mengen $U_j^{\pm}$ ist offen, und diese Mengen überdecken $\mathbb{S}^n$. Wir betrachten die Abbildungen

$$\varphi_j^{\pm} : U_j^{\pm} \to \mathbb{R}^{n+1}, \quad \varphi_j^{\pm}(x) = \Big(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n, x_j \mp \Big(1 - \sum_{i \neq j} x_i^2\Big)^{\frac{1}{2}}\Big).$$

Dann ist $\varphi_j^{\pm}$ ein Diffeomorphismus auf $\varphi_j^{\pm}(U_j^{\pm})$ (Nachweis als Übung!), und es gilt

$$\varphi_j^{\pm}(U_j^{\pm} \cap M) = \varphi_j^{\pm}(U_j^{\pm}) \cap (\mathbb{R}^n \times \{0\}).$$

Hiermit sieht man, dass die $2(n+1)$ Umgebungskarten $\varphi_j^{\pm}$, $j = 1, \ldots, n+1$ einen Umgebungsatlas von $\mathbb{S}^n$ bilden. ■

Aufgabe XII.1. Sei $M \subseteq \mathbb{R}^n$ eine Teilmenge. Zeigen Sie: Die Eigenschaft von M, eine k-dimensionale C^m-Untermannigfaltigkeit zu sein, ist in dem folgenden Sinne lokal. Die Teilmenge M ist genau dann eine k-dimensionale C^m-Untermannigfaltigkeit von $\mathbb{R}^n$, wenn für jeden Punkt $p \in M$ eine offene Umgebung U existiert, so dass $U \cap M$ eine k-dimensionale C^m-Untermannigfaltigkeit ist. ■

Bemerkung XII.1.10. (a) Erfüllt $f : U \times V \to \mathbb{R}^k$ ($U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^k$) in (x_0, y_0) die Voraussetzung des Satzes über implizite Funktionen XII.1.1, d.h., ist $d_2 f(x_0, y_0)$ invertierbar, so besagt Definition XII.1.8, dass die Menge

$$M := \{(x, y) \in U_1 \times V_1 : f(x, y) = 0\}$$

eine k-dimensionale Untermannigfaltigkeit von $\mathbb{R}^{n+k}$ ist, denn sie ist ein Funktionsgraph. Die Abbildung

$$\varphi \ : \ U_1 \times V_1 \to \mathbb{R}^{n+k}$$

aus dem Beweis ist eine Umgebungskarte. Möchte man zeigen, dass die gesamte Lösungsmenge $\widetilde{M} := \{(x, y) \in U \times V : f(x, y) = 0\}$ eine Untermannigfaltigkeit ist, so hat man allerdings für jeden Punkt $(x_0, y_0) \in \widetilde{M}$ eine Umgebungskarte zu finden (vgl. Beispiel XII.1.4).
(b) Ist speziell $f : \mathbb{R}^n \to \mathbb{R}^k$ eine surjektive lineare Abbildung, so ist $M := \ker f$ eine $(n-k)$-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$ (vgl. Rangsatz: $\dim \ker f = \dim \mathbb{R}^n - \dim\big(\operatorname{im} f\big) = n - \dim\big(\operatorname{im} f\big)$). ■

Definition XII.1.11. Ist $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^k$ eine differenzierbare Funktion, so heißt

$$\operatorname{rg}_u(f) := \operatorname{rg}\big(\mathtt{d}f(u)\big)$$

der *Rang* von f in u. Der Punkt u heißt *kritischer Punkt*, wenn $\operatorname{rg}_u(f) < k$ ist. In diesem Fall heißt $f(u) \in \mathbb{R}^k$ *kritischer Wert*. Ein Punkt $y \in \mathbb{R}^k$ heißt *regulärer Wert*, wenn y kein kritischer Wert ist, d.h., wenn die Menge $f^{-1}(y)$ keine kritischen Punkte enthält. Beachte, dass alle Punkte $y \notin f(U)$ reguläre Werte sind, weil $f^{-1}(y) = \emptyset$ ist. ■

Den folgenden Satz kann man als eine globale Version des Satz über implizite Funktionen verstehen, denn er gibt nicht nur Auskunft über die lokale Struktur der Lösungsmenge einer Gleichung, sondern über ihre globale Struktur.

