VIII Umkehrabbildungen und Implizite Funktionen

In diesem Abschnitt untersuchen wir vier Themenkomplexe, die Frage wann eine stetig differenzierbare Funktion eine ebensolche Umkehrfunktion besitzt, das Auflösen von Gleichungen in $\mathbb{R}^n$ und damit zusammenhängend den Satz über implizite Funktionen, Untermannigfaltigkeiten des $\mathbb{R}^n$ und damit zusammenhängend die Bestimmung von Extremwerten unter Nebenbedingungen.

Die Beantwortung der beiden ersten Fragen erweist sich als deutlich schwieriger als in der eindimensionalen Situation, da der Zwischenwertsatz aus der Analysis I kein n-dimensionales Analogon besitzt. Unser zentrales Hilfsmittel zur Herleitung des Satzes über die lokale Umkehrbarkeit einer stetig differenzierbaren Abbildung ist der Banachsche Fixpunktsatz. Wir zeigen mit seiner Hilfe zunächst, dass unter gewissen Bedingungen lokal eine stetige Umkehrung existiert. Die Kettenregel impliziert dann, dass diese wiederum stetig differenzierbar ist.

Der Satz über implizite Funktionen beschäftigt sich dann mit der Frage unter welchen Bedingungen eine Gleichung $f(x,y) = 0$ in der Nähe einer Nullstelle von f eine differenzierbare Auflösung $y = g(x)$ besitzt. Dieser Satz führt aus geometrischer Sicht zum Begriff der Untermannigfaltigkeit in $\mathbb{R}^n$; dies sind diejenigen Teilmengen die lokal wie offene Teilmengen eines $\mathbb{R}^d$ aussehen. Die zunächst geometrisch motivierte Einführung der Begriffe des Tangential- bzw. Normalenraums an eine Untermannigfaltigkeit erlauben dann auch einen eleganten Beweis der sogenannten Multiplikatorenregel von Lagrange, einer notwendigen Bedingung für die Existenz von Extremwerten unter Nebenbedingungen.

1 Der Satz über die Umkehrabbildung

Wir beginnen diesen Abschnitt mit der Definition des Begriffs des Diffeomorphismus.

1.1 Definition. Eine bijektive und stetig differenzierbare Funktion $f : U \to V$ zwischen offenen Mengen $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ heißt *Diffeomorphismus*, falls die Umkehrabbildung $f^{-1} : V \to U$ ebenfalls stetig differenzierbar ist.

1.2 Bemerkungen.
a) Ist $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \to J$ eine stetig differenzierbare Funktion mit $f'(x) \neq 0$ für alle $x \in I$, so ist f streng monoton, die Umkehrfunktion existiert im Intervall $J = f(I)$ und sie ist nach dem Satz über die Umkehrfunktion IV.1.8 aus Analysis I auch differenzierbar.
b) Ein zu a) analoges Resultat gilt nicht für Funktionen in $\mathbb{R}^n$ für $n \geq 2$! Als Gegenbeispiel betrachten wir die Polarkoordinatenabbildung

$$f : (0, \infty) \times \mathbb{R} \to \mathbb{R}^2 \backslash \{(0,0)\}, \quad f(r, \varphi) := (r \cos \varphi, r \sin \varphi).$$

Diese ist surjektiv und stetig differenzierbar, $J_f(r, \varphi)$ ist invertierbar für alle $(r, \varphi) \in (0, \infty) \times \mathbb{R}$, aber f ist nicht injektiv. Also ist f auch nicht invertierbar.
c) Es seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ und $f : U \to V$ ein Diffeomorphismus mit der Umkehrabbildung $g := f^{-1} : V \to U$. Nach der Kettenregel gilt

$$\begin{aligned} g'(f(x)) \cdot f'(x) &= (g \circ f)'(x) = (\mathrm{id}_U)'(x) = \mathrm{id}_{\mathbb{R}^n}, \quad x \in U, \\ f'(g(y)) \cdot g'(y) &= (f \circ g)'(y) = (\mathrm{id}_V)'(y) = \mathrm{id}_{\mathbb{R}^m}, \quad y \in V, \end{aligned}$$

und aus der Linearen Algebra folgt, dass $n = m$ gilt und dass für $y = f(x)$ die Abbildungen $g'(y)^{-1} \in \mathcal{L}(\mathbb{R}^n)$ und $f'(x) \in \mathcal{L}(\mathbb{R}^n)$ zueinander inverse Isomorphismen sind.
d) Die Aussagen b) und c) implizieren, dass $Df(x)$ für alle $x \in U$ invertierbar ist, falls $f : U \to V$ ein Diffeomorphismus ist. Die Umkehrung gilt nur, falls man f geeignet einschränkt.

