

18

Analysis im \mathbb{R}^n

Wir betrachten nun Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $m \geq 2$, die man auch als vektorwertige Funktionen bezeichnet. Zunächst formulieren wir den lokalen Umkehrsatz für Abbildungen des \mathbb{R}^n in sich selbst. Wir beweisen ihn – recht ausführlich – zuerst innerhalb der Kategorie der Lipschitzstetigen Abbildungen. Höhere Regularität betrachten wir erst danach und bereitet keine neuen Probleme.

Eine unmittelbare Folge des Umkehrsatzes ist der fundamentale Satz über implizite Funktionen. Er bildet die Grundlage für die Definition gleichungsdefinierter Mannigfaltigkeiten, und daran anknüpfend die Diskussion von Extrema mit Nebenbedingungen und der Methode der Lagrangemultiplikatoren.

18.1

Umkehrabbildungen

Wir betrachten vektorwertige Abbildungen

$$\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto u = \varphi(x).$$

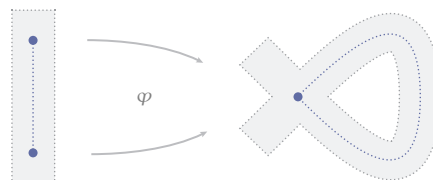
Zuerst wollen wir die Frage studieren, unter welchen Bedingungen eine solche Abbildung *umkehrbar*, also die Gleichung $u = \varphi(x)$ nach x eindeutig auflösbar ist. Mit anderen Worten: Wann können wir das System von m Gleichungen in n Variablen,

$$\begin{aligned} u_1 &= \varphi_1(x_1, \dots, x_n), \\ u_2 &= \varphi_2(x_1, \dots, x_n), \\ &\vdots \\ u_m &= \varphi_m(x_1, \dots, x_n), \end{aligned}$$

eindeutig nach x_1, \dots, x_n auflösen?

Abb 1

Lokal, aber nicht global
injektive Abbildung



Den eindimensionalen Fall kennen wir bereits 7.2. Ist I ein Intervall und $f: I \rightarrow \mathbb{R}$ stetig, so ist f umkehrbar dann und nur dann, wenn f streng monoton ist. In diesem Fall ist $J = f(I)$ ein Intervall, die Umkehrfunktion f^{-1} auf $f(I)$ wohldefiniert und ebenfalls stetig und streng monoton 7.13.

In höheren Dimensionen steht uns das Monotoniekriterium jedoch nicht zur Verfügung, und die Sache ist komplizierter. So kann man zum Beispiel die Abbildung φ in Abbildung 1 so definieren, dass sie *lokal* injektiv ist. Das heißt, jeder Punkt besitzt eine kleine Umgebung, die bijektiv abgebildet wird. Sie ist aber nicht *global* injektiv, denn die beiden hervorgehobenen Punkte links werden auf denselben Punkt rechts abgebildet. Also ist φ insgesamt nicht umkehrbar.

In einem ersten Schritt vereinfachen wir daher die Aufgabe, indem wir sie *lokalisieren* – was ohnehin bei vielen analytischen Problemen eine sinnvolle Herangehensweise ist. Wir fixieren also einen Punkt x_0 und dessen Bild

$$u_0 = \varphi(x_0)$$

und fragen, ob es offene Umgebungen U von x_0 und V von u_0 gibt, in denen die Gleichung $u = \varphi(x)$ eindeutig nach x auflösbar ist.

Beschränken wir uns auf kleine Umgebungen, so können wir in einem zweiten Schritt das Problem *linearisieren*, indem wir die typischerweise *nichtlineare* Abbildung φ lokal durch ihre *Linearisierung* im Punkt x_0 approximieren. Dies ist ohnehin ein wichtiger Spezialfall des allgemeinen Problems und führt zu der Gleichung

$$u = u_0 + D\varphi(x_0)(x - x_0).$$

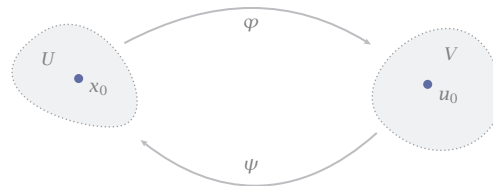
Die lineare Algebra lehrt nun, dass diese Gleichung *uneingeschränkt* – also ohne weitere Annahmen für alle u – lösbar ist genau dann, wenn $D\varphi(x_0)$ invertierbar ist, also

$$\det D\varphi(x_0) \neq 0$$

gilt. Insbesondere müssen u und x von derselben Dimension sein.

Für die lokale Umkehrbarkeit einer differenzierbaren Abbildung φ um einen Punkt x_0 ist es somit sicher sinnvoll zu verlangen, dass $D\varphi(x_0)$ regulär ist. Das fundamentale Ergebnis ist, dass diese Eigenschaft auch *hinreichend* ist.

Abb 2
Abbildung und
Umkehrabbildung



1 **Umkehrsatz** Es sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und

$$\det D\varphi(x_0) \neq 0.$$

Dann existieren Umgebungen U von x_0 und V von $u_0 = \varphi(x_0)$ sowie eine stetig differenzierbare Abbildung $\psi: V \rightarrow U$, so dass

$$\psi \circ \varphi = id_U, \quad \varphi \circ \psi = id_V. \quad \times$$

Im Detail bedeutet dies, dass φ die offene Menge U bijektiv auf die offene Menge V abbildet und deren lokale Umkehrabbildung $\psi = (\varphi|_U)^{-1}$ ebenfalls stetig differenzierbar ist.

■ Etwas Terminologie

Der Satz über die Existenz lokaler Umkehrabbildungen gehört zu den wichtigsten Hilfsmitteln der Analysis. Wir wollen seine wesentlichen Aspekte deshalb auch begrifflich herausstellen.

Definition Eine C^1 -Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt *regulär im Punkt x_0* , und der Punkt selbst *regulärer Punkt* der Abbildung φ , wenn

$$\det D\varphi(x_0) \neq 0.$$

Die Abbildung φ heißt *regulär*, wenn sie in jedem Punkt ihres Definitionsbereichs regulär ist. \times

Ein nicht-regulärer Punkt heißt *singulärer Punkt*, dort ist $\det D\varphi(x_0) = 0$.

Definition Sei $\Omega \subset \mathbb{R}^n$ offen und nicht leer. Eine stetig differenzierbare Abbildung $\varphi: \Omega \rightarrow \mathbb{R}^n$ heißt *Diffeomorphismus*, wenn gilt:

- (i) $\Omega' = \varphi(\Omega)$ ist offen,
- (ii) $\varphi: \Omega \rightarrow \Omega'$ ist bijektiv,
- (iii) $\varphi^{-1}: \Omega' \rightarrow \Omega$ ist ebenfalls stetig differenzierbar.

Genauer heißt dann φ ein *Diffeomorphismus von Ω auf Ω'* . \times

Analog sind zum Beispiel *Homöomorphismus* und *Lipeomorphismus* definiert, hierfür ist nur *stetig differenzierbar* durch *stetig* respektive *lipschitzstetig* zu ersetzen.

▶ A. Eine affine Abbildung

$$\Lambda: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \rightarrow u = Ax + b$$

ist ein Diffeomorphismus des \mathbb{R}^n genau dann, wenn $\det A \neq 0$.

B. Ist I ein offenes Intervall und $f: I \rightarrow \mathbb{R}$ regulär, so verschwindet die Ableitung f' nirgends. Also ist gemäß dem Satz über Umkehrfunktionen 8.15 f ein Diffeomorphismus von I auf $I' = f(I)$.

C. Sind $\varphi: \Omega \rightarrow \Omega'$ und $\psi: \Omega' \rightarrow \Omega''$ Diffeomorphismen, so sind

$$\varphi^{-1}: \Omega' \rightarrow \Omega, \quad \psi \circ \varphi: \Omega \rightarrow \Omega''$$

ebenfalls Diffeomorphismen.

D. Eine reguläre Kurve $\gamma: I \rightarrow \mathbb{R}^2$ kann kein Diffeomorphismus sein. ◀

2 **Notiz** Ein Diffeomorphismus φ ist in jedem Punkt seines Definitionsbereichs regulär. ✕

◀◀◀ Da φ und φ^{-1} beide differenzierbar sind, können wir auf die Identität $\varphi^{-1} \circ \varphi = id$ die Kettenregel 14.7 anwenden und erhalten

$$(D\varphi^{-1} \circ \varphi)D\varphi = Id$$

auf ganz Ω . Dies geht aber nur, wenn $D\varphi$ in jedem Punkt regulär ist. ▶▶▶

Wie bereits am Beispiel von Abbildung 1 bemerkt, gilt die Umkehrung dieser Feststellung im *Globalen* im Allgemeinen *nicht*. Aus der Regularität, einer lokalen Eigenschaft, kann man nicht auf die globale Eigenschaft der Umkehrbarkeit schließen. Dies ist nur lokal möglich, und das ist die Quintessenz des Umkehrsatzes.

3 **Kurzfassung des Umkehrsatzes** Lokal um einen regulären Punkt ist eine stetig differenzierbare Abbildung diffeomorph. ✕

Das bedeutet, dass die Einschränkung auf eine hinreichend kleine Umgebung dieses Punktes einen Diffeomorphismus dieser Umgebung auf sein Bild ergibt. Wichtige Beispiele sind Polar- und Kugelkoordinaten, die wir am Ende dieses Abschnitts betrachten.

■ Rückführung auf einen Spezialfall

Der Beweis des Umkehrsatzes wird übersichtlicher, wenn wir folgenden Spezialfall betrachten. Auf ihn führen wir den allgemeinen Fall zurück.

4 **Spezialfall** Es sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und

$$\varphi(0) = 0, \quad D\varphi(0) = \text{Id}.$$

Dann ist φ lokal um 0 diffeomorph. ✕

⟨⟨⟨ Beweis des Umkehrsatz mithilfe des Spezialfalls Sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine C^1 -Abbildung mit regulärer Ableitung $\Lambda = D\varphi(x_0)$. Verschieben wir den Nullpunkt nach x_0 mittels der Translation $\tau: x \mapsto x + x_0$ und wenden nach φ die affine Transformation $\lambda: u \mapsto \Lambda^{-1}(u - u_0)$ mit $u_0 = \varphi(x_0)$ an, so erhalten wir die *normalisierte Abbildung*

$$\tilde{\varphi} = \lambda \circ \varphi \circ \tau: \tilde{\varphi}(x) = \Lambda^{-1}(\varphi(x + x_0) - u_0).$$

Diese erfüllt die Voraussetzungen des Spezialfalls, denn $\tilde{\varphi}$ ist stetig differenzierbar, $\tilde{\varphi}(0) = 0$, und

$$D\tilde{\varphi}(0) = \Lambda^{-1}D\varphi(x_0) = \text{Id}.$$

Also besitzt $\tilde{\varphi}$ eine stetig differenzierbare lokale Inverse $\tilde{\psi}$. Es gilt also

$$\tilde{\psi} \circ \tilde{\varphi} = \tilde{\psi} \circ \lambda \circ \varphi \circ \tau = \text{id}_{U_0},$$

$$\tilde{\varphi} \circ \tilde{\psi} = \lambda \circ \varphi \circ \tau \circ \tilde{\psi} = \text{id}_{V_0},$$

mit gewissen Umgebungen U_0 und V_0 von 0. Verknüpfen wir die erste Gleichung von links mit τ und von rechts mit τ^{-1} und verfahren analog in der zweiten Gleichung mit λ , so erhalten wir

$$(\tau \circ \tilde{\psi} \circ \lambda) \circ \varphi = \text{id}_U, \quad U = \tau(U_0)$$

$$\varphi \circ (\tau \circ \tilde{\psi} \circ \lambda) = \text{id}_V, \quad V = \lambda^{-1}(V_0).$$

Somit ist φ ein Diffeomorphismus von U auf V mit der stetig differenzierbaren Umkehrabbildung $\psi = \tau \circ \tilde{\psi} \circ \lambda$. ⟩⟩⟩

■ Der Umkehrsatz für lipschitzstetige Abbildungen

Andererseits können wir den Spezialfall 4 etwas allgemeiner fassen. Es stellt sich heraus, dass der Umkehrsatz bereits innerhalb der Kategorie der lipschitzstetigen Abbildungen gilt, ohne dass der Beweis dadurch komplizierter würde.

