

27. Implizite Funktionen

Bevor wir uns ab dem nächsten Kapitel der Integration widmen, wollen wir zum Abschluss unseres Studiums differenzierbarer Abbildungen noch auf das in der Praxis sehr wichtige Thema der sogenannten impliziten Funktionen eingehen, bei dem es um die Auflösbarkeit von Gleichungen nach bestimmten Variablen geht. Die Idee dieser Situation lässt sich am besten an einem Beispiel erklären.

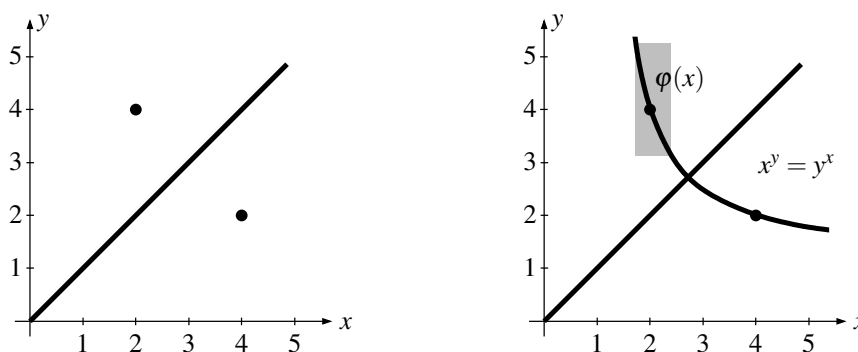
Beispiel 27.1. Für $x, y \in \mathbb{R}_{>0}$ wollen wir die Lösungsmenge der Gleichung $x^y = y^x$ untersuchen.

Wenn wir alle Lösungen einer solchen Gleichung in zwei Variablen bestimmen wollen, würden wir sie normalerweise vermutlich nach einer der Variablen – z. B. y – auflösen wollen, um dann daraus für jeden Wert von x die zugehörigen Werte von y berechnen zu können. Allerdings lässt sich diese Gleichung mit den uns bekannten speziellen Funktionen aus Kapitel 9 leider nicht nach y (und auch nicht nach x) auflösen, da beide Variablen sowohl in der Basis als auch im Exponenten auftreten. Um diese Gleichung zu untersuchen, brauchen wir also andere Methoden.

Als Erstes fällt uns vermutlich auf, dass wir schon ein paar Lösungen sehen können:

- $x = y$ für beliebige $x, y > 0$, sowie
- $(x, y) = (2, 4)$ und $(x, y) = (4, 2)$ (denn $2^4 = 4^2 = 16$),

also insgesamt die im folgenden Bild links eingezeichnete Punktmenge.



Dieses Bild sieht natürlich sehr merkwürdig aus: Ist z. B. der Punkt $(2, 4)$ wirklich ein isolierter Punkt der Lösungsmenge, oder gibt es in einer Umgebung davon noch weitere Lösungen? Um dies zunächst einmal numerisch herauszufinden, könnten wir die Lösungsmenge der betrachteten Gleichung von einem Computer berechnen lassen, der einfach alle Punkte der Ebene abtastet und diejenigen Paare (x, y) zeichnet, bei denen x^y gleich bzw. sehr nahe bei y^x ist. Das Ergebnis, das wir so erhalten würden, ist im Bild oben rechts dargestellt.

Danach sieht es so aus, als ob die Lösungsmenge der gegebenen Gleichung aus zwei Zweigen besteht: den Punkten mit $y = x$, und einer weiteren Kurve, die sich als Graph einer (stetig differenzierbaren) Funktion schreiben lässt. In einer (im Bild oben grau eingezeichneten) Umgebung des Punktes $(2, 4)$ kann man die gegebene Gleichung $x^y = y^x$ also z. B. anscheinend nach y auflösen und als Funktionsgleichung $y = \varphi(x)$ schreiben – auch wenn wir diese Funktion nicht explizit angeben können. Man sagt, dass diese Funktion φ *implizit* durch die Gleichung $x^y = y^x$ definiert ist. Am Kreuzungspunkt der beiden Zweige oben lässt sich die gegebene Gleichung jedoch nicht nach einer der beiden Variablen auflösen und z. B. y als Funktion von x schreiben, weil in einer Umgebung dieses Punktes ja für jeden Wert von x zwei mögliche Werte y mit $x^y = y^x$ existieren.

Ziel dieses Kapitels ist es, derartige Aussagen exakt zu beweisen. Dabei werden wir auch sehen, dass man mit solchen nicht nach einer Variablen aufgelösten Funktionsdefinitionen durchaus arbeiten

kann. So werden wir z. B. im obigen Bild den Schnittpunkt der beiden Zweige und die Ableitung der Funktion φ im Punkt $(2, 4)$ bestimmen können, auch ohne diese Funktion explizit zu kennen (siehe Beispiel ?? ??).