1.3 Theorem. (Satz über die Umkehrabbildung). *Es seien $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \to \mathbb{R}^n$ eine stetig differenzierbare Funktion, für welche $Df(a) \in \mathcal{L}(\mathbb{R}^n)$ für $a \in U$ invertierbar ist. Dann existiert eine offene Umgebung V von $b := f(a)$ und eine Umgebung $\tilde{U} \subset U$ von a, derart dass die Einschränkung $\tilde{f} : \tilde{U} \to V$ von f auf $\tilde{U}$ ein Diffeomorphismus ist. Für die Umkehrabbildung gilt außerdem*

$$D\tilde{f}^{-1}(b) = (Df(a))^{-1}.$$

Für den Beweis benötigen wir die folgenden beiden Hilfssätze.

1.4 Lemma. (Schrankensatz). *Es seien $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ eine differenzierbare Funktion und $a, x \in U$ mit $\overline{ax} \in U$. Gilt*

$$\sup_{t \in [0,1]} \|Df(a + t(x - a))\| =: L < \infty,$$

so ist $\|f(x) - f(a)\| \leq L\|x - a\|$.

Für den Beweis verweisen wir auf die Übungen.

1.5 Lemma. (Stetigkeit der Inversion). *Die Menge* $GL_n(\mathbb{R}) := \{A \in M_n(\mathbb{R}) : A$ *in-vertierbar* $\}$ *ist offen in* $\mathbb{R}^{n^2}$ *und die Abbildung*

$$\text{Inv} : GL_n(\mathbb{R}) \to M_n(\mathbb{R}),\ A \mapsto A^{-1}$$

ist stetig.

Beweis. Da $GL_n(\mathbb{R}) = \det^{-1}(\{x \in \mathbb{R} : x \neq 0\})$ gilt, ist die Menge $GL_n(\mathbb{R})$ als Urbild der Menge $\{x \in \mathbb{R} : x \neq 0\}$ unter der stetigen Funktion $\det : M_n(\mathbb{R}) \to \mathbb{R}$ offen. Für die Stetigkeit der Inversenabbildung verweisen wir auf die Übungen.

□

Beweis von Theorem 1.3. Wir unterteilen den relativ umfangreichen Beweis in sechs Teilschritte und beginnen mit der folgenden Vorbemerkung.
Betrachtet man anstelle von f die Funktion $Df(a)^{-1} \circ f : U \to \mathbb{R}^n$ so folgt, dass wir oBdA

$$Df(a) = \text{Id}_{\mathbb{R}_n} \tag{1.1}$$

wählen können. Ferner dürfen wir oBdA ebenfalls

$$a = 0 \quad \text{und} \quad f(a) = 0$$

wählen, indem wir die Funktion $x \mapsto f(x+a) - f(a)$ für $x \in \{z \in \mathbb{R}^n : z + a \in U\}$ betrachten.
Schritt 1: Ziel ist es die Gleichung $y = f(x)$ für „kleine" $y \in \mathbb{R}^n$ nach x aufzulösen. Hierzu setzen wir für $y \in \mathbb{R}^n$ und $x \in U$

$$\varphi_y(x) := y + x - f(x).$$

Es gilt dann $y = f(x)$ genau dann, wenn $\varphi_y(x) = x$ ist, d.h. genau dann, wenn x ein Fixpunkt der Abbildung φ_y ist.
Schritt 2: Wir wenden auf die obige Gleichung den Banachschen Fixpunktsatz an. Zunächst ist $\varphi_0 : U \to \mathbb{R}^n$ stetig differenzierbar und die Skalierung (1.1) impliziert, dass

$$D\varphi_0(0) = \text{Id}_{\mathbb{R}^n} - \text{Id}_{\mathbb{R}^n} = 0$$

gilt. Daher existiert ein $r > 0$ so, dass $\overline{U}_{2r}(0) = \{z \in \mathbb{R}^n : |z| \leq 2r\} \subset U$ und

