27. Implizite Funktionen