RANGSATZ

Satz XII.1.12. *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^k$ eine C^m-Abbildung. Ist $w \in f(U) \subseteq \mathbb{R}^k$ ein regulärer Wert von f, so ist das Urbild $f^{-1}(w)$ eine $(n-k)$-dimensionale C^m-Untermannigfaltigkeit von $\mathbb{R}^n$.*

Für $n = 2$ und $k = 1$ heißen die Mengen $f^{-1}(w)$ Höhenlinien von f. Im allgemeinen spricht man von *Niveaumengen* oder *Niveauflächen*.

Beweis. Sei $u \in f^{-1}(w)$. Wegen der Regularität von u ist $\mathrm{rg}_u(f) = k$. Also ist die Abbildung $g : U \to \mathbb{R}^k$, $x \mapsto f(x) - w$ im Punkt u regulär, d.h., es gilt $\mathrm{rg}\left(\frac{\partial g_i}{\partial x_j}(u)\right)_{\substack{i=1\ldots,k\\ j=1,\ldots,n}} = k$. Nach geeigneter Umnumerierung der Koordinaten dürfen wir o.B.d.A. annehmen, dass die ersten k Spalten der Matrix $J_u(g)$ linear unabhängig sind, d.h., es ist $\det\left(\frac{\partial g_i}{\partial x_j}\right)_{i,j=1\ldots,k} \neq 0$ (vgl. die Diskussion in Bemerkung XII.1.2). Sei nun

$$\varphi : U \to \mathbb{R}^n, \quad x \mapsto \big(g_1(x), \ldots, g_k(x), x_{k+1}, \ldots, x_n\big).$$

Dann ist

$$J_u(\varphi) = \begin{pmatrix} & \left(\frac{\partial g_i}{\partial x_j}\right)_{\substack{i=1\ldots,k\\ j=1,\ldots,n}} & & & & \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ & \ddots & & & \ddots & \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix}$$

eine $(n \times n)$-Matrix mit $\det\big(J_u(\varphi)\big) = \det\left(\frac{\partial g_i}{\partial x_j}\right)_{i,j=1\ldots,k} \neq 0$. Nach dem Satz über die Umkehrfunktion gibt es also eine offene Umgebung V von u, so dass $\varphi|_V : V \to \varphi(V)$ ein Diffeomorphismus ist. Dann ist

$$\varphi\big(f^{-1}(w) \cap V\big) = \varphi\big(g^{-1}(0) \cap V\big) = \big(\{0\} \times \mathbb{R}^{n-k}\big) \cap \varphi(V).$$

Also ist φ eine Umgebungskarte von $f^{-1}(w)$ um u, und $\dim\big(f^{-1}(w)\big) = n-k$. ■

Beispiel XII.1.13. (a) Sei $A \in \mathrm{GL}_n(\mathbb{R})$ eine invertierbare symmetrische Matrix und $f : \mathbb{R}^n \to \mathbb{R}$ gegeben durch $f(x) = \langle Ax, x\rangle = x^\top Ax$. Nach der Produktregel ist dann

$$\mathtt{d}f(x)(h) = 2\langle Ax, h\rangle,$$

also $J_x(f) = 2x^\top A$ (Nachweis als Übung). Ein Punkt $x \in \mathbb{R}^n$ ist genau dann ein kritischer Punkt, wenn $Ax = 0$ ist. Da A invertierbar ist, gilt dies genau dann, wenn $x = 0$ ist. Der einzige kritische Wert ist also $t = 0$. Somit ist die Menge $f^{-1}(t)$ für alle $t \neq 0$ eine Untermannigfaltigkeit von $\mathbb{R}^n$. Für $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_i > 0$, ergeben sich Ellipsoide:

$$f^{-1}(t) = \Big\{x \in \mathbb{R}^n : \sum_{j=1}^n \lambda_j x_j^2 = t\Big\}.$$