Sei $\Omega \subset \mathbb{R}^n$ beliebig, aber nicht leer, und $|\cdot|$ eine beliebige Norm auf \mathbb{R}^n . Für eine Lipschitzstetige Abbildung $f: \Omega \rightarrow \mathbb{R}^n$ ist dann

$$[f]_{\Omega} := \sup_{\substack{u \neq v \\ u, v \in \Omega}} \frac{|f(u) - f(v)|}{|u - v|}$$

die bestmögliche Lipschitzkonstante von f auf Ω bezüglich $|\cdot|$. Für eine stetig differenzierbare Abbildung f auf einer konvexen Menge Ω gilt hierfür folgendes

5 **Lemma** *Ist $f: \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar und Ω konvex, so gilt*

$$[f]_{\Omega} = \|Df\|_{\Omega},$$

wobei $\|\cdot\|$ die von $|\cdot|$ induzierte Operatornorm bezeichnet. \times

««« Mit $u, v \in \Omega$ ist auch $[u, v] \subset \Omega$ aufgrund der Konvexität von Ω . Aufgrund des Schrankensatzes 14.16 gilt dann

$$|f(u) - f(v)| \leq \max_{z \in [u, v]} \|Df(z)\| |u - v| \leq \|Df\|_{\Omega} |u - v|.$$

Also gilt $[f]_{\Omega} \leq \|Df\|_{\Omega}$. — Umgekehrt gilt für $z \in \Omega$ und $h \neq 0$

$$\begin{aligned} |Df(z)h| &= \lim_{t \rightarrow 0} \left| \frac{1}{t} (f(z + th) - f(z)) \right| \\ &\leq \lim_{t \rightarrow 0} \frac{1}{|t|} [f]_{\Omega} |th| = [f]_{\Omega} |h|. \end{aligned}$$

Da dies für alle $h \neq 0$ gilt, folgt hieraus $\|Df(z)\| \leq [f]_{\Omega}$. Und da dies für jedes $z \in \Omega$ gilt, folgt auch $\|Df\|_{\Omega} \leq [f]_{\Omega}$. »»»

Für eine Abbildung φ wie im Spezialfall 4 gilt aufgrund der Stetigkeit der ersten Ableitung

$$[\varphi - id]_{B_r(0)} = \|D\varphi - \text{Id}\|_{B_r(0)} \rightarrow 0, \quad r \rightarrow 0.$$

Für hinreichend kleines r ist also $[\varphi - id]_{B_r(0)} < 1$. Dann ist φ bereits injektiv:

Proposition A *Ist $\varphi: \Omega \rightarrow \mathbb{R}^n$ Lipschitzstetig und $[\varphi - id]_{\Omega} < 1$, so ist φ auf Ω injektiv. \times*

««« Angenommen, es ist $\varphi(x) = \varphi(y)$ für $x, y \in \Omega$. Dann ist

$$|x - y| = |(\varphi(x) - x) - (\varphi(y) - y)| \leq [\varphi - id]_{\Omega} |x - y|.$$

Wegen $[\varphi - id]_{\Omega} < 1$ folgt hieraus $|x - y| = 0$, also $x = y$. »»»

Das Problem besteht somit nicht im Nachweis der Injektivität von φ , sondern im Nachweis der Stetigkeit der Umkehrabbildung φ^{-1} auf einer offenen Menge. Dies zeigen wir im Folgenden, indem wir die Umkehrabbildung innerhalb der Klasse Lipschitzstetiger Abbildungen konstruieren.

Proposition B Sei $\varphi: B_r(0) \rightarrow \mathbb{R}^n$ lipschitzstetig mit

$$\varphi(0) = 0, \quad [\varphi - id]_{B_r(0)} \leq 1/4.$$

Dann existiert eine lipschitzstetige Abbildung $\psi: B_{r/2}(0) \rightarrow B_r(0)$ mit

$$\psi(0) = 0, \quad [\psi - id]_{B_{r/2}(0)} \leq 1/2,$$

so dass $\varphi \circ \psi = id_{B_{r/2}(0)}$. \times

■ Formulierung als Fixpunktproblem

Für den Beweis formulieren wir die Aussage als Fixpunktproblem. Dazu schreiben wir $\varphi = id + \hat{\varphi}$ und die gesuchte Umkehrabbildung als $\psi = id + \hat{\psi}$. Zu lösen ist dann die Gleichung

$$id = \varphi \circ \psi = (id + \hat{\varphi}) \circ (id + \hat{\psi}) = id + \hat{\psi} + \hat{\varphi} \circ (id + \hat{\psi}),$$

was äquivalent ist zu

$$\hat{\psi} = -\hat{\varphi} \circ (id + \hat{\psi}).$$

Zu gegebenem $\hat{\varphi}$ suchen wir also ein $\hat{\psi}$, das in einer hinreichend kleinen Umgebung von 0 definiert ist und diese Gleichung erfüllt.

Die Abbildung $\hat{\psi}$ erscheint hier als *Fixpunkt* des Operators

$$T: u \mapsto Tu = -\hat{\varphi} \circ (id + u),$$

also als Lösung der *Fixpunktgleichung*

$$Tu = u.$$

Wir haben das *Invertierungsproblem* ›Gesucht ist die Inverse der Abbildung φ ‹ somit in das *Fixpunktproblem* ›Gesucht ist ein Fixpunkt des Operators T ‹ umformuliert. Dies ist ein oft angewandter Kunstgriff, da für Fixpunktprobleme vielfältige und weitreichende Sätze zur Verfügung stehen. Einer der vielseitigsten ist der Banachsche Fixpunktsatz, den wir bereits beim Beweis des Existenz- und Eindeutigkeitsatzes für gewöhnliche Differenzialgleichungen kennengelernt haben. Wir wiederholen ihn hier der Vollständigkeit halber.

- 6 **Banachscher Fixpunktsatz** Sei E ein Banachraum mit Norm $\|\cdot\|$, sei X eine abgeschlossene Teilmenge von E , und $T: X \rightarrow X$ eine *Kontraktion*. Das heißt, es existiert eine Konstante $\theta \in (0, 1)$, so dass

$$\|Tu - Tv\| \leq \theta \|u - v\|, \quad u, v \in X.$$

Dann besitzt T in X genau einen Fixpunkt ξ . \times

Die Herausforderung bei der Anwendung des Banachschen Fixpunktsatzes besteht darin, im jeweiligen Fall einen geeigneten Banachraum E und eine geeignete Teilmenge X zu finden. Der folgende Beweis gibt dafür ein Beispiel.

««« *Beweis von Proposition B* Schreibe also $\varphi = id + \hat{\varphi}$ und $\psi = id + \hat{\psi}$. Die Gleichung $\varphi \circ \psi = id$ ist dann äquivalent zu

$$\hat{\psi} = -\hat{\varphi} \circ (id + \hat{\psi}).$$

Diese Gleichung lösen wir mit dem Banachschen Fixpunktsatz im Banachraum

$$E = C(K, \mathbb{R}^n), \quad K = \bar{B}_{r/2}(0),$$

versehen mit der Supremumsnorm $\|\cdot\|$. Als Teilmenge betrachten wir

$$X = \{u \in E : u(0) = 0, [u]_K \leq 1/2\}.$$

Es ist leicht zu sehen, dass X in E abgeschlossen ist. Auf X definieren wir

$$Tu = -\hat{\varphi} \circ (id + u).$$

Nun sind eine Reihe von Aussagen zu verifizieren.

T ist wohldefiniert, also Tu auf K definiert: Für $u \in X$ und $x \in K$ ist

$$|u(x)| = |u(x) - u(0)| \leq [u]_K |x|,$$

also

$$|x + u(x)| \leq (1 + [u]_K) |x| < 2|x| < r.$$

Also gilt $id + u: K \rightarrow B_r(0)$, und $\hat{\varphi} \circ (id + u)$ ist auf K definiert.

T bildet X in X ab: Tu ist sicher wieder stetig und

$$Tu(0) = -\hat{\varphi}(u(0)) = -\hat{\varphi}(0) = 0.$$

Mit $[\hat{\varphi}] := [\hat{\varphi}]_{B_r(0)} \leq 1/4$ haben wir für beliebige $x, y \in K$ die Abschätzung

$$\begin{aligned} |Tu(x) - Tu(y)| &= |\hat{\varphi}(x + u(x)) - \hat{\varphi}(y + u(y))| \\ &\leq [\hat{\varphi}] |(x + u(x)) - (y + u(y))| \\ &\leq [\hat{\varphi}] [id + u]_K |x - y| \\ &\leq [\hat{\varphi}] (1 + [u]_K) |x - y|. \end{aligned}$$

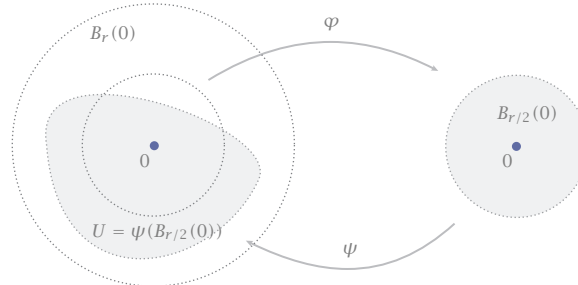
Daraus folgt

$$[Tu]_K \leq [\hat{\varphi}] (1 + [u]_K) \leq \frac{1}{4} \left(1 + \frac{1}{2}\right) \leq \frac{1}{2}.$$

Somit ist $Tu \in X$.

Abb 3

Zum Beweis von Proposition C



T ist eine Kontraktion auf X bezüglich der Supremumsnorm: Es ist

$$\begin{aligned} |Tu(x) - Tv(x)| &= |\hat{\varphi}(x + u(x)) - \hat{\varphi}(x + v(x))| \\ &\leq [\hat{\varphi}] |(x + u(x)) - (x + v(x))| \\ &= [\hat{\varphi}] |u(x) - v(x)| \\ &\leq [\hat{\varphi}] \|u - v\|_K. \end{aligned}$$

Da dies für alle $x \in K$ gilt, folgt

$$\|Tu - Tv\|_K \leq [\hat{\varphi}] \|u - v\|_K.$$

Wegen $[\hat{\varphi}] \leq 1/4$ ist also T eine Kontraktion auf X .

Abschluss des Beweises: Der Banachsche Fixpunktsatz ist somit anwendbar, und T besitzt einen eindeutigen Fixpunkt $\hat{\psi} \in X$. Die Fixpunktgleichung ist aber äquivalent mit $\varphi \circ \psi = id$. Damit ist Proposition B gezeigt. \gggg

- 7 **Proposition C** Sei $\varphi: B_r(0) \rightarrow \mathbb{R}^n$ lipschitzstetig wie in Proposition B und $V = B_{r/2}(0)$. Dann ist φ ein Lipeomorphismus der Nullpunktsumgebung $U = \psi(V)$ auf V mit Umkehrabbildung $\varphi^{-1} = \psi$. \times

\gggg Nach Proposition A ist φ injektiv, und nach Proposition B gibt es eine lipschitzstetige Abbildung $\psi: V \rightarrow B_r(0)$ mit $\varphi \circ \psi = id_V$. Setzen wir also

$$U := \psi(V) \subset B_r(0),$$

so ist $\varphi: U \rightarrow V$ injektiv und surjektiv, also bijektiv, mit lipschitzstetiger Umkehrabbildung $\psi = \varphi^{-1}: V \rightarrow U$. Außerdem ist U offen, denn $U = \varphi^{-1}(V)$ ist das Urbild einer offenen Menge unter der stetigen Abbildung φ . \gggg

■ Regularität

Nun zeigen wir, dass ψ auch differenzierbar ist, wenn φ es ist. Dies geschieht wie im Beweis der Umkehrregel in einer Dimension 8.15, nur tritt an die Stelle der Ableitung φ' die totale Ableitung $D\varphi$.