27.A Lokale Umkehrfunktionen

Wir beginnen unser Studium impliziter Funktionen mit einem wichtigen Spezialfall: Sind $D \subset \mathbb{R}^n$ offen und $f: D \rightarrow \mathbb{R}^n$ eine gegebene Funktion, so wollen wir untersuchen, ob wir die Gleichung $y = f(x)$ (mit $x \in D$ und $y \in \mathbb{R}^n$) nach x auflösen, also eine Umkehrfunktion $x = f^{-1}(y)$ finden können. Im eindimensionalen Fall wissen wir dies bereits:

Beispiel 27.2 (Umkehrbarkeit im Eindimensionalen). Sind $D = (a, b)$ ein offenes Intervall und $f: D \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $f'(x) \neq 0$ für alle $x \in D$, so ist f' zunächst einmal entweder überall positiv oder überall negativ, da f' sonst nach dem Zwischenwertsatz 8.21 auch irgendwo den Wert 0 annehmen müsste. Also ist f nach Folgerung 10.24 dann streng monoton und damit injektiv. Auf dem Bildbereich $f(D)$, der nach dem Zwischenwertsatz ebenfalls ein Intervall ist, existiert also eine Umkehrfunktion f^{-1} von f , und diese ist nach Satz 10.11 ebenfalls differenzierbar mit Ableitung $(f^{-1})'(f(x)) = \frac{1}{f'(x)}$. Mit anderen Worten ist die Gleichung $y = f(x)$ somit auf dem betrachteten Intervall nach x auflösbar, nämlich durch die Umkehrfunktion $x = f^{-1}(y)$.

Bemerkung 27.3 (Ableitung einer Umkehrfunktion). Die Formel für die Ableitung der Umkehrfunktion lässt sich sofort auf den mehrdimensionalen Fall verallgemeinern: Ist $f: D \rightarrow D'$ eine differenzierbare Funktion zwischen offenen Teilmengen D und D' von \mathbb{R}^n , und wissen wir bereits, dass f bijektiv ist und eine ebenfalls differenzierbare Umkehrfunktion $f^{-1}: D' \rightarrow D$ besitzt, so erhält man durch Differenzieren der Gleichung $f^{-1}(f(x)) = x$ für alle $x \in D$ mit der Kettenregel aus Satz 25.30 sofort $(f^{-1})'(f(x)) \cdot f'(x) = E$, da die Ableitung von $x \mapsto x = Ex$ nach Beispiel 25.5 die Einheitsmatrix ist. Also muss die Matrix $f'(x) \in \mathbb{R}^{n \times n}$ invertierbar sein, d. h. es muss $\det f'(x) \neq 0$ gelten, und für die Ableitung von f^{-1} ist analog zum Eindimensionalen die inverse Matrix

$$(f^{-1})'(f(x)) = (f'(x))^{-1}.$$

Wo wir im Eindimensionalen $f'(x) \neq 0$ für alle $x \in D$ vorausgesetzt haben, müssen wir nun also verlangen, dass die Matrix $f'(x)$ überall invertierbar ist. Überraschend ist dabei allerdings, dass die Invertierbarkeit von $f'(x)$ für $n > 1$ im Gegensatz zum eindimensionalen Fall in Beispiel 27.2 nicht mehr hinreichend für die Existenz einer Umkehrfunktion ist, wie das folgende einfache Beispiel zeigt.

Beispiel 27.4 (Umkehrbarkeit der Polarkoordinaten). Wir betrachten noch einmal die Polarkoordinatenabbildung in $\mathbb{C} = \mathbb{R}^2$

$$f: \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \mathbb{R}^2, \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} =: \begin{pmatrix} x \\ y \end{pmatrix}$$

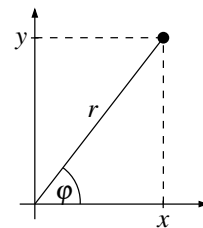
(siehe Satz 9.27 und Definition 9.12), die dem Betrag und Winkel einer komplexen Zahl ihren Real- und Imaginärteil zuordnet. Dann ist die Ableitungsmatrix von f

$$f' = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}, \quad (1)$$

und damit nach Satz 9.14 (b)

$$\det f' = r(\cos^2 \varphi + \sin^2 \varphi) = r > 0.$$

Die Ableitungsmatrix f' ist also überall invertierbar. Trotzdem ist f aber nicht injektiv, da die Addition von Vielfachen von 2π zum Winkel φ nichts am Funktionswert ändert. Um eine bijektive Abbildung zu erhalten, müssen wir f einschränken: Betrachten wir z. B. wie im Bild unten nur die Werte von r und φ mit $1 < r < 2$ und $0 < \varphi < \frac{\pi}{2}$, so ist f auf dieser offenen Teilmenge U des Definitionsbereichs injektiv. Das Bild dieser Teilmenge unter f ist der unten rechts im Bild dargestellte