Für $A = \mathbf{1}$ erhalten wir einen neuen Beweis dafür, dass die Sphäre

$$\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$$

eine $(n-1)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ ist (Beispiel XII.1.9(b)).
(b) Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^4 - y^4$. Dann ist $\nabla f(x, y) = (4x^3, -4y^3)$. Also ist (x, y) genau dann kritischer Punkt, wenn $(x, y) = (0, 0)$ ist, und der einzige kritische Wert ist $t = 0$. Für $t \neq 0$ sind also alle Höhenlinien von f Untermannigfaltigkeiten von $\mathbb{R}^2$. Für $t = 0$ ist die zugehörige Höhenlinie keine Untermannigfaltigkeit, da sie aus zwei Geraden besteht, die sich im Nullpunkt schneiden.
(c) Ist $f : \mathbb{R}^n \to \mathbb{R}^k$ eine lineare Abbildung, so ist ein Punkt $w \in f(\mathbb{R}^n) \subseteq \mathbb{R}^k$ genau dann ein regulärer Wert, wenn $\mathrm{rg}(f) = k$ ist (beachte, dass $f = \mathtt{d}f(u)$ für alle $u \in \mathbb{R}^n$ gilt, so dass $\mathrm{rg}_u(f) = \mathrm{rg}\big(\mathtt{d}f(u)\big) = \mathrm{rg}(f)$ von u unabhängig ist). Dann ist für jeden Punkt $w \in \mathbb{R}^k$ die Menge $f^{-1}(w)$ ein affiner Unterraum von $\mathbb{R}^n$: Ist $f(x) = w$, so ist $f^{-1}(w) = x + \ker f$. Dies sind spezielle Untermannigfaltigkeiten der Dimension $\dim(\ker f) = n - k$. ∎

Aufgabe XII.1. Sei $1 < p < \infty$. Wir betrachten die Einheitssphäre

$$M := \{x \in \mathbb{R}^n : \|x\|_p = 1\} = \{x \in \mathbb{R}^n : |x_1|^p + \ldots + |x_n|^p = 1\}.$$

Für welche p ist die Funktion $f(x) := |x_1|^p + \ldots + |x_n|^p$ auf $\mathbb{R}^n$ stetig differenzierbar? Wie hoch ist die Differenzierbarkeitsordnung? Diskutieren sie zuerst die Funktion $\mathbb{R} \to \mathbb{R}, x \mapsto |x|^p$. Für welche k ist M eine C^k-Untermannigfaltigkeit? ∎

XII.2. Extrema mit Nebenbedingungen

Nachdem wir in Abschnitt X.4 Extrema von Funktionen studiert haben, die auf offenen Mengen $U \subseteq \mathbb{R}^n$ definiert sind, wenden wir uns nun einer Situation zu, die in praktischen Problemen viel häufiger zu finden ist. Wir werden Extrema mit Nebenbedingungen studieren, d.h. Extrema von Funktionen auf Untermannigfaltigkeiten $M \subseteq \mathbb{R}^n$. Der wesentliche Punkt hierbei ist, dass man dies nicht direkt durch eine Parametrisierung der Untermannigfaltigkeit auf die Situation von Abschnitt X.4 zurückführen möchte, da dies im allgemeinen recht kompliziert sein kann. Vielmehr möchte man direkter notwendige Bedingungen ableiten, die sich mit Daten formulieren lassen, die sich aus der Funktion g ergeben, die die Untermannigfaltigkeit als Niveaumenge $g^{-1}(0)$ beschreibt.

Definition XII.2.1. Sei $M \subseteq \mathbb{R}^n$ eine k-dimensionale Untermannigfaltigkeit und $p \in M$. Ein Vektor $v \in \mathbb{R}^n$ heißt *Tangentialvektor an M in p*, wenn eine stetig differenzierbare Kurve $\gamma :]-\varepsilon, \varepsilon[\to M \subseteq \mathbb{R}^n$ mit $\gamma(0) = p$ und $\dot\gamma(0) = v$ existiert. Die Menge $T_p(M)$ aller Tangentialvektoren von M in p heißt *Tangentialraum von M in p*. Die Menge $p + T_p(M)$ heißt *Tangente an p*.■

Satz XII.2.2. *Der Tangentialraum $T_p(M)$ ist ein k-dimensionaler Untervektorraum von $\mathbb{R}^n$. Beschreiben kann man ihn wie folgt:*

(a) *Ist $\varphi : U \to U' \subseteq \mathbb{R}^n$ eine Umgebungskarte um p mit $\varphi(p) = 0$ und $\varphi(U \cap M) = \varphi(U) \cap \big(\mathbb{R}^k \times \{0\}\big)$, so ist*

$$T_p(M) = \big(\mathtt{d}\varphi(p)\big)^{-1}\big(\mathbb{R}^k \times \{0\}\big) = \mathtt{d}\big(\varphi^{-1}\big)(0)\big(\mathbb{R}^k \times \{0\}\big).$$

(b) *Ist $U \subseteq \mathbb{R}^n$ offen, $g : U \to \mathbb{R}^{n-k}$ eine stetig differenzierbare Funktion, $w \in \mathbb{R}^{n-k}$ ein regulärer Wert von g und*

$$M := g^{-1}(w) = \{x \in U : g(x) = w\} \neq \emptyset,$$

so ist

$$T_p(M) = \ker\big(\mathtt{d}g(p)\big).$$

Beweis. Es ist klar, dass aus (a) und (b) jeweils folgt, dass $T_p(M)$ ein Vektorraum ist.