Proposition D *Ist die Abbildung φ in Proposition C stetig differenzierbar, so ist φ ein Diffeomorphismus von U auf V . \times*

««« Wir hatten gezeigt 5, dass

$$\|D\hat{\varphi}\|_{B_r(0)} = [\hat{\varphi}]_{B_r(0)} \leq 1/4.$$

Somit ist $D\varphi = I + D\hat{\varphi}$ in jedem Punkt invertierbar 6.19. Aufgrund des Lemmas von Hadamard 14.15 gilt nun lokal um jeden Punkt $x_0 \in U$

$$\varphi(x) = \varphi(x_0) + \Lambda(x)(x - x_0)$$

mit

$$\Lambda(x) = \int_0^1 D\varphi((1-t)x_0 + tx) dt.$$

Es ist also

$$\Lambda(x_0) = D\varphi(x_0).$$

Da $\Lambda(x_0)$ invertierbar ist, ist es aus Stetigkeitsgründen auch $\Lambda(x)$ für alle x in einer hinreichend kleinen Umgebung von x_0 . Dort gilt dann auch

$$x = x_0 + \Lambda^{-1}(x)(\varphi(x) - \varphi(x_0)).$$

Mit $x = \psi(u)$ und $x_0 = \psi(u_0)$ ist dies gleichbedeutend mit

$$\begin{aligned} \psi(u) &= \psi(u_0) + \Lambda^{-1}(\psi(u))(u - u_0) \\ &= \psi(u_0) + \Lambda^{-1}(\psi(u_0))(u - u_0) \\ &\quad + [\Lambda^{-1}(\psi(u)) - \Lambda^{-1}(\psi(u_0))](u - u_0). \end{aligned}$$

Aufgrund der Stetigkeit von ψ und Λ verschwindet der Ausdruck in eckigen Klammern für $u \rightarrow u_0$. Wir erhalten somit

$$\psi(u) = \psi(u_0) + \Lambda^{-1}(\psi(u_0))(u - u_0) + o(u - u_0).$$

Also 14.5 ist ψ in u_0 differenzierbar mit Ableitung

$$D\psi(u_0) = (D\varphi)^{-1}(\psi(u_0)).$$

Dies zeigt auch, dass $D\psi$ ebenfalls stetig ist. Damit ist alles gezeigt. »»»

Der Kern des Beweises des Umkehrsatzes betraf lipschitzstetige Abbildungen. Die stetige Differenzierbarkeit der Umkehrabbildung war dann wesentlich einfacher zu zeigen. Höhere Ableitungen bereiten keine weiteren Mühen:

- 8 **Zusatz zum Umkehrsatz** *Gelten die Voraussetzungen des Umkehrsatzes, und ist φ von der Klasse C^r mit $1 \leq r \leq \infty$, so ist auch die lokale Umkehrabbildung ψ von der Klasse C^r .* ✕

⟨⟨⟨ Die Behauptung gilt für $r = 1$ aufgrund des Umkehrsatzes, und es ist

$$D\psi = (D\varphi)^{-1} \circ \psi.$$

Gilt nun die Behauptung für ein $r \geq 1$ und ist φ von der Klasse C^{r+1} , so sind auf der rechten Seite dieser Formel $(D\varphi)^{-1}$ und ψ von der Klasse C^r . Also gilt dasselbe auch für $D\psi$, und damit ist ψ selbst von der Klasse C^{r+1} . ⟩⟩⟩

Wir merken noch an, dass die Lipschitzstetigkeit für die Existenz einer stetigen Umkehrabbildung tatsächlich nicht notwendig ist. Wissen wir bereits, dass die Abbildung injektiv ist, so ist sie auch offen. Der folgende Satz ist auch als Satz von der Invarianz der Dimension bekannt.

Brouwerscher Umkehrsatz *Ist $\varphi: \Omega \rightarrow \mathbb{R}^n$ eine stetige und injektive Abbildung einer offenen Menge Ω im \mathbb{R}^n , so ist $\Omega' = \varphi(\Omega)$ offen und die Umkehrabbildung $\varphi^{-1}: \Omega' \rightarrow \Omega$ ebenfalls stetig.* ✕

Ein relativ einfacher Beweis basiert auf der Theorie des Abbildungsgrads, doch geht dies über den Rahmen dieser Vorlesung hinaus.

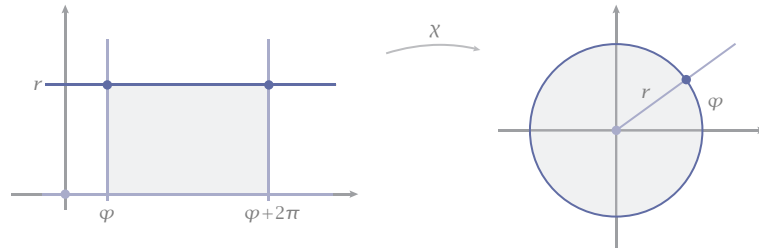
■ Koordinatentransformationen

Einen Diffeomorphismus

$$\varphi: \Omega \rightarrow \Omega', \quad x \mapsto u = \varphi(x),$$

kann man auffassen als eine *Koordinatentransformation*, die auf der Zielmenge Ω' mit Koordinaten u neue Koordinaten x aus der Menge Ω einführt. Koordinatentransformationen sind ein wichtiges Hilfsmittel, um mathematische Probleme zu lösen, und viele mathematische und physikalische Probleme haben ihre eigenen speziellen Koordinatensysteme. Die wichtigsten sind die *Polar-* und *Kugelkoordinaten*.

Abb 4 Polarkoordinaten



Polarkoordinaten In der euklidischen Ebene lässt sich jeder Punkt durch seinen Abstand r zum Nullpunkt und den Winkel φ seines Ortsvektors mit der positiven x -Achse beschreiben. Umgekehrt wird durch

$$\chi: \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

jedem Koordinatenpaar (r, φ) der entsprechende Punkt (x, y) in der Ebene zugeordnet. Für die so definierte Abbildung

$$\chi: [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2, \quad (r, \varphi) \mapsto (x, y)$$

ist

$$D\chi = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}, \quad \det D\chi = r.$$

Somit ist χ regulär in allen Punkten mit $r > 0$ und definiert dort einen lokalen Diffeomorphismus.

Für $r = 0$ dagegen ist χ nicht einmal injektiv, denn

$$\chi(0, \varphi) = (0, 0), \quad \varphi \in \mathbb{R}.$$

Aber auch für $r > 0$ ist χ nicht injektiv, denn aufgrund der Periodizität der Kreisfunktionen gilt ja

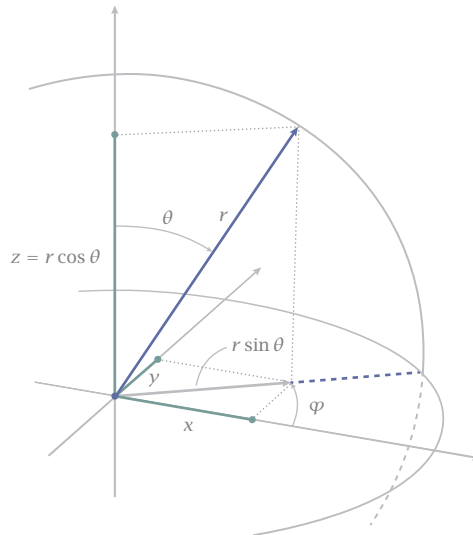
$$\chi(r, \varphi + 2\pi n) = \chi(r, \varphi), \quad n \in \mathbb{Z}.$$

Daher ist χ ein Koordinatensystem nur nach Einschränkung auf geeignete Teilgebiete, wie zum Beispiel $(0, \infty) \times (0, 2\pi)$ oder allgemeiner

$$(0, \infty) \times (\alpha, \alpha + 2\pi), \quad \alpha \in \mathbb{R}.$$

Für die Umkehrabbildung ist dann jeweils der geeignete Zweig der Arcusfunktionen zu wählen. In vielen Fällen ist eine Rücktransformation jedoch nicht erforderlich, und die Mehrdeutigkeit der Polarkoordinaten kein Problem.

Abb 5
Kugelkoordinaten



► **Beispiel** Gesucht sind *rotationssymmetrische* harmonische Funktionen u in der Ebene. Es soll also

$$\Delta u = u_{xx} + u_{yy} = 0$$

gelten, und in Polarkoordinaten soll

$$v(r, \varphi) := u(r \cos \varphi, r \sin \varphi)$$

nicht von φ abhängen, also $v_\varphi = 0$ gelten. — Eine kleine Rechnung ergibt A-15.4

$$u_{xx} + u_{yy} = v_{rr} + \frac{1}{r}v_r + \frac{1}{r^2}v_{\varphi\varphi}.$$

Somit suchen wir eine Lösung der Gleichung

$$v_{rr} + \frac{1}{r}v_r = 0.$$

Deren allgemeine Lösung ist $v = a + b \log r$. Diese Funktion ist allerdings im Nullpunkt stetig und differenzierbar dann und nur dann, wenn $b = 0$. Somit sind die einzigen rotationssymmetrischen und auf ganz \mathbb{R}^2 harmonischen Funktionen die konstanten Funktionen. ◀

Kugelkoordinaten Im euklidischen Raum lässt sich jeder Punkt durch seinen Abstand r zum Nullpunkt und zwei Winkel θ und φ beschreiben, dem **Azimutwinkel** θ seines Ortsvektors zur z -Achse und dem **Polarwinkel** φ seiner Projektion auf die xy -Ebene. Umgekehrt wird durch

$$\chi: \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \varphi \sin \theta \\ r \sin \varphi \sin \theta \\ r \cos \theta \end{pmatrix}$$

jedem Koordinatentripel (r, θ, φ) der entsprechende Punkt (x, y, z) im Raum zugeordnet. Für die so definierte Abbildung

$$\chi: [0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad (r, \theta, \varphi) \mapsto (x, y, z)$$

ist

$$D\chi = \begin{pmatrix} \cos \varphi \sin \theta & r \cos \varphi \cos \theta & -r \sin \varphi \sin \theta \\ \sin \varphi \sin \theta & r \sin \varphi \cos \theta & r \cos \varphi \sin \theta \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}$$

und

$$\det D\chi = r^2 \sin \theta.$$

Diese Jacobideterminante verschwindet somit genau auf der z -Achse.

► Suchen wir im \mathbb{R}^3 nach rotationssymmetrischen harmonischen Funktionen, so führt der Ansatz $u(x, y, z) = v(r)$ zu der Differenzialgleichung

$$v_{rr} + \frac{2}{r}v_r = 0.$$

Die allgemeine Lösung ist in diesem Fall $v_r = -b/r^2$ und damit $v = a + b/r$. Auch in diesem Fall sind die einzigen rotationssymmetrischen und auf ganz \mathbb{R}^3 harmonischen Funktionen die konstanten Funktionen. ◀

18.2

Implizite Funktionen

Der Umkehrsatz ₁ sagt aus, dass eine stetig differenzierbare Abbildung

$$f: \mathbb{R}^n \hookrightarrow \mathbb{R}^n, \quad w = f(x)$$

lokal um $f(x_0) = w_0$ eine stetig differenzierbare Umkehrabbildung genau dann besitzt, wenn $\det Df(x_0) \neq 0$. In diesem Fall werden die n Koordinaten von x

durch die n Gleichungen

$$w_i = f_i(x_1, \dots, x_n), \quad i = 1, \dots, n,$$

lokal eindeutig und in stetig differenzierbarer Weise als Funktionen von w in der Nähe von w_0 bestimmt. Daher sagt man auch › n Gleichungen bestimmen im Allgemeinen eindeutig n Unbekannte‹.