$$\|D\varphi_0(x) - \underbrace{D\varphi_0(0)}_{=0}\| = \|D\varphi_0(x)\| \leq \frac{1}{2}, \quad x \in \overline{U}_{2r}(0), \tag{1.2}$$

gilt. Da $D\varphi_y = D\varphi_0$ ist, folgt aus dem Schrankensatz Lemma 1.4

$$\|\varphi_y(x_1) - \varphi_y(x_0)\| \leq \frac{1}{2}\|x_1 - x_0\|, \quad x_0,\ x_1 \in \overline{U}_{2r}(0) \tag{1.3}$$

und somit

$$\|\varphi_y(x)\| \leq \|\varphi_y(x) - \varphi_y(0)\| + \|\underbrace{\varphi_y(0)}_{=y}\| \leq \frac{1}{2}\|x\| + \|y\| < 2r, \tag{1.4}$$

falls $\|x\| \leq 2r$ und $\|y\| < r$ ist. Mit anderen Worten bedeutet dies, dass für $\|y\| < r$ die Abbildung

$$\varphi_y : \overline{U}_{2r}(0) \to \overline{U}_{2r}(0)$$

eine strikte Kontraktion ist. Da $\overline{U}_{2r}(0)$ als abgeschlossene Teilmenge des $\mathbb{R}^n$ vollständig ist, folgt aus dem Banachschen Fixpunktsatz, dass für alle $y \in \mathbb{R}^n$ mit $\|y\| < r$ genau ein Fixpunkt $x \in \overline{U}_{2r}(0)$ von φ_y existiert, für welchen wegen (1.4) sogar $\|x\| < 2r$ gilt. Setzt man

$$V := U_r(0) \text{ und } \tilde{U} := f^{-1}(V) \cap U_{2r}(0),$$

so bedeutet dies, dass die Einschränkung $\tilde{f}$ von f auf $\tilde{U}$ bijektiv ist.
Schritt 3: Wir zeigen als nächstes, dass $g := (\tilde{f})^{-1} : V \to \tilde{U}$ stetig ist. Für jedes $x \in \tilde{U}$ gilt $x = \varphi_0(x) + f(x)$. Damit und wegen (1.3) gilt für alle $y_0, y_1 \in V$ die Abschätzung

$$\|g(y_1) - g(y_0)\| \leq \underbrace{\|\varphi_0(g(y_1)) - \varphi_0(g(y_0))\|}_{\leq \frac{1}{2}\|g(y_1) - g(y_0)\|} + \|f(g(y_1)) - f(g(y_0))\|$$

und somit

$$\|g(y_1) - g(y_0)\| \leq 2\|f(g(y_1)) - f(g(y_0))\| = 2\|y_1 - y_0\|. \tag{1.5}$$

Dies bedeutet, dass g sogar Lipschitz-stetig ist.
Schritt 4: Wir zeigen, dass $Df(x) \in \mathcal{L}(\mathbb{R}^n)$ für alle $x \in \tilde{U}$ invertierbar ist. Zunächst gilt

$$f(x) = x - \varphi_0(x), \quad \text{und} \quad Df(x) = \mathrm{Id}_{\mathbb{R}^n} - D\varphi_0(x), \quad x \in \tilde{U}.$$

Gilt $Df(x)v = 0$ für ein $v \in \mathbb{R}^n$, so folgt $v = [D\varphi_0(x)]v$ und wegen (1.2) gilt

$$\|v\| \leq \|D\varphi_0(x)\| \cdot \|v\| \leq \frac{1}{2}\|v\|,$$

also $v = 0$. Daher ist $Df(x)$ injektiv und die Dimensionsformel aus der Linearen Algebra impliziert, dass $Df(x)$ auch surjektiv ist. Also ist $Df(x)$ invertierbar.
Schritt 5: Wir zeigen, dass g differenzierbar ist. Hierzu sei $y_0 \in V$ und $k \in \mathbb{R}^n$ mit $y_0 + k \in V$. Setzt man $x_0 := g(y_0)$ und $h := g(y_0 + k) - g(y_0)$, so ist $k = f(x_0 + h) - f(x_0)$ und

$$g(y_0 + k) - g(y_0) - [Df(x_0)]^{-1}k = h - [Df(x_0)]^{-1}(f(x_0 + h) - f(x_0))$$