Zuerst führen wir (b) auf (a) zurück. In der Situation von (b) sei $\varphi(x) = \big(x_1, \ldots, x_k, g(x)\big)$ wie im Beweis von Theorem XII.1.12, wobei die Matrix

$$\left(\frac{\partial g_i}{\partial x_j}(p)\right)_{i=1,\ldots,n-k, j=k+1,\ldots,n}$$

invertierbar ist. Dann liefert φ eine Umgebungskarte um p in M und

$$J_p(\varphi) = \begin{pmatrix} \mathbf{1}_k & \mathbf{0} \\ & \\ \multicolumn{2}{c}{\left(\frac{\partial g_i}{\partial x_j}(p)\right)_{i,j}} \end{pmatrix} = \begin{pmatrix} \mathbf{1}_k & \mathbf{0} \\ \multicolumn{2}{c}{J_p(g)} \end{pmatrix}$$

wobei $\mathbf{1}_k$ die $(k \times k)$-Einheitsmatrix ist. Dann ist

$$\big(\mathtt{d}\varphi(p)\big)^{-1}\big(\mathbb{R}^k \times \{0\}\big) = \ker \mathtt{d}g(p).$$

Nun beweisen wir (a). Ist $\gamma :]-\varepsilon, \varepsilon[\to M \subseteq \mathbb{R}^n$ eine stetig differenzierbare Kurve mit $\gamma(0) = p$ und $\dot\gamma(0) = v$, so ist

$$\big(\varphi \circ \gamma\big)'(0) = \mathtt{d}\varphi\big(\gamma(0)\big)\big(\dot\gamma(0)\big) \in \mathbb{R}^k \times \{0\},$$

d.h., $v = \dot\gamma(0) \in \big(\mathtt{d}\varphi(p)\big)^{-1}\big(\mathbb{R}^k \times \{0\}\big)$. Ist umgekehrt

$$v \in \big(\mathtt{d}\varphi(p)\big)^{-1}\big(\mathbb{R}^k \times \{0\}\big)$$

und ist $\varepsilon > 0$, so dass $t \cdot \mathtt{d}\varphi(p)(v) \in \varphi(U)$ für alle t mit $|t| < \varepsilon$ gilt, so definieren wir

$$\gamma :]-\varepsilon, \varepsilon[\to M, \ t \mapsto \varphi^{-1}\big(t \cdot \mathtt{d}\varphi(p)(v)\big).$$

Dann ist γ stetig differenzierbar mit $\gamma(0) = \varphi^{-1}(0) = p$ und

$$\dot\gamma(0) = \mathtt{d}\big(\varphi^{-1}\big)(0)\big(\mathtt{d}\varphi(p)(v)\big) = \Big(\big(\mathtt{d}\varphi(p)\big)^{-1} \circ \mathtt{d}\varphi(p)\Big)(v) = v.$$

Es folgt $T_p(M) = \big(\mathtt{d}\varphi(p)\big)^{-1}\big(\mathbb{R}^k \times \{0\}\big)$. ■

Wir fassen zusammen:
(a) Lokal sieht eine Untermannigfaltigkeit M aus wie eine verbogene Kopie einer offenen Teilmenge von $\mathbb{R}^k$ im $\mathbb{R}^n$. Der Tangentialraum $T_p(\mathbb{R}^k)$ an $\mathbb{R}^k$ ist in allen Punkten p gleich $\mathbb{R}^k$ selbst; entsprechend überträgt sich der Tangentialraum durch das Differential der Parametrisierungsabbildung auf die Mannigfaltigkeit.
(b) Wird M durch die Gleichung $g(x) = 0$ beschrieben, so wird der Tangentialraum $T_p(M)$ durch die Gleichung $\mathtt{d}g(p)(v) = 0$ beschrieben. Ist speziell $g : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$, so ist $M = g^{-1}(w)$ eine Niveaufläche der Funktion g. Ist w ein regulärer Wert, so ist

$$T_p(M) = \ker \mathtt{d}g(p) = \{v \in \mathbb{R}^n : \langle \nabla g(p), w\rangle = 0\}.$$

Der Gradient ist also orthogonal zu den Niveauflächen bzw. dem Tangentialraum.