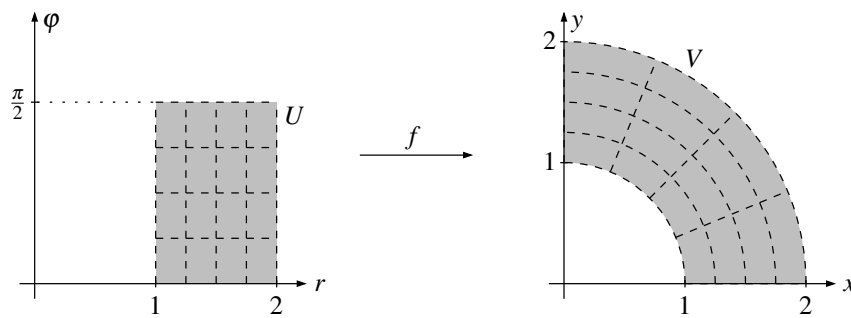
Viertelkreisring V , so dass die Einschränkung $f|_U: U \rightarrow V$ nun bijektiv ist und damit eine Umkehrabbildung $f^{-1}: V \rightarrow U$ besitzt. In der Tat können wir diese Umkehrabbildung auch sofort aus dem geometrischen Bild oben rechts ablesen: Man kann r und φ in diesem Winkelbereich offensichtlich mit den (stetig differenzierbaren) Formeln

$$r = \sqrt{x^2 + y^2} \quad \text{und} \quad \varphi = \arctan \frac{y}{x}$$

aus x und y zurückgewinnen, so dass also

$$f^{-1}: V \rightarrow U, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan \frac{y}{x} \end{pmatrix} \quad (2)$$

dort die Umkehrfunktion ist.



Gemäß Bemerkung 27.3 können wir die Ableitung dieser Umkehrfunktion g nun als die zu f' inverse Matrix berechnen: Mit (1) und Beispiel 18.22 erhalten wir

$$(f^{-1})' = \frac{1}{r} \cdot \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ -\frac{y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix},$$

wie man durch Differenzieren der expliziten Formel (2) für f^{-1} natürlich auch direkt bestätigen könnte.

Auch bei überall invertierbarer Ableitungsmatrix können wir also nur lokal (d. h. nach geeigneter Einschränkung des Definitionsbereichs auf eine offene Umgebung eines Punktes) erwarten, eine bijektive Abbildung zu erhalten. Dass dieses Phänomen erst bei einer Raumdimension $n > 1$ sichtbar wird, hat letztlich topologische Gründe: Wählen wir in unserem Beispiel einen zu großen Definitionsbereich wie z. B. $U = (1, 2) \times (0, 3\pi)$, so würde das Bild von $f|_U$ den Kreisring mit innerem Radius 1 und äußerem Radius 2 eineinhalbmal durchlaufen, so dass f dort dann nicht mehr injektiv wäre – für solche „Schleifen“ um den Nullpunkt herum ist auf der eindimensionalen reellen Zahlengeraden aber „nicht genug Platz“.

Nach diesen Vorbemerkungen wollen wir nun den technisch etwas aufwendigen, aber sicher nicht mehr unerwarteten Satz beweisen, dass eine stetig differenzierbare Abbildung mit invertierbarer Ableitungsmatrix stets lokal umkehrbar ist. Wir beginnen dazu mit dem folgenden Lemma, in dem wir einige spezielle Koordinatenwahlen getroffen haben, um den Beweis einfacher zu halten. Es enthält die eigentliche technische Arbeit, die für die Sätze in diesem Kapitel erforderlich ist – die weiteren Aussagen werden sich daraus dann durch geeignete Koordinatentransformationen als Anwendungen ergeben.

Lemma 27.5 (Lemma über lokale Umkehrfunktionen). *Es seien $D \subset \mathbb{R}^n$ offen mit $0 \in D$, und $f: D \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung mit $f(0) = 0$ und $f'(0) = E_n$.*

Dann ist die Funktion f lokal um den Ursprung umkehrbar, d. h. es gibt offene Umgebungen U und V von $0 \in \mathbb{R}^n$ mit $U \subset D$, so dass die Einschränkung $f|_U: U \rightarrow V$ bijektiv ist. Darüber hinaus ist die dann existierende Umkehrfunktion $f^{-1}: V \rightarrow U$ ebenfalls differenzierbar in 0 mit Ableitung E_n .