Da f in x_0 differenzierbar ist, gilt

$$f(x_0+h)-f(x_0)=Df(x_0)h+r(x_0+h)$$

mit $\lim_{h\to 0}\frac{r(x_0+h)}{\|h\|}=0$. Daher ist

$$g(y_0+k)-g(y_0)=[Df(x_0)]^{-1}k-[Df(x_0)]^{-1}r(x_0+h)$$

und es bleibt zu zeigen, dass

$$\lim_{k\to 0}\frac{[Df(x_0)]^{-1}r(x_0+h)}{\|k\|}=0 \tag{1.6}$$

gilt. Aufgrund von Gleichung (1.5) gilt

$$\|h\|=\|(x_0+h)-x_0\|\le 2\|\underbrace{f(x_0+h)-f(x_0)}_{=k}\|=\|2k\|,$$

und für $k\to 0$, also auch $h\to 0$, folgt

$$\frac{\|r(x_0+h)\|}{\|k\|}\le\frac{2}{\|h\|}\|r(x_0+h)\|\overset{h\to 0}{\longrightarrow}0,$$

da f differenzierbar ist. Da $[Df(x_0)]^{-1}\in\mathcal{L}(\mathbb{R}^n)$ stetig ist, folgt schließlich (1.6).
Schritt 6: Anwenden der Kettenregel analog zu Bemerkung 1.2 c) liefert

$$D_g(y)=[Df(g(y))]^{-1}$$

und Lemma 1.5 impliziert, dass $D_g:V\to\mathcal{L}(\mathbb{R}^n)$ stetig ist. Daher ist $\tilde{f}:\tilde{U}\to V$ ein Diffeomorphismus und der Satz ist vollständig bewiesen.

□

Der obige Satz über die lokale Umkehrbarkeit hat zahlreiche Konsequenzen. Als unmittelbare Folgerung notieren wir die Sätze über offene Abbildungen und die Diffeomorphie.

1.6 Korollar. (Satz von der offenen Abbildung). *Es sei $U\subset\mathbb{R}^n$ offen und $f:U\to\mathbb{R}^n$ eine stetig differenzierbare Funktion, derart dass $Df(x)\in\mathcal{L}(\mathbb{R}^n)$ für alle $x\in U$ invertierbar ist. Dann ist $f(U)$ offen in $\mathbb{R}^n$.*

Beweis. Nach dem Umkehrsatz existiert zu jedem $x\in U$ eine offene Umgebung $U_x\subset U$ von x so, dass $f(U_x)\subset\mathbb{R}^n$ offen ist. Wegen

$$f(U)=\bigcup_{x\in U}f(U_x)$$

und da beliebige Vereinigungen offener Mengen wiederum offen sind, ist auch $f(U)$ offen.

□

Die obige Aussage bedeutet, dass $f(O)$ offen ist für alle offenen Mengen $O \subset U$. Abbildungen mit dieser Eigenschaft heißen auch *offene Abbildungen*.

1.7 Korollar. *Ist in der Situation des Satzes von der offenen Abbildung die Funktion f zusätzlich injektiv, so ist f ein Diffeomorphismus von U auf die offene Menge $f(U) \subset \mathbb{R}^n$.*

Beweis. Die Umkehrabbildung $g : f(U) \to U$ ist stetig, da für jede offene Menge $O \subset U$ das Urbild $g^{-1}(O) = f(O)$ nach dem Satz über die offene Abbildung offen ist. Nach dem Satz über die Umkehrabbildung ist f sogar ein Diffeomorphismus.

□

1.8 Bemerkung. Verwendet man den Satz über die Umkehrabbildung um nichtlineare Gleichungssysteme zu lösen, so erhält man folgende Aussage:
Ist $\det Df(x_0) \neq 0$ für ein $x_0 \in U$, so existieren Umgebungen U von x_0 und V von $f(x_0)$ derart, dass das Gleichungssystem

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= y_1 \\ \vdots \quad &\quad \vdots \\ f_n(x_1, \ldots, x_n) &= y_n \end{aligned}$$

für jeden Wert $(y_1, \ldots, y_n) \in V$ genau eine Lösung

$$x_1 = x_1(y_1, \ldots, y_n), \ldots, x_n = x_n(y_1, \ldots, y_n)$$

in U besitzt. Ferner besitzen die Funktionen $x_1, \ldots, x_n$ dieselbe Regularität wie $f_1, \ldots, f_n$.