Beispiel XII.2.3. (a) Sei $M = \mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ die Einheitssphäre. Man kann sie als Nullstellenmenge der Funktion $g(x) = \|x\|^2 - 1 = \langle x, x\rangle - 1$ beschreiben, d.h., es ist $M = g^{-1}(0)$. Null ist ein regulärer Wert. Wegen $\nabla g(x) = 2x$ ist

$$T_p(M) = \{v \in \mathbb{R}^n : \langle v, p\rangle = 0\},$$

und die Tangente in p ist gleich

$$p + T_p(M) = \{v \in \mathbb{R}^n : \langle v, p\rangle = 1\}.$$

(b) Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und $f : U \to \mathbb{R}^k$ eine C^m-Abbildung. Dann ist der Funktionsgraph

$$M := \Gamma(f) := \big\{\big(x, f(x)\big) : x \in U\big\} \subseteq U \times \mathbb{R}^k$$

eine n-dimensionale Untermannigfaltigkeit, denn für

$$g : U \times \mathbb{R}^k \to \mathbb{R}^k, \quad g(x,y) := y - f(x)$$

ist $M = g^{-1}(0)$ und

$$J_{(x,y)}(g) = \left.\left(\begin{array}{c|c} \underbrace{\quad -J_x(f) \quad}_{n} & \underbrace{\begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix}}_{k} \end{array}\right)\right\} k,$$

d.h., für alle $p \in M$ gilt die Beziehung $\mathrm{rg}_p(g) = k$; insbesondere ist 0 ein regulärer Wert. Um den Tangentialraum zu berechnen, verwenden wir, dass genau dann $\mathtt{d}g(x,y)(v,w) = -\mathtt{d}f(x)(v) + w = 0$ gilt, wenn $w = \mathtt{d}f(x)(v)$ ist. Damit erhalten wir

$$T_{(x,f(x))}(M) = \big\{\big(v, \mathtt{d}f(x)(v)\big) : v \in \mathbb{R}^n\big\} = \Gamma\big(\mathtt{d}f(x)\big);$$

die Tangente als „affine Approximation" von M ist dann

$$\big(x, f(x)\big) + T_{(x,f(x))}(M) = \big\{\big(x+v, f(x) + \mathtt{d}f(x)(v)\big) : v \in \mathbb{R}^n\big\}.$$

Passenderweise ist sie also der Graph der affinen Funktion

$$T_x^1(f)(v) = f(x) + \mathtt{d}f(x)(v). \qquad \blacksquare$$

Definition XII.2.4. Sei M eine C^1-Untermannigfaltigkeit der offenen Teilmenge $U \subseteq \mathbb{R}^n$ und $f : U \to \mathbb{R}$ eine stetig differenzierbare Funktion. Dann heißt $p \in M$ *kritischer Punkt von* $f|_M$, wenn

$$\mathtt{d}f(p)|_{T_p(M)} = 0$$

gilt. Er heißt dann ein *kritischer Punkt unter der Nebenbedingung* M. ■

Wir fügen an dieser Stelle einen Satz ein, der eigentlich in die lineare Algebra gehört; wir brauchen ihn im Beweis des nachfolgenden Satzes.

Satz XII.2.5. *Ist* V *ein Vektorraum, und sind* $\alpha, \beta_1, \ldots, \beta_n : V \to \mathbb{R}$ *lineare Abbildungen, so ist die Bedingung*

$$\ker\alpha \supseteq \bigcap_{j=1}^n \ker\beta_j \tag{2.1}$$

äquivalent zur Existenz von $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *mit*

$$\alpha = \lambda_1\beta_1 + \ldots + \lambda_n\beta_n. \tag{2.2}$$

Beweis. Die Richtung (2.2) ⇒ (2.1) ist trivial. Zum Beweis der Richtung (2.1) ⇒ (2.2) betrachten wir die lineare Abbildung

$$\varphi : V \to \mathbb{R}^n,\ v \mapsto \big(\beta_1(v), \ldots, \beta_n(v)\big).$$