Beweis. Die Beweisidee dieses Lemmas besteht darin, Urbilder unter f als Fixpunkte einer geeigneten Hilfsfunktion umzuschreiben, und diese Fixpunkte dann mit Hilfe des Banachschen Fixpunktsatzes 24.27 zu untersuchen. Wir verwenden im Beweis durchgehend die Maximumsnorm auf \mathbb{R}^n und die davon induzierte Zeilensummennorm auf $\mathbb{R}^{n \times n}$ (siehe Beispiel 24.40 (b)).

Da D nach Voraussetzung eine offene Umgebung von 0 ist und f' stetig mit $f'(0) = E$, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(0) \subset D$, so dass

$$\|E - f'(x)\| < \frac{1}{2} \quad (*)$$

für alle $x \in U_\varepsilon(0)$ gilt. Es sei nun $y \in U_{\frac{\varepsilon}{4}}(0)$ zunächst fest. Um Urbilder von y unter f zu suchen, betrachten wir die (von y abhängige) Hilfsfunktion

$$\varphi: D \rightarrow \mathbb{R}^n, \quad x \mapsto x - f(x) + y,$$

so dass $f(x) = y$ genau dann gilt wenn $\varphi(x) = x$, die Urbilder von y unter f also genau die Fixpunkte von φ sind. Um den Banachschen Fixpunktsatz anwenden zu können, zeigen wir nun, dass φ auf der abgeschlossenen Kugel $K_{\frac{\varepsilon}{2}}(0) \subset D$ eine Kontraktion ist:

- (a) Da $K_{\frac{\varepsilon}{2}}(0)$ in der offenen und nach Beispiel 24.22 (c) konvexen Menge $U_\varepsilon(0) \subset D$ liegt, gilt für alle $x, \tilde{x} \in K_{\frac{\varepsilon}{2}}(0)$

$$\|\varphi(x) - \varphi(\tilde{x})\| \stackrel{26.19}{\leq} \|\varphi'|_{\overline{K_{\frac{\varepsilon}{2}}(0)}} \cdot \|x - \tilde{x}\| = \|(E - f')|_{\overline{K_{\frac{\varepsilon}{2}}(0)}} \cdot \|x - \tilde{x}\| \stackrel{(*)}{\leq} \frac{1}{2} \|x - \tilde{x}\|.$$

- (b) Für alle $x \in K_{\frac{\varepsilon}{2}}(0)$ gilt

$$\|\varphi(x)\| \leq \|\varphi(x) - \varphi(0)\| + \|\varphi(0)\| \stackrel{(a)}{\leq} \frac{1}{2} \|x\| + \|y\| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2},$$

und damit auch $\varphi(x) \in K_{\frac{\varepsilon}{2}}(0)$.

Weil $K_{\frac{\varepsilon}{2}}(0)$ als abgeschlossene Teilmenge von \mathbb{R}^n nach Satz 23.29 und Folgerung 23.43 vollständig ist, hat φ nach Satz 24.27 dort also genau einen Fixpunkt, d. h. y besitzt in $K_{\frac{\varepsilon}{2}}(0)$ genau ein Urbild unter f . Da für $\varphi(x) = x$ aus (b) außerdem $\|x\| < \frac{\varepsilon}{2}$ folgt, liegt dieses Urbild sogar in der offenen Kugel $U_{\frac{\varepsilon}{2}}(0)$. Insgesamt ist damit die eingeschränkte Abbildung

$$f: \underbrace{U_{\frac{\varepsilon}{2}}(0) \cap f^{-1}(U_{\frac{\varepsilon}{4}}(0))}_{=:U} \rightarrow \underbrace{U_{\frac{\varepsilon}{4}}(0)}_{=:V}$$

bijektiv. Da U und V nach Lemma 23.34 (a) und Satz 24.17 (b) offen sind, zeigt dies die behauptete Existenz einer lokalen Umkehrabbildung.

Es bleibt also nur noch zu zeigen, dass f^{-1} im Nullpunkt mit Ableitung E differenzierbar ist. Wir prüfen dies wie in Algorithmus 25.15 nach: Ist $y = f(x)$, also $\varphi(x) = x$, so folgt $\|x\| \leq \frac{1}{2}\|x\| + \|y\|$ und damit $\|x\| \leq 2\|y\|$ aus (b), und damit wie gewünscht

$$\frac{f^{-1}(y) - f^{-1}(0) - Ey}{\|y\|} = \frac{x - f(x)}{\|y\|} = - \underbrace{\frac{f(x) - f(0) - Ex}{\|x\|}}_{\rightarrow 0} \cdot \underbrace{\frac{\|x\|}{\|y\|}}_{\leq 2} \rightarrow 0$$

für $y \rightarrow 0$, da mit $y \rightarrow 0$ wegen $\|x\| \leq 2\|y\|$ auch $x \rightarrow 0$ folgt und f in 0 differenzierbar mit Ableitung E ist. \square