Nun betrachten wir in gleicher Weise eine Abbildung

$$f: \mathbb{R}^n \hookrightarrow \mathbb{R}^m, \quad w = f(x)$$

mit $m \neq n$. Ist $n < m$, so handelt es sich um ein *überbestimmtes System*, das nicht für alle w in der Nähe von w_0 gelöst werden kann, wenn überhaupt. Dies ist bereits bei linearen Gleichungen offensichtlich. Diesen Fall werden wir daher nicht weiter betrachten.

Bleibt der Fall $n > m$. Da die Anzahl der Gleichungen kleiner ist als die Anzahl der Koordinaten, spricht man von einem *unterbestimmten System*. Um die Notation zu vereinfachen, schreiben wir die Dimension des Urbildraumes als $n + m$ mit $m \geq 1$ und $n \geq 1$, und betrachten eine Abbildung

$$f: \mathbb{R}^{n+m} \hookrightarrow \mathbb{R}^m, \quad w = f(x).$$

Die Vermutung liegt nahe, dass hier durch m Gleichungen auch nur m Koordinaten von x bestimmt werden, während n Koordinaten frei gewählt werden können. Somit sind keine eindeutigen Lösungen zu erwarten, sondern *Familien* von Lösungen, die von n Parametern abhängen.

Wir betrachten wieder die Linearisierung des Problems lokal um $w_0 = f(x_0)$. Statt $w = f(x)$ betrachten wir also

$$w = w_0 + Df(x_0)(x - x_0), \tag{1}$$

wobei die Jacobische

$$Df(x_0) = \left(\frac{\partial f_k}{\partial x_l}(x_0) \right)_{1 \leq k \leq m, 1 \leq l \leq n+m}$$

durch eine $m \times (n + m)$ -Matrix dargestellt wird. Diese ist natürlich nicht invertierbar. Hat sie aber *maximalen Rang* m , so gibt es eine Umordnung der Spalten von $Df(x_0)$ derart, dass die hintere $m \times m$ -Untermatrix

$$A = \left(\frac{\partial f_k}{\partial x_l}(x_0) \right)_{1 \leq k \leq m, n+1 \leq l \leq n+m}$$

maximalen Rang m hat und somit invertierbar ist. Nummerieren wir die Koordinaten entsprechend dieser Umordnung um und schreiben

$$x = (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) = (u_1, \dots, u_n, v_1, \dots, v_m) = (u, v),$$

so geht Gleichung (1) über in

$$w = w_0 + B(u - u_0) + A(v - v_0), \quad (2)$$

wobei die Matrix B aus den ersten n Spalten von $Df(x_0)$ besteht. Da A regulär sein soll, können wir diese Gleichung nach v auflösen und erhalten

$$v = v_0 + A^{-1}(w - w_0) - A^{-1}B(u - u_0).$$

Für jedes feste w ist die Lösungsmenge von (2) somit ein n -dimensionaler affiner Unterraum, wo die n Koordinaten u eindeutig die übrigen m Koordinaten v bestimmen.

Soweit das linearisierte Problem. Der Satz über implizite Funktionen sagt nun aus, dass *lokal* dasselbe auch für das nichtlineare Problem gilt.

Um die Formulierung dieses Satzes zu vereinfachen, gehen wir davon aus, dass wir die Koordinaten bereits so nummeriert haben, dass die hintere quadratische Untermatrix von $Df(x_0)$ maximalen Rang hat, und bezeichnen diese Koordinaten wie zuvor mit (u, v) . Dementsprechend sei

$$f_u = \left(\frac{\partial f_k}{\partial u_l} \right)_{1 \leq k \leq m, 1 \leq l \leq n}, \quad f_v = \left(\frac{\partial f_k}{\partial v_l} \right)_{1 \leq k, l \leq m}.$$

Mit den obigen Bezeichnungen ist also $Df = (f_u | f_v) = (B | A)$.

9 Satz über implizite Funktionen (IFS) Sei

$$f : \mathbb{R}^n \times \mathbb{R}^m \hookrightarrow \mathbb{R}^m, \quad w = f(u, v)$$

stetig differenzierbar und $w_0 = f(u_0, v_0)$. Ist

$$\det f_v(u_0, v_0) \neq 0,$$

so existieren eine Umgebung $U \times V$ von (u_0, v_0) sowie eine stetig differenzierbare Abbildung

$$\varphi : U \rightarrow V, \quad u \mapsto v = \varphi(u),$$

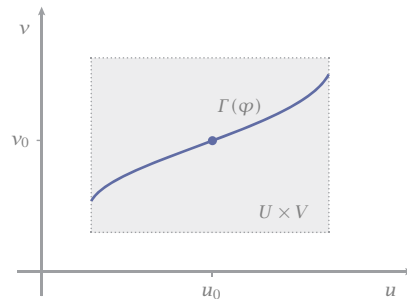
so dass

$$\{(u, v) \in U \times V : f(u, v) = w_0\} = \{(u, \varphi(u)) : u \in U\} = \Gamma(\varphi). \quad \times$$

Innerhalb des ›Fensters‹ $U \times V$ um (u_0, v_0) sind die einzigen Lösungen der Gleichung $f(u, v) = w_0$ also genau diejenigen, die auf dem Graphen $\Gamma(\varphi)$ von φ liegen. *Andere Lösungen gibt es in $U \times V$ nicht.* Insbesondere ist $\varphi(u_0) = v_0$. In diesem Sinne wird die implizite Gleichung $f(u, v) = w_0$ lokal eindeutig nach v aufgelöst durch die stetig differenzierbare Funktion $v = \varphi(u)$.

Diese Funktion φ lässt sich allerdings nur in den wenigsten Fällen explizit angeben. Es handelt sich eben um eine nur *implizit* durch f definierte Funktion.

Abb 6 Satz über implizite Funktionen



Wir wissen aber immerhin, dass $f(u, \varphi(u))$ konstant ist, wodurch wir praktisch ebenso viel über φ erfahren können wie durch eine explizite Formel.

««« *Beweis des Satzes* Wir erhalten diesen Satz aus dem Umkehrsatz ₁, indem wir die Abbildung f geeignet erweitern. Wir definieren dazu in einer Umgebung von (u_0, v_0) die Abbildung

$$F: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m, \quad F(u, v) = (u, f(u, v)).$$

Diese ist ebenfalls stetig differenzierbar, und es ist

$$F(u_0, v_0) = (u_0, f(u_0, v_0)) = (u_0, w_0).$$

Die Jacobimatrix von F ist

$$DF = \begin{pmatrix} I & 0 \\ f_u & f_v \end{pmatrix},$$

denn die erste Komponente von F ist die Identität in u . Aufgrund der Rechenregeln für Determinanten ist daher

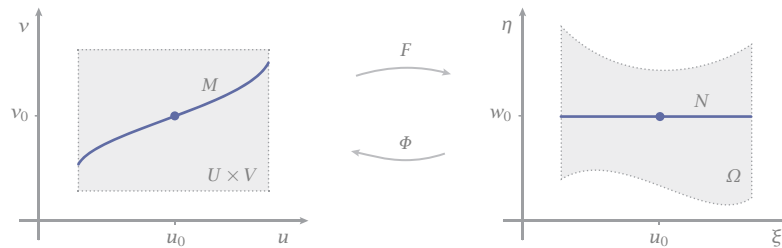
$$\det DF(u_0, v_0) = \det f_v(u_0, v_0) \neq 0.$$

Also ist der Umkehrsatz ₁ anwendbar und F ein lokaler Diffeomorphismus.

Es gibt also eine offene Umgebung $U \times V$ von (u_0, v_0) , die von F diffeomorph auf eine Umgebung Ω von (u_0, w_0) abgebildet wird. Da F in der ersten Komponente die Identität ist, hat die Menge Ω die in Abbildung 7 skizzierte vertikal gescherte Gestalt, und die Umkehrabbildung ist von der Form

$$\Phi: \Omega \rightarrow U \times V, \quad \Phi(\xi, \eta) = (\xi, \varphi(\xi, \eta)).$$

Aus Stetigkeitsgründen können wir U auch noch so klein wählen, dass der horizontale Schnitt

Abb 7 Erweiterte Abbildung F und ihre Umkehrabbildung Φ 

$$N := U \times \{w_0\}$$

ganz in Ω enthalten ist.

Betrachte nun die Menge M aller Lösungen der Gleichung $f(u, v) = w_0$ im Fenster $U \times V$, also

$$M = \{(u, v) \in U \times V : f(u, v) = w_0\}.$$

Deren Bild unter F ist die Menge

$$F(M) = \{(\xi, \eta) \in \Omega : \eta = w_0\} = U \times \{w_0\} = N.$$

Da F umkehrbar ist mit Umkehrabbildung Φ , folgt hieraus

$$\begin{aligned} M &= \Phi(N) \\ &= \{(\xi, \varphi(\xi, w_0)) : (\xi, w_0) \in \Omega\} \\ &= \{(u, \varphi(u, w_0)) : u \in U\} \\ &= \Gamma(\varphi_0) \end{aligned}$$

mit

$$\varphi_0 : U \rightarrow V, \quad u \mapsto v = \varphi(u, w_0).$$

Dies ist die im Satz mit φ bezeichnete gesuchte Abbildung. \gggg

10 Zusatz Für die Ableitung der impliziten Funktion φ gilt

$$\varphi_u(u) = -f_v^{-1} f_u \Big|_{(u, \varphi(u))}.$$

Ist außerdem f von der Klasse C^r mit $1 \leq r \leq \infty$, so gilt dies auch für φ . \times

\llll Für $u \in U$ gilt ja $f(u, \varphi(u)) = w_0$. Da f und φ stetig differenzierbar sind, können wir die Kettenregel $_{14.7}$ anwenden und erhalten

$$0 = D(f(u, \varphi(u))) = f_u + f_v \varphi_u,$$

wobei f_u und f_v an der Stelle $(u, \varphi(u))$ ausgewertet werden. Da im vorangehenden Beweis F als Diffeomorphismus von $U \times V$ nach Ω konstruiert wurde, ist $\det f_v \neq 0$ für alle $u \in U$. Wir können daher die letzte Gleichung nach φ_u auflösen und erhalten die Behauptung. Die Regularitätsaussage folgt unmittelbar aus dem entsprechenden Ergebnis für den Umkehrsatz 8. \gggg

Tatsächlich haben wir im vorangehenden Beweis eine stärkere Aussage bewiesen. Die Menge $M = \{(u, v) \in U \times V : f(u, v) = w_0\}$ können wir ebensogut für alle w in einer hinreichend kleinen Umgebung von w_0 betrachten. Dazu müssen wir nur φ als Funktion von u wie auch w betrachten. Dies führt zu folgendem Satz.

11 Erweiterter Satz über implizite Funktionen (IFS) Sei

$$f : \mathbb{R}^n \times \mathbb{R}^m \hookrightarrow \mathbb{R}^m, \quad w = f(u, v)$$

stetig differenzierbar und $w_0 = f(u_0, v_0)$. Gilt $\det f_v(u_0, v_0) \neq 0$, so existieren Umgebungen $U \times V$ von (u_0, v_0) und W von w_0 sowie eine stetig differenzierbare Abbildung

$$\Phi : U \times W \rightarrow V, \quad v = \Phi(u, w),$$

so dass für jedes $w \in W$

$$\{(u, v) \in U \times V : f(u, v) = w\} = \{(u, \Phi(u, w)) : u \in U\}.$$

Ist außerdem f von der Klasse C^r mit $1 \leq r \leq \infty$, so gilt dies auch für Φ . \times

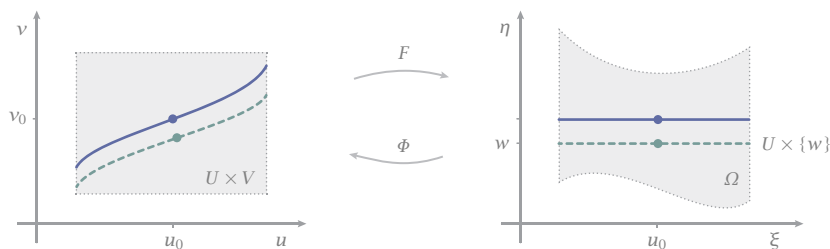
In dem Fenster $U \times V$ ist also nicht nur die Menge $\{f = w_0\}$ der Graph einer Abbildung. Dasselbe gilt auch für jede Menge $\{f = w\}$ mit w in einer hinreichend kleinen Umgebung von w_0 , und die Abhängigkeit von w ist ebenso regulär wie die Abbildung f . Alle diese Mengen werden durch die Abbildung Φ beschrieben.

\gggg Im vorangehenden Beweis wählen wir die Umgebung U noch so klein, dass Ω ein Rechteck $U \times W$ mit einer Umgebung W von w_0 enthält. Dann können wir überall im Beweis w_0 durch $w \in W$ ersetzen und erhalten damit die implizite Lösungsfunktion

$$\Phi : U \times W \rightarrow V, \quad (u, w) \mapsto v = \Phi(u, w).$$

Alles Weitere ist dann klar. \gggg

Abb 8 Zum allgemeinen Satz über implizite Funktionen



■ Skalare Funktionen

Als erste Anwendung des Satzes über implizite Funktionen betrachten wir skalare Funktionen. Sei zunächst

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y)$$

eine Funktion zweier Variablen, (x_0, y_0) ein Punkt im Definitionsbereich von f und $c_0 = f(x_0, y_0)$ sein Bildpunkt. Wir wollen wissen, wie die *Niveaumengen*

$$f^{-1}(c) = \{(x, y) : f(x, y) = c\}$$

für c in der Nähe von c_0 lokal um (x_0, y_0) aussehen.

Ist

$$f_y(x_0, y_0) \neq 0,$$

so können wir die Gleichung $c = f(x, y)$ lokal nach y auflösen und jede Niveaumenge als Graph einer Funktion $g: y = g(x)$ darstellen. Ist

$$f_x(x_0, y_0) \neq 0,$$

so können wir diese Gleichung lokal nach x auflösen und jede Niveaumenge als Graph einer Funktion $h: x = h(y)$ darstellen. Sind beide Bedingungen erfüllt, so sind beide Darstellungen möglich. Außerdem gilt dann

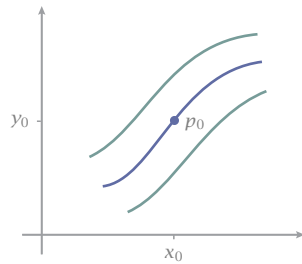
$$g'(x_0) = -\frac{f_x}{f_y} \Big|_{(x_0, y_0)} \neq 0, \quad h'(y_0) = -\frac{f_y}{f_x} \Big|_{(x_0, y_0)} \neq 0.$$

Das lokale Bild entspricht somit qualitativ dem in Abbildung 9.

Gilt dagegen

$$f_x(x_0, y_0) = 0, \quad f_y(x_0, y_0) \neq 0,$$

so ist auch $g'(x_0) = 0$. Die Niveaulinie durch p_0 hat somit einen *Flachpunkt* und sieht im Allgemeinen wie in Abbildung 10 links aus. Entsprechendes gilt, wenn die Rollen von x und y vertauscht sind.

Abb 9 Niveaulinien im Fall $f_y(x_0, y_0) \neq 0$ und $f_x(x_0, y_0) \neq 0$ 

Der nächste Satz überträgt diese Betrachtungen auf Funktionen $\mathbb{R}^n \hookrightarrow \mathbb{R}$.

Satz Sei $f: \mathbb{R}^{n+1} \hookrightarrow \mathbb{R}$ stetig differenzierbar mit $n \geq 1$ und

$$\nabla f(p_0) \neq 0.$$

Dann ist lokal um p_0 jede Niveaumenge $f^{-1}(c)$ mit c in einer hinreichend kleinen Umgebung von $c_0 = f(p_0)$ darstellbar als Graph einer stetig differenzierbaren Funktion $\mathbb{R}^n \hookrightarrow \mathbb{R}$. \times

«««« Wegen $\nabla f(p_0) \neq 0$ ist mindestens eine partielle Ableitung von f nicht Null. Nennen wir diese Koordinate v und die restlichen n Koordinaten u , so ist der IFS anwendbar, und lokal auf jeder Niveaufläche die Koordinate v als Funktion der Koordinaten u darstellbar. »»»»

Lokal um einen Punkt p_0 mit

$$\nabla f(p_0) = 0$$

gelten solche Aussagen im Allgemeinen *nicht*. Dies zeigen die folgenden beiden Beispiele.

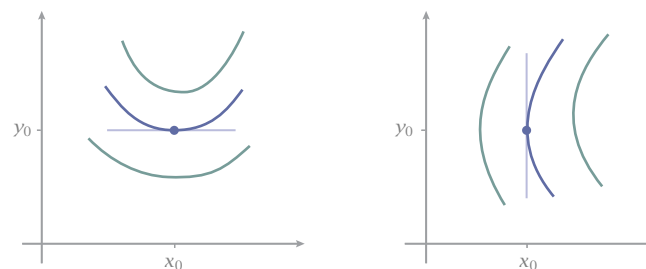
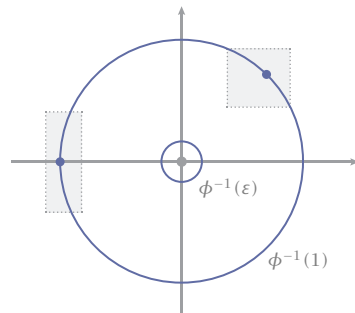
Abb 10 Niveaulinien im Fall $f_x(x_0, y_0) = 0$ respektive $f_y(x_0, y_0) = 0$ 

Abb 11 Niveaulinien von $x^2 + y^2$ mit zwei regulären Punkten

► A. Betrachte

$$\phi : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \phi(x, y) = x^2 + y^2.$$

Der Gradient von ϕ verschwindet im Nullpunkt und sonst nicht. Die Niveaumengen dieser Funktion sind

$$\phi^{-1}(c) = \begin{cases} \text{Kreislinie,} & c > 0, \\ \text{Nullpunkt,} & c = 0, \\ \text{leere Menge,} & c < 0. \end{cases}$$

Also sind in keiner Umgebung von 0 die Niveaumengen von ϕ als Graphen einer Funktion auf einem offenen Intervall darstellbar. In jedem anderen Punkt ist dies möglich: auf den beiden Koordinatenachsen gibt es jeweils eine Möglichkeit, innerhalb der vier offenen Quadranten deren zwei.

B. Dasselbe gilt für

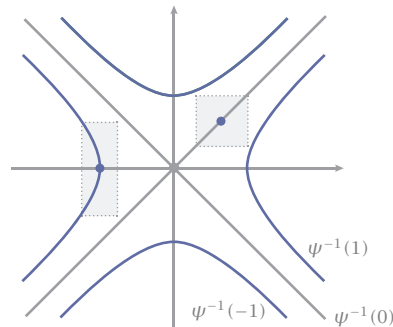
$$\psi : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \psi(x, y) = x^2 - y^2.$$

Die Niveaumengen dieser Funktion sind

$$\psi^{-1}(c) = \begin{cases} \text{zwei Hyperbelbögen durch } x\text{-Achse,} & c > 0, \\ \text{beide Winkelhalbierende,} & c = 0, \\ \text{zwei Hyperbelbögen durch } y\text{-Achse,} & c < 0. \end{cases}$$

C. Die klassische newtonsche Bewegungsgleichung eines reibungsfreien Teilchens der Masse 1 auf der reellen Achse unter dem Einfluss eines Potentials $V: \mathbb{R} \rightarrow \mathbb{R}$ ist

$$\ddot{x} = -V'(x), \quad x \in \mathbb{R}.$$

Abb 12 Niveaulinien von $x^2 - y^2$ mit zwei regulären Punkten

Als System erster Ordnung lauten die Gleichungen

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= -V'(x).\end{aligned}$$

Die *Gesamtenergie* dieses System ist die Summe aus kinetischer und potentieller Energie,

$$E: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad E(x, \dot{x}) = \frac{\dot{x}^2}{2} + V(x).$$

Diese ist konstant entlang jeder Lösung, denn

$$\frac{d}{dt}E(x, \dot{x}) = E_x \dot{x} + E_{\dot{x}} \ddot{x} = V'(x) \dot{x} - \dot{x} V'(x) = 0.$$

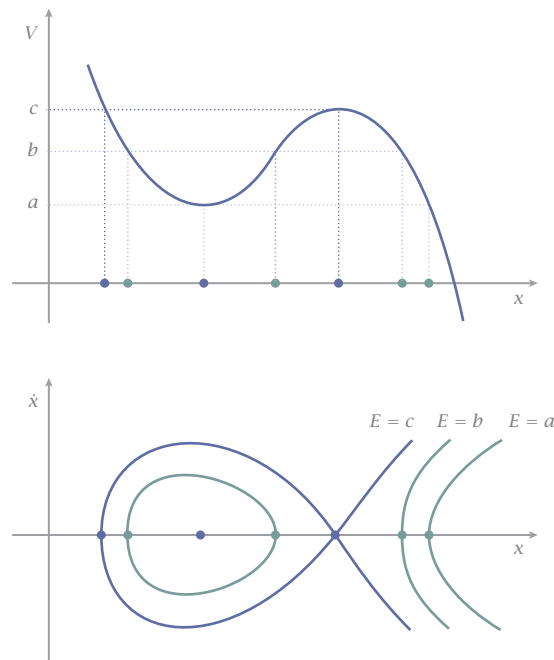
Das ist der klassische *Energieerhaltungssatz*. Jede Lösungskurve ist somit in einer Niveaumenge der Energiefunktion enthalten, und die Niveaumengen liefern bereits Aufschlüsse über die Lösungen der Differentialgleichung $\ddot{x} = -V'(x)$.

Betrachte also die Niveaumengen von E . Wegen $\nabla E = (V'(x), \dot{x})^\top$ liegen kritische Punkte genau dann vor, wenn

$$V'(x) = 0 \wedge \dot{x} = 0.$$

Diese entsprechen kritischen Punkten des Potentials V auf der x -Achse. Alle anderen Punkte sind regulär, und die zugehörigen Niveaulinien sind reguläre, stetig differenzierbare Kurven. Wegen $E(x, -y) = E(x, y)$ verlaufen diese symmetrisch zur x -Achse und schneiden die x -Achse immer senkrecht, da

$$E_{\dot{x}} = 0 \Leftrightarrow \dot{x} = 0,$$

Abb 13 Ein Potential V und sein Phasenportrait

während an diesen Stellen $V'(x) \neq 0$. Mit diesen Überlegungen lässt sich bereits das sogenannte *Phasenportrait* zur Gleichung $\ddot{x} = -V'(x)$ vollständig beschreiben. ◀

■ Reguläre Punkte

Die Formulierung des IFS geht davon aus, dass bereits eine geeignete Zerlegung der Koordinaten in $x = (u, v)$ vorliegt. Im Allgemeinen ist diese jedoch nicht gegeben, und oft gibt es auch mehr als eine solche Zerlegung. Wir formulieren den IFS daher ein weiteres Mal ohne Bezug auf spezielle Koordinaten.

Zentral ist hier der Begriff des *regulären Punktes*. Bisher hatten wir erklärt, dass ein Punkt p regulär heißt für eine stetig differenzierbare Abbildung

$$f : \left\{ \begin{array}{l} \mathbb{R} \hookrightarrow \mathbb{R} \\ \mathbb{R}^n \hookrightarrow \mathbb{R} \\ \mathbb{R}^n \hookrightarrow \mathbb{R}^n \end{array} \right\} \text{ genau dann, wenn } \left\{ \begin{array}{l} f'(p) \neq 0 \\ \nabla f(p) \neq 0 \\ \det Df(p) \neq 0 \end{array} \right\} .$$

Wir verallgemeinern nun diese Definition auf beliebige Abbildungen $\mathbb{R}^n \hookrightarrow \mathbb{R}^m$.

Definition Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig differenzierbar. Ein Punkt p im Definitionsbereich von f heißt *regulärer Punkt* von f , falls $Df(p)$ surjektiv ist. Andernfalls heißt er *singulärer oder kritischer Punkt* von f . \times

Bemerkungen a. Für eine C^1 -Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist $Df(p)$ surjektiv genau dann, wenn $\text{rang } Df(p) = m$. Somit beinhaltet diese Definition die obigen Spezialfälle.

b. Andererseits gilt

$$\text{rang } Df(p) \leq \min \{n, m\}.$$

Im Fall $n < m$ kann also f keine regulären Punkte haben.

c. Insbesondere haben Kurven $\gamma: \mathbb{R} \rightarrow \mathbb{R}^m$ für $m \geq 2$ keine regulären Punkte im Sinne dieser Definition. Der früher eingeführte Begriff des regulären Punktes einer Kurve 13.12 fällt somit *nicht* darunter. Der Begriff ›regulär‹ wird in so vielen Kontexten verwendet, dass solche Kollisionen manchmal auftreten. \rightarrow

12 **Niveaufächensatz** Sei $f: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^m$ stetig differenzierbar mit $n \geq 1$. Dann ist lokal um einen regulären Punkt jede Niveaumenge von f der Graph einer stetig differenzierbaren Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$. \times

⟨⟨⟨ Nach Voraussetzung ist $\text{rang } Df(p) = m$. Durch geeignete Nummerierung der Koordinaten können wir daher erreichen, dass die letzten m Spalten von $Df(p)$ linear unabhängig sind. Schreiben wir jetzt $x = (u, v) \in \mathbb{R}^n \times \mathbb{R}^m$ wie üblich, so gilt

$$\text{rang } f_v(p) = m \Leftrightarrow \det f_v(p) \neq 0.$$

Wir können damit den IFS 11 anwenden und erhalten eine Umgebung

$$X = U \times V \subset \mathbb{R}^n \times \mathbb{R}^m$$

von p , eine Umgebung $W \subset \mathbb{R}^m$ von $f(p)$ und eine stetig differenzierbare Abbildung $\Phi: U \times W \rightarrow V$, so dass für jedes $w \in W$

$$\begin{aligned} f^{-1}(w) \cap X &= \{(u, v) \in X : f(u, v) = w\} \\ &= \{(u, v) \in X : v = \Phi(u, w), u \in U\} \\ &= \Gamma(\Phi(\cdot, w)). \end{aligned}$$

Somit ist lokal um p jede Niveaumenge von f der Graph einer stetig differenzierbaren Abbildung. $\rangle\rangle\rangle$

Bemerkung Der Satz gilt im Prinzip auch für $n = 0$. In diesem Fall ist der Umkehrsatz 1 anwendbar, und die Niveaumengen bestehen aus isolierten Punkten. \rightarrow

18.3 Mannigfaltigkeiten

Der Graph einer stetig differenzierbaren Funktion $\mathbb{R} \hookrightarrow \mathbb{R}$ respektive $\mathbb{R}^2 \hookrightarrow \mathbb{R}^3$ ist ein eindimensionales Kurvenstück in der Ebene respektive ein zweidimensionales Flächenstück im Raum. Aber nicht alle Kurven oder Flächen sind als Graphen einer einzigen Funktion darstellbar. Das zeigen schon die Kreislinie \mathbb{S}^1 und die Sphäre \mathbb{S}^2 .

Man kann diese Mengen aber als *Niveaumengen* einer einzigen, stetig differenzierbaren Funktion beschreiben. So können wir die n -dimensionale Einheitskugel schreiben als $\mathbb{S}^n = g^{-1}(1)$ mit

$$g: \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad g(x) = x_1^2 + \dots + x_{n+1}^2.$$

Jeder Punkt in \mathbb{S}^n ist ein *regulärer* Punkt von g , da $Dg(x) = 2x \neq 0$ für $x \neq 0$. Somit ist die Sphäre \mathbb{S}^n *lokal* immer als Graph einer Funktion $\mathbb{R}^n \hookrightarrow \mathbb{R}$ darstellbar¹². Der Begriff der Mannigfaltigkeit verallgemeinert diese Überlegung.

- 13 **Definition** Sei $n \geq 0$ und $m \geq 1$. Eine nichtleere Teilmenge $M \subset \mathbb{R}^{n+m}$ heißt *Mannigfaltigkeit der Kodimension m* , wenn es eine offene Umgebung Ω von M in \mathbb{R}^{n+m} und eine C^1 -Abbildung $g: \Omega \rightarrow \mathbb{R}^m$ ohne singuläre Punkte gibt, so dass $M = g^{-1}(0) := \{x \in \Omega : g(x) = 0\}$. \times

Bemerkungen a. Der Wert 0 hat keine besondere Bedeutung und kann durch jeden anderen Wert ersetzt werden.

b. Mannigfaltigkeiten der Kodimension 1 werden auch *Hyperflächen* genannt. Im \mathbb{R}^2 und \mathbb{R}^3 spricht man von *Kurven* respektive *Flächen*.

c. Im Fall $n = 0$ besteht M aus isolierten Punkten. Es ist aber sinnvoll, auch solche Objekte als Mannigfaltigkeiten zu betrachten.

d. Genauer haben wir hier *gleichungsdefinierte Untermannigfaltigkeiten* des \mathbb{R}^{n+m} definiert. Der differenzialtopologische Begriff der Mannigfaltigkeit ist wesentlich allgemeiner, aber wir benötigen ihn hier nicht. \rightarrow

Bevor wir zu den Beispielen kommen, stellen wir fest, dass *lokal* jede Mannigfaltigkeit wie ein euklidischer Raum aussieht.

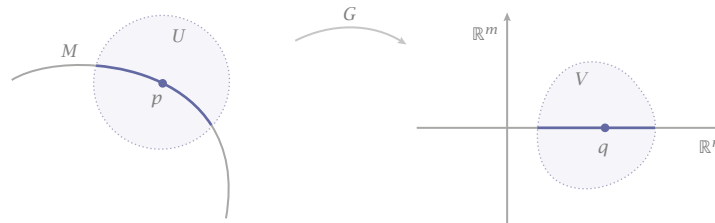
- 14 **Satz** Sei M eine Mannigfaltigkeit im \mathbb{R}^{n+m} der Kodimension m . Dann existiert zu jedem Punkt in M eine Umgebung U und ein Diffeomorphismus

$$F: U \rightarrow V \subset \mathbb{R}^n \times \mathbb{R}^m,$$

so dass

$$F(M \cap U) = (\mathbb{R}^n \times 0^m) \cap V. \quad \times$$

Abb 14 Trivialisierende Koordinaten



Man sagt, der Diffeomorphismus F *trivialisert* lokal die Mannigfaltigkeit, da ein Ausschnitt um p diffeomorph auf einen Ausschnitt des \mathbb{R}^n abgebildet wird.

«» Sei $M = g^{-1}(0)$ mit einer C^1 -Abbildung $g: \mathbb{R}^{n+m} \supset \Omega \rightarrow \mathbb{R}^m$ ohne singuläre Punkte. Sei $p \in M$. Dann hat $Dg(p)$ Rang m , und wir können die Koordinaten so umordnen, dass die hintere $m \times m$ -Untermatrix von $Dg(p)$ maximalen Rang hat. Setzen wir in einer Umgebung von p jetzt

$$G: \mathbb{R}^{n+m} \hookrightarrow \mathbb{R}^{n+m}, \quad (u, v) \mapsto (x, y) = (u, g(u, v)),$$

so ist

$$\det DG(p) = \det \begin{pmatrix} I & 0 \\ g_u & g_v \end{pmatrix}(p) = \det(g_v(p)) \neq 0.$$

Also ist G ein lokaler Diffeomorphismus von einer Umgebung U von p auf eine Umgebung $V \subset \mathbb{R}^n \times \mathbb{R}^m$. Für diesen gilt

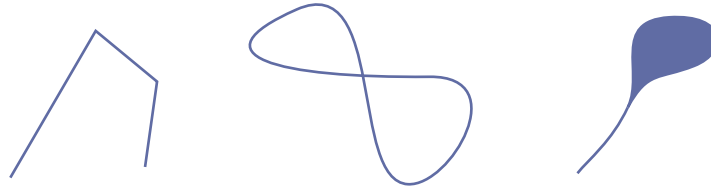
$$G(M \cap U) = G(g^{-1}(0) \cap U) = V \cap \{y = 0\}.$$

Also leistet G das Gewünschte. »»»

Man kann jede solche Abbildung G - beziehungsweise ihre Umkehrabbildung - als ein lokales *Koordinatensystem* auf M betrachten. Die Anzahl n dieser *Koordinaten* ist überall dieselbe und wird als *Dimension* der Mannigfaltigkeit M bezeichnet. Ihre *Kodimension* m ist die Anzahl der *Gleichungen*, durch die M bestimmt wird. Dies gilt auch für die Dimension 0. Eine 0-dimensionale Mannigfaltigkeit ist eine Menge isolierter Punkte. Ihre Kodimension ist die Dimension des Gesamtraumes.

Die Dimension einer Mannigfaltigkeit hängt nicht von ihrem Umgebungsraum ab, lediglich ihre Kodimension. Die Summe aus Dimension und Kodimension einer Mannigfaltigkeit ist immer die Dimension des Gesamtraumes.

Abb 15 Keine Mannigfaltigkeiten



► **Beispiele** A. Im $\mathbb{R}^n \times \mathbb{R}^m \simeq \mathbb{R}^{n+m}$ ist $\mathbb{R}^n \simeq \mathbb{R}^n \times \{0\}$ eine Mannigfaltigkeit der Dimension n und Kodimension m .

B. Ist $A: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ linear und surjektiv, so ist $\ker A = A^{-1}(0)$ eine Mannigfaltigkeit der Dimension n und Kodimension m .

C. Die Sphären $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ sind Mannigfaltigkeiten der Dimension n und Kodimension 1. Dies gilt auch für $\mathbb{S}^0 = \{-1, 1\}$ als Teilmenge von \mathbb{R} .

D. Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ regulär, so ist jede nichtleere Menge $f^{-1}(c)$ eine 0-dimensionale Mannigfaltigkeit, nämlich genau ein Punkt.

E. Der Graph einer C^1 -Abbildung

$$\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad v = \varphi(u)$$

ist eine n -dimensionale Mannigfaltigkeit im \mathbb{R}^{n+m} . Denn ist $D \subset \mathbb{R}^n$ der offene Definitionsbereich von φ , so ist $\Omega = D \times \mathbb{R}^m$ offen in $\mathbb{R}^n \times \mathbb{R}^m$ und

$$g: \Omega \rightarrow \mathbb{R}^m, \quad g(u, v) = v - \varphi(u)$$

stetig differenzierbar. Wegen $g_v = \text{Id}$ sind alle Punkte von g regulär, und es ist

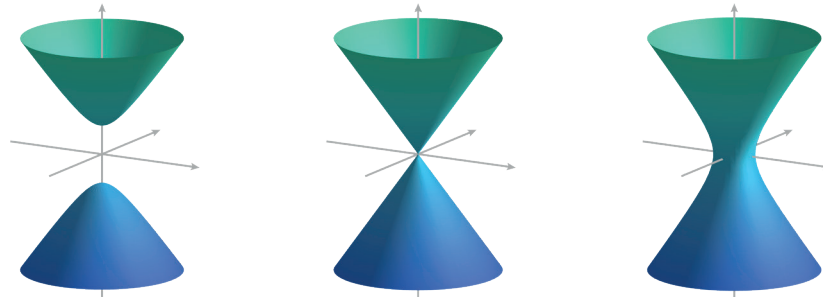
$$\begin{aligned} \Gamma(\varphi) &= \{(u, v) \in \Omega : v = \varphi(u)\} \\ &= \{(u, v) \in \Omega : g(u, v) = 0\} \\ &= g^{-1}(0). \end{aligned}$$

Also ist $\Gamma(\varphi)$ eine n -dimensionale Mannigfaltigkeit in \mathbb{R}^{n+m} .

F. Die Menge $M = \{1/n : n \geq 1\}$ ist eine 0-dimensionale Mannigfaltigkeit in \mathbb{R} , die Menge $M \cup \{0\}$ dagegen *nicht*.

G. Die geometrischen Gebilde in Abbildung 15 sind *keine* Mannigfaltigkeiten, denn solche besitzen keine Ecken, keine Selbstschnitte, und haben überall dieselbe Dimension. ◀

Abb 16 Zweischaliges Hyperboloid, Kegel, und einschaliges Hyperboloid



- 15 ▶ **Hyperboloid und Kegel** Betrachte die Funktion

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

und ihre Niveaumengen

$$M_c = g^{-1}(c) = \{(x, y, z) : x^2 + y^2 - z^2 = c\}.$$

Jeder Punkt mit Ausnahme des Koordinatenursprungs ist ein regulärer Punkt von g , und dieser liegt auf M_0 . Daher ist jede Menge M_c mit $c \neq 0$ eine Mannigfaltigkeit der Kodimension 1 und Dimension 2, also eine Fläche im \mathbb{R}^3 . Für $c > 0$ ist dies ein *einschaliges Hyperboloid*, für $c < 0$ ein *zweischaliges Hyperboloid*. Für $c = 0$ erhält man einen *Kegel*, der aufgrund seiner Spitze im Nullpunkt keine Mannigfaltigkeit bildet. ◀

■ Reguläre Werte

Die Definition einer Mannigfaltigkeit $M = g^{-1}(0)$ verlangt von der definierenden Funktion g mehr als tatsächlich erforderlich ist. Es genügt, dass jeder Punkt *auf M selbst* ein regulärer Punkt von g ist. Dies führt zum Begriff des *regulären Wertes*.

Definition Sei $n \geq 0$ und $m \geq 1$, und $g: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ sei stetig differenzierbar.

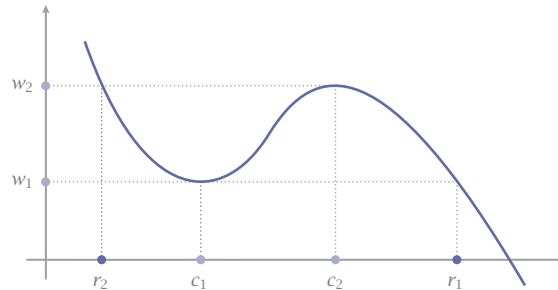
Ein Punkt $w \in \mathbb{R}^m$ heißt *regulärer Wert* von g , wenn die Menge $g^{-1}(w)$ entweder leer ist oder nur aus regulären Punkten besteht. Andernfalls heißt w ein *singulärer* oder *kritischer Wert* von g . ✕

Bemerkungen a. Ein regulärer Punkt ist also ein Punkt im *Definitionsbereich*, ein regulärer Wert ein Punkt im *Wertebereich* einer Abbildung.

b. Der Wert eines regulären Punktes muss kein regulärer Wert sein, denn auch nichtreguläre Punkte können auf denselben Wert abgebildet werden Abb 17.

Abb 17

Zwei kritische Punkte,
zwei reguläre Punkte
sowie zwei kritische
Werte



c. Ob w ein regulärer Wert von g ist, hängt auch vom Definitionsbereich ab. Für die Funktion $g: (x, y, z) \mapsto x^2 + y^2 - z^2$ des letzten Beispiels 15 ist 0 kein regulärer Wert, da

$$0 \in g^{-1}(0), \quad \nabla g(0) = 0.$$

Entfernen wir den Nullpunkt aus dem Definitionsbereich von g , so wird 0 ein regulärer Wert. Die zugehörige Niveaulfläche ist ein Kegel ohne seine Spitze, und diese Menge ist eine Mannigfaltigkeit. ∞

Mit diesem Begriff erhalten wir folgende äquivalente Charakterisierung einer Mannigfaltigkeit.

- 16 Satz** Eine nichtleere Teilmenge M des \mathbb{R}^{n+m} ist eine Mannigfaltigkeit der Dimension n genau dann, wenn es eine C^1 -Funktion $g: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ mit regulärem Wert 0 gibt, so dass $M = g^{-1}(0)$. \times

««« Ist $M = g^{-1}(0)$ eine Mannigfaltigkeit im Sinne unserer Definition 13 , so ist insbesondere jeder Punkt in $g^{-1}(0)$ ein regulärer Punkt und damit 0 selbst ein regulärer Wert von g .

Sei umgekehrt 0 ein regulärer Wert von g . Dann ist jeder Punkt p in $M = g^{-1}(0)$ ein regulärer Punkt von g und damit $Dg(p)$ surjektiv. Dies ist äquivalent mit der Eigenschaft, dass die Determinante einer geeigneten Auswahl von m Spalten von $Dg(p)$ nicht verschwindet. Diese Determinante hängt stetig von p ab und ist somit auch auf einer offenen Umgebung von p nicht Null. Es existiert also zu jedem Punkt $p \in M$ eine Umgebung $U(p)$ in \mathbb{R}^{n+m} , so dass

$$\text{rang } Dg(x) = m, \quad x \in U(p).$$

Setzen wir $\Omega = \bigcup_{p \in M} U(p)$, so erhalten wir eine M umfassende offene Menge Ω ohne singuläre Punkte von g . Also ist M eine Mannigfaltigkeit im Sinne unserer Definition 13 . »»»

Vereinbarung Im Folgenden heiÙe eine nichtleere Teilmenge M des \mathbb{R}^{n+m} eine g -definierte Mannigfaltigkeit, wenn $M = g^{-1}(0)$ mit einer stetig differenzierbaren Funktion $g: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ mit regulärem Wert 0. \times

■ Tangentialraum und Normalraum

Jedem Punkt p einer Mannigfaltigkeit M können wir *Tangentialvektoren* zuordnen. Dazu betrachten wir beliebige Kurven auf M durch p und deren Geschwindigkeitsvektoren in diesem Punkt.

- 17 **Definition** Ein Vektor $v \in \mathbb{R}^{n+m}$ heißt *Tangentialvektor* an M im Punkt p , wenn es eine C^1 -Kurve $\gamma: \mathbb{R} \rightarrow M$ gibt mit

$$\gamma(0) = p, \quad \dot{\gamma}(0) = v.$$

Die Menge aller Tangentialvektoren an M im Punkt p heißt der *Tangentialraum* von M an p und wird mit T_pM bezeichnet. \times

Man stellt sich Tangentialvektoren üblicherweise als im Punkt p angeheftete ›gebundene‹ Ortsvektoren vor. Tatsächlich handelt es sich um ›ungebundene‹ Vektoren, also Elemente eines Vektorraumes. Der Raum T_pM ist ein *Vektorraum* derselben Dimension wie M :

- 18 **Satz** Sei M eine g -definierte Mannigfaltigkeit. Dann ist

$$T_pM = \ker Dg(p), \quad p \in M.$$

Insbesondere ist jeder Tangentialraum T_pM ein Vektorraum derselben Dimension wie M . \times

⟨⟨⟨ Sei $G: U \rightarrow V$ ein lokaler trivialisierender Diffeomorphismus um p 14, so dass

$$G(M \cap U) = (\mathbb{R}^n \times 0^m) \cap V.$$

Für die Umkehrabbildung $\Phi = G^{-1}$ um $q = F(p)$ gilt dann $g \circ \Phi \equiv 0$ entlang aller Kurven in $(\mathbb{R}^n \times 0^m) \cap V$. Insbesondere gilt

$$(g \circ \Phi)(q + te_i) \equiv 0, \quad 1 \leq i \leq n,$$

auf einem t -Intervall um 0, und damit

$$Dg(p)D\Phi(q)e_i = 0, \quad 1 \leq i \leq n.$$

Die Vektoren $D\Phi(q)e_i$ liegen somit sämtlich im Kern von $Dg(p)$. Außerdem sind sie linear unabhängig, da $D\Phi(q)$ regulär ist. Da der Kern von $Dg(p)$ aber

Dimension n hat, folgt

$$\text{span} \{D\Phi(q)e_i : 1 \leq i \leq n\} = \ker Dg(p).$$

Auf der anderen Seite ist klar, dass der Raum links die Geschwindigkeitsvektoren sämtlicher möglicher Kurven auf M durch p enthält, also

$$T_p M = \text{span} \{D\Phi(q)e_i : 1 \leq i \leq n\}$$

gilt. Also ist $T_p M = \ker Dg(p)$. \gggg

Versehen wir den \mathbb{R}^{n+m} mit einem beliebigen Skalarprodukt $\langle \cdot, \cdot \rangle$, so ist auch die Orthogonalität zweier Vektoren und das orthogonale Komplement einer Teilmenge erklärt. Damit können wir auch den *Normalraum* einer Mannigfaltigkeit definieren.

Definition Sei M eine Mannigfaltigkeit im \mathbb{R}^{n+m} . Ist dieser mit einem Skalarprodukt versehen, so heißt das orthogonale Komplement zum Tangentialraum $T_p M$ der *Normalraum* von M in p und wird mit $T_p^\perp M$ bezeichnet. Seine Elemente heißen die *Normalenvektoren* von M in p . \times

Für jeden Punkt $p \in M$ gilt also

$$T_p M \oplus T_p^\perp M = \mathbb{R}^{n+m}.$$

19 Satz Sei M eine g -definierte Mannigfaltigkeit der Dimension n im \mathbb{R}^{n+m} . Ist dieser mit einem Skalarprodukt versehen, so ist

$$T_p^\perp M = \text{span} \{\nabla g_1(p), \dots, \nabla g_m(p)\}, \quad p \in M. \quad \times$$

Die Gradienten der Komponentenfunktionen von g bezüglich dieses Skalarproduktes stehen also überall senkrecht auf der Niveaumenge $M = g^{-1}(0)$ und spannen in jedem Punkt deren Normalraum auf.

\llll Nach dem letzten Satz ist $T_p M = \ker Dg(p)$. Für jede Komponentenfunktion g_k von g gilt also

$$Dg_k(p)v = \langle \nabla g_k(p), v \rangle = 0, \quad v \in T_p M.$$

Also ist

$$\nabla g_k(p) \in T_p^\perp M, \quad 1 \leq k \leq m.$$

Andererseits sind alle diese Gradienten linear unabhängig, da g im Punkt p regulär ist. Da der Normalraum $T_p^\perp M$ genau die Dimension m hat, müssen diese m linear unabhängigen Vektoren diesen Raum aufspannen. \gggg

► Wir betrachten noch einmal

$$g: \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad f(x) = \|x\|^2.$$

Der Gradient von g verschwindet nur im Nullpunkt, und dort ist $g(0) = 0$. Also ist 1 ein regulärer Wert von g , und $\mathbb{S}^n = g^{-1}(1)$ eine Mannigfaltigkeit der Kodimension 1 und Dimension n . Ein Normalenvektor im Punkt $x \in \mathbb{S}^n$ ist x selbst. Somit gilt

$$T_x^\perp \mathbb{S}^n = \text{span} \{x\}, \quad T_x \mathbb{S}^n = \{x\}^\perp.$$

Die *Tangentialebene* an \mathbb{S}^n im Punkt x ist übrigens die zu $T_x M$ parallele affine Ebene durch den Punkt x , also

$$\begin{aligned} x + T_x \mathbb{S}^n &= x + \{x\}^\perp = x + \{v : \langle v, x \rangle = 0\} \\ &= \{v : \langle v - x, x \rangle = 0\} \\ &= \{v : \langle v, x \rangle = 1\}. \quad \blacktriangleleft \end{aligned}$$

18.4

Extrema mit Nebenbedingungen

Ein oft auftretendes Problem ist, Extremwerte einer skalaren Funktion zu bestimmen, während gleichzeitig eine Anzahl von Nebenbedingungen in Form von Gleichungen einzuhalten sind.

Ein anschauliches Beispiel ist die Aufgabe, auf einer Fläche diejenigen Punkte mit dem größten oder kleinstem Abstand zu einem Referenzpunkt zu finden. Als Zielfunktion dient hier bequemerweise das *Quadrat* des euklidischen Abstands, also

$$f(x) = x_1^2 + \dots + x_n^2.$$

Die Fläche selbst soll durch m Gleichungen

$$g_i(x) = 0, \quad 1 \leq i \leq m.$$

beschrieben werden. Ist 0 ein regulärer Wert der Vektorfunktion

$$g = (g_1, \dots, g_m)^\top : \mathbb{R}^n \rightarrow \mathbb{R}^m,$$

so handelt es sich bei dieser Fläche um eine durch g bestimmte Mannigfaltigkeit M der Kodimension m . Es geht dann darum, Extremwerte der *Einschränkung* $f|_M$ von f auf M zu bestimmen. Die primäre Aufgabe ist dabei, deren *kritischen*

Punkte zu finden. Die Frage, ob dort ein Maximum oder Minimum vorliegt, ergibt sich meist aus geometrischen Überlegungen.

Naheliegender ist die Idee, die Mannigfaltigkeit M mehr oder weniger geschickt zu parametrisieren und dadurch die Nebenbedingungen aufzulösen. Die Zielfunktion könnte man dann wie gewohnt untersuchen^{15.14}. Dies ist allerdings im Allgemeinen sehr mühsam – und auch nicht nötig. Denn solche kritischen Punkte lassen sich geometrisch sehr einfach charakterisieren.

Im Folgenden hat der Gesamttraum wieder die Dimension n . Eine Mannigfaltigkeit der Kodimension m hat jetzt also die Dimension $n - m$.

- 20 **Satz** Sei M eine Mannigfaltigkeit und f eine in einer Umgebung von M stetig differenzierbare skalare Funktion. Dann besitzt $f|_M$ einen kritischen Punkt in $p \in M$ genau dann, wenn

$$\nabla f(p) \in T_p^\perp M. \quad \times$$

⟨⟨⟨ Die Einschränkung der Funktion f auf M besitzt im Punkt $p \in M$ einen kritischen Punkt genau dann, wenn sie entlang aller Kurven auf M durch p dort einen kritischen Punkt besitzt. Somit müssen sämtliche Richtungsableitungen in *tangentialen* Richtungen an M verschwinden:

$$Df(p)v = \langle \nabla f(p), v \rangle = 0, \quad v \in T_p M.$$

Also gilt $\nabla f(p) \perp T_p M$ und damit $\nabla f(p) \in T_p^\perp M$. ⟩⟩⟩

Dies ist ein abstraktes geometrisches Resultat. Es ist unabhängig davon, wie die Mannigfaltigkeit M gegeben ist. Konkreter wird es für gleichungsdefinierte Mannigfaltigkeiten.

- 21 **Korollar** Sei M eine g -definierte Mannigfaltigkeit im \mathbb{R}^n der Kodimension m und $f: \mathbb{R}^n \rightarrow \mathbb{R}$ in einer Umgebung von M stetig differenzierbar. Dann besitzt $f|_M$ einen kritischen Punkt in $p \in M$ genau dann, wenn es reelle Zahlen $\lambda_1, \dots, \lambda_m$ gibt, genannt *Lagrangemultiplikatoren*, so dass

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \dots + \lambda_m \nabla g_m(p). \quad \times$$

⟨⟨⟨ Für einen kritischen Punkt p von $f|_M$ gilt ja^{19, 20}

$$\nabla f(p) \in T_p^\perp M = \text{span} \{ \nabla g_1(p), \dots, \nabla g_m(p) \}.$$

Also ist $\nabla f(p)$ eine Linearkombination von $\nabla g_1(p), \dots, \nabla g_m(p)$. ⟩⟩⟩

Für die praktische Anwendung formulieren wir dieses Ergebnis noch einmal für den Standardfall. Aus kritischen Punkten der Funktion f mit Nebenbedingungen werden kritische Punkte einer Hilfsfunktion F ohne Nebenbedingungen, wobei allerdings die Dimension des Problems vergrößert wird.

Satz Es seien $f: \mathbb{R}^n \rightarrow \mathbb{R}$ und $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ auf einem gemeinsamen Definitionsbereich stetig differenzierbar und 0 ein regulärer Wert von g . Dann besitzt f unter der Nebenbedingung $g = 0$ einen kritischen Punkt p genau dann, wenn die *erweiterte Funktion*

$$F: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}, \quad F(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle$$

einen kritischen Punkt (p, μ) besitzt. \times

«««« Besitzt F einen kritischen Punkt (p, μ) , so ist erstens

$$0 = \partial_{\lambda_k} F(p, \mu) = g_k(p), \quad 1 \leq k \leq m,$$

und zweitens

$$0 = \partial_{x_l} F(p, \mu) = \partial_{x_l} f(p) + \sum_{k=1}^m \mu_k \partial_{x_l} g_k(p), \quad 1 \leq l \leq n.$$

Ersteres ist äquivalent mit $p \in M = g^{-1}(0)$, und Letzteres ist äquivalent mit

$$0 = \nabla F(p, \mu) = \nabla f(p) + \sum_{k=1}^m \mu_k \nabla g_k(p),$$

also mit

$$\nabla f(p) \in T_p^\perp M.$$

Also ist p ein kritischer Punkt von f unter den Nebenbedingungen $g = 0$. Die Umkehrung beweist man mit denselben Argumenten. »»»»

Damit erhalten wir folgendes Verfahren zur Lösung von Extremalaufgaben mit Nebenbedingungen. Sind die Zielfunktion f und die Nebenbedingungen $g = (g_1, \dots, g_m) = 0$ gegeben, und ist 0 ein regulärer Wert von g , so bilden wir die erweiterte Funktion

$$F = f + \langle \lambda, g \rangle$$

und bestimmen deren kritische Punkte bezüglich x und λ . Dies führt zu einem System aus $n + m$ Gleichungen für ebensoviele Unbekannte x_1, \dots, x_n und $\lambda_1, \dots, \lambda_m$, und zwar

$$\begin{aligned} 0 &= \partial_{x_k} f(x) + \lambda_1 \partial_{x_k} g_1(x) + \dots + \lambda_m \partial_{x_k} g_m(x), & 1 \leq k \leq n, \\ 0 &= g_l(x), & 1 \leq l \leq m. \end{aligned}$$

Dieses ist zu lösen. Der Erfolg ist allerdings nicht garantiert, da es sich im Allgemeinen um *nichtlineare* Gleichungen handelt. — Dazu zwei Beispiele.

► *Youngsche Ungleichung* Betrachte

$$f(u, v) = \frac{u^p}{p} + \frac{v^q}{q},$$

wobei

$$p, q > 1, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Auf dem positiven Quadranten $\{(u, v) : u, v > 0\}$ besitzt f keine lokalen Extrema. Wir beschränken uns deshalb auf die Teilmenge mit $uv = 1$, stellen also die Nebenbedingung

$$g(u, v) = uv - 1 = 0.$$

Die erweiterte Funktion ist dann

$$F(u, v, \lambda) = \frac{u^p}{p} + \frac{v^q}{q} + \lambda(uv - 1).$$

Für die kritischen Punkte erhalten wir das Gleichungssystem

$$\begin{aligned} u^{p-1} + \lambda v &= 0, \\ v^{q-1} + \lambda u &= 0, \\ uv &= 1. \end{aligned}$$

Multiplizieren der ersten Gleichung mit u und der zweiten mit v ergibt $u^p = v^q$, und zusammen mit $uv = 1$ folgt $u = v = 1$. Offensichtlich handelt es sich hierbei um ein absolutes Minimum. Als Ergebnis erhalten wir

$$\frac{u^p}{p} + \frac{v^q}{q} \geq \frac{1}{p} + \frac{1}{q} = 1$$

für alle $u, v > 0$ mit $uv = 1$.

Um die allgemeine Ungleichung für $u, v > 0$ zu erhalten, setzen wir

$$\tilde{u} = \frac{u}{(uv)^{1/p}}, \quad \tilde{v} = \frac{v}{(uv)^{1/q}}.$$

Es ist $\tilde{u}\tilde{v} = 1$, also

$$\frac{\tilde{u}^p}{p} + \frac{\tilde{v}^q}{q} = \frac{u^p}{puv} + \frac{v^q}{quv} \geq 1.$$

Also gilt allgemein

$$\frac{u^p}{p} + \frac{v^q}{q} \geq uv, \quad u, v > 0. \quad \blacktriangleleft$$

► *Eigenwerte* Sei $A \in L(\mathbb{R}^n)$ ein symmetrischer Operator bezüglich eines beliebigen Skalarprodukts $\langle \cdot, \cdot \rangle$ auf dem \mathbb{R}^n . Die quadratische Form

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) = \langle Ax, x \rangle$$

ist auf dem Gesamtraum unbeschränkt, wenn sie nicht identisch Null ist. Auf der kompakten Einheitskugel $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \langle x, x \rangle = 1\}$ nimmt sie aber ihr Minimum und Maximum an, besitzt also dort im Fall $n \geq 2$ mindestens zwei kritische Punkte.

Um diese zu charakterisieren, schreiben wir die Nebenbedingung als

$$g(x) = 1 - \langle x, x \rangle.$$

Die erweiterte Funktion lautet dann

$$F(x, \lambda) = \langle Ax, x \rangle + \lambda(1 - \langle x, x \rangle).$$

Das zugehörige Gleichungssystem ist

$$\begin{aligned} Ax - \lambda x &= 0, \\ \langle x, x \rangle &= 1. \end{aligned}$$

Ein kritischer Punkt der quadratischen Form f auf \mathbb{S}^n ist demnach ein normierter *Eigenvektor* von A , sein Lagrangemultiplikator der zugehörige *Eigenwert*.

Da die quadratische Form kritische Punkte besitzen *muss*, erhalten wir damit einen Beweis, dass jeder symmetrische Operator mindestens einen reellen Eigenwert besitzt.

Eine Erweiterung dieses Arguments liefert sogar die Existenz einer orthonormalen Basis für A . Ist v der im ersten Schritt gefundene normierte Eigenvektor, so suchen wir im nächsten Schritt ein Maximum auf $\mathbb{S}^{n-1} \cap \{v\}^\perp$. Mit anderen Worten, wir betrachten nun die Nebenbedingungen

$$\begin{aligned} g_0(x) &= 1 - \langle x, x \rangle = 0, \\ g_1(x) &= \langle v, x \rangle = 0. \end{aligned}$$

Eine Maximalstelle auf $g^{-1}(0) = \mathbb{S}^{n-1} \cap \{v\}^\perp$ tritt dann an einem weiteren Eigenvektor auf, der aufgrund der Konstruktion orthogonal zu v ist. Dabei spielt es keine Rolle, ob der zugehörige Eigenwert vom ersten verschieden ist oder nicht.

Führt man so fort, so erhält man eine orthogonale Basis des Gesamtraumes, bestehend auf Eigenvektoren des Operators A . ◀