Dann ist $\ker\varphi = \bigcap_{i=1}^n \ker\beta_i$. Wegen $\alpha(\ker\varphi) = \{0\}$ wird durch

$$\widetilde{\alpha} : \varphi(V) \to \mathbb{R}, \quad \varphi(v) \mapsto \alpha(v)$$

eine lineare Abbildung definiert, die sich zu einer linearen Abbildung $\widetilde{\alpha} : \mathbb{R}^n \to \mathbb{R}$ mit $\widetilde{\alpha} \circ \varphi = \alpha$ fortsetzen lässt. Nun existieren Zahlen $\lambda_1, \ldots, \lambda_n$ mit $\widetilde{\alpha}(x) = \sum_{i=1}^n \lambda_i x_i$ für alle $x \in \mathbb{R}^n$. Dann gilt $\alpha(v) = \widetilde{\alpha}\big(\varphi(v)\big) = \sum_{i=1}^n \lambda_i\big(\beta_i(v)\big)$ für alle $v \in V$, d.h.,

$$\alpha = \lambda_1\beta_1 + \ldots + \lambda_n\beta_n. \qquad \blacksquare$$

Satz XII.2.6. *Sei $U \subseteq \mathbb{R}^{m+n}$ eine offene Teilmenge, und die m-dimensonale C^1-Untermannigfaltigkeit $M \subseteq U$ sei durch*

$$M := \{x \in U : g(x) = 0\}$$

gegeben, wobei $g : U \to \mathbb{R}^n$ eine stetig differenzierbare Funktion und 0 ein regulärer Wert sei.

(a) *Dann ist $p \in M$ genau dann ein kritischer Punkt von $f \in C^1(U)$ unter der Nebenbedingung $g = 0$, wenn Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ so existieren, dass*

$$\mathtt{d}f(p) = \sum_{j=1}^{n} \lambda_j \mathtt{d}g_j(p)$$

gilt, d.h. $\mathtt{d}f(p)$ ist von den $\mathtt{d}g_j(p)$ linear abhängig.

$\boxed{\text{NOTWENDIGE BEDINGUNG FÜR EXTREMA}}$

(b) *Hat f in $p \in M$ ein lokales Extremum unter der Nebenbedingung $g = 0$, so ist p ein kritischer Punkt von $f|_M$. Dass p ein lokales Maximum ist, bedeutet in diesem Kontext, dass ein $\delta > 0$ so existiert, dass $f(x) \leq f(p)$ für alle $x \in M$ mit $\|x - p\| \leq \delta$ gilt.*

Beweis. (b) Sei p ein lokales Extremum von $f|_M$. Es ist zu zeigen, dass $\mathtt{d}f(p)|_{T_p(M)} = 0$ gilt, dass also $T_p(M) \subseteq \ker \mathtt{d}f(p)$ ist. Sei $v \in T_p(M)$. Dann existiert eine Kurve $\gamma :]-\varepsilon, \varepsilon[\to M$ mit $\gamma(0) = p$ und $\dot\gamma(0) = v$. Nun ist 0 ein lokales Extremum der Funktion $]-\varepsilon, \varepsilon[\to \mathbb{R}, t \mapsto f\big(\gamma(t)\big)$, also gilt

$$0 = \frac{d}{dt}\Big|_{t=0} f\big(\gamma(0)\big) = \mathtt{d}f\big(\gamma(0)\big)\big(\dot\gamma(0)\big) = \mathtt{d}f(p)(v).$$

Damit ist $T_p(M) \subseteq \ker \mathtt{d}f(p)$, also p ein kritischer Punkt von $f|_M$.

(a) Nach Satz XII.2.2 ist

$$T_p(M) = \ker \mathtt{d}g(p) = \bigcap_{i=1}^{n} \ker \mathtt{d}g_i(p) \qquad \text{für} \quad g = (g_1, \ldots, g_n).$$

Die Bedingung, dass p kritischer Punkt ist, ist also zu $\ker \mathtt{d}f(p) \supseteq \bigcap_{i=1}^{n} \ker \mathtt{d}g_i(p)$ äquivalent. Nach Satz XII.2.5 bedeutet dies, dass Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit

$$\mathtt{d}f(p) = \sum_{i=1}^{n} \lambda_i \mathtt{d}g_i(p)$$

existieren. ■

$\boxed{\text{METHODE DER LAGRANGE-MULTIPLIKATOREN}}$

Die Zahlen $\lambda_1, \ldots, \lambda_m$ in Satz XII.2.6 heißen *Lagrange-Multiplikatoren.* Will man die lokalen Extrema von $f|_M$ bestimmen, so hat man also die $2n + m$ Gleichungen

$$g(x) = 0 \quad \text{und} \quad \mathtt{d}f(x) - \sum_{i=1}^{n} \lambda_i \mathtt{d}g_i(p) = 0$$

zu lösen. In ihnen kommen $2n + m$ Unbekannte vor, nämlich $x_1, \ldots, x_{m+n}$ und $\lambda_1, \ldots, \lambda_n$.