1.9 Beispiele.
a) *Ebene Polarkoordinaten*
Jeder Punkt $(x, y) \in \mathbb{R}^2$ besitzt eine Darstellung durch Polarkoordinaten, d.h.

$$x = r\cos\varphi,\ y = r\sin\varphi \qquad \text{mit } r = (x^2 + y^2)^{1/2}.$$

Für die Abbildung $f : (r, \varphi) \mapsto (x, y)$ gilt

$$\det Df(r, \varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r \neq 0, \quad (x, y) \neq (0, 0).$$

Zu jedem Punkt $(x,y) \neq (0,0)$ existieren unendlich viele Urbilder $(r, \varphi + 2k\pi)$; in Übereinstimmung mit dem Umkehrsatz gibt es jedoch in einer Umgebung von $(x_0, y_0) = (r_0 \cos\varphi_0, r_0 \sin\varphi_0) \neq (0,0)$ eine unendlich oft differenzierbare Umkehrfunktion gegeben durch

$$r = (x^2 + y^2)^{1/2} \quad \text{und } \varphi = \arctan \frac{y}{x}, \quad x_0 \neq 0$$

wobei jener Funktionszweig des Arcustangens zu wählen ist, für welchen (x_0, y_0) den Wert φ_0 ergibt. Gilt $x_0 = 0$, so wählen wir φ als $\varphi = \operatorname{arccot} \frac{x}{y}$.
Die Polarkoordinatendarstellung bildet also den Streifen

$$S = \{(r, \varphi) \in \mathbb{R}^2 : r > 0, |\varphi| < \pi\}$$

diffeomorph auf die längs der negativen reellen Achse geschlitzten Ebene $\mathbb{R}^2_- := \mathbb{R}^2 \setminus \{(x,y) : x \leq 0, y = 0\}$ ab.
b) *Der komplexe Logarithmus*
Die komplexe Zahl $z \in \mathbb{C}$ besitze die Darstellung $z = x + iy$. Beschränkt man z auf den Streifen $S = \mathbb{R} \times (-\pi, \pi]$ und setzt man $w := e^z$, so folgt $|w| = e^x$ und $\arg w = y$. Dies bedeutet, dass die Exponentialfunktion den Streifen S bijektiv auf $\mathbb{C} \setminus \{0\}$ abbildet. Ihre Umkehrabbildung ist durch $(x,y) \mapsto (\log|w|, \arg w)$ mit $y \in (-\pi, \pi]$ gegeben. Die Periodizität der Exponentialfunktion, d.h. $e^z = e^{z+2\pi i}$, impliziert, dass auch jeder um $2\pi k i$ verschobene Streifen $S_k = 2\pi k i + S$ mit $k \in \mathbb{Z}$ bijektiv auf $\mathbb{C} \setminus \{0\}$ abgebildet wird. Die obige Umkehrformel bleibt gültig mit dem Zusatz, dass jetzt $(2k-1)\pi < y \leq (2k+1)\pi$ gefordert werden muss.
Bezeichnet man für gegebenes $w \neq 0$ jede die Gleichung $e^z = w$ erfüllende komplexe Zahl z als Logarithmus von w, so sehen wir dass es in jedem Streifen S_k genau einen Logarithmus von w gibt. Die verschiedenen Logarithmen unterscheiden sich lediglich um Vielfache von $2\pi i$ und sie sind durch

$$\log w = \log |w| + i \arg w$$

gegeben. Wird das Argument durch $\arg w \in (-\pi, \pi)$ eingeschränkt, so spricht man vom *Hauptzweig* des Logarithmus. In der Sprache des Umkehrsatzes dreht es sich bei $w = e^z$ um die Funktion $f = (f_1, f_2)$ mit

$$f_1(x,y) = e^x \cos y, \quad f_2(x,y) = e^x \sin y$$

und

$$\det\, Df(x,y) = \begin{vmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{vmatrix} = e^{2x} \neq 0.$$

Beschränkt man (x,y) auf den Streifen $S = \mathbb{R} \times (-\pi, \pi)$, so beschreibt $f = (f_1, f_2)$ einen Diffeomorphismus mit dem Bildbereich $\mathbb{R}^2_-$. Die Umkehrabbildung ist dann der Hauptzweig des Logarithmus gegeben durch

$$x = \frac{1}{2} \log(f_1^2 + f_2^2), \quad y = \arg(f_1, f_2) \in (-\pi, \pi).$$