Beispiel XII.2.7. (a) Sei $M = \mathbb{S}^{k-1} \subseteq \mathbb{R}^k$ die Einheitskugel, also die Nullstellenmenge von $g(x) = \|x\|^2 - 1$, und $f(x) = \langle Ax, x \rangle$ für eine symmetrische Matrix A. Wir suchen die kritischen Punkte von $f\,|_M$. Hier ist $n = 1$ und $m = k - 1$, also $m + 2n = k + 1$. Es ist $\nabla f(x) = 2Ax$ und $\nabla g(x) = 2x$. Die Gleichungen, die wir lösen müssen, sind also $g(x) = 0$ und $\nabla f(x) - \lambda \cdot \nabla g(x) = 0$, d.h.

$$\|x\|_2 = 1 \quad \text{und} \quad Ax - \lambda x = 0.$$

Ein Punkt $x \in M$ ist also genau dann kritischer Punkt von $f\,|_M$, wenn er Eigenvektor von A zum Eigenwert λ ist. Wegen $f(x) = \langle Ax, x \rangle$ ergibt sich als zugehöriger Funktionswert in $x \in \mathbb{S}^{k-1}$:

$$f(x) = \langle Ax, x \rangle = \lambda \langle x, x \rangle = \lambda.$$

Folglich ist x genau dann ein Minimum von $f|_M$, wenn x Eigenvektor zum minimalen Eigenwert von A ist und ein Maximum genau dann, wenn x Eigenvektor zum maximalen Eigenwert ist.

Man beachte: Aus dem Satz vom Maximum folgt die Existenz eines Maximums von f auf der kompakten Menge $M = \mathbb{S}^{k-1}$, was bedeutet, dass A mindestens einen reellen Eigenwert hat! Induktiv schließt man hieraus leicht, dass die nach Voraussetzung symmetrische Matrix A reell diagonalisierbar ist (Lineare Algebra II).

(b) Sei $M = \{x \in U : g(x) = 0\} \subseteq \mathbb{R}^n$ eine Untermannigfaltigkeit, wobei 0 ein regulärer Wert der stetig differenzierbaren Funktion $g: U \to \mathbb{R}$ ist. Weiter sei $p \in \mathbb{R}^n \setminus M$. Wir suchen in M einen Punkt minimalen Abstands von p. Wir betrachten dazu die Funktion

$$f: U \to \mathbb{R}, \quad f(x) = \|x - p\|^2 = \sum_{j=1}^{n} (x_j - p_j)^2.$$

Wir nehmen an, dass $x \in M$ ein lokales Minimum der Funktion f unter der Nebenbedingung $g = 0$ ist. Nun ist

$$\nabla f(x) = 2(x - p),$$

so dass genau dann ein $\lambda \in \mathbb{R}$ mit

$$\mathtt{d}f(x) + \lambda \mathtt{d}g(x) = 0$$

existiert, wenn $\mathtt{d}f(x)$ kollinear zu $\mathtt{d}g(x)$ ist. Da der Tangentialraum $T_x(M)$ mit $\ker \mathtt{d}g(x)$ übereinstimmt, bedeutet dies, dass

$$(x - p) \perp T_x(M)$$

gilt. Wir finden also die notwendige Bedingung, dass die Verbindungsstrecke von p und x senkrecht zu $T_p(M)$ ist, d.h., diese Verbindungsstrecke trifft orthogonal auf die Untermannigfaltigkeit M. ■

Aufgabe XII.2. Gegeben sei die Untermannigfaltigkeit

$$M = \{(x, y, z) \in \mathbb{R}^3 : z = y^2 + x^2\}$$

und $p = (0, 0, t)$. Bestimmen sie (in Abhängigkeit von t), alle Punkte auf M, die von p minimalen Abstand haben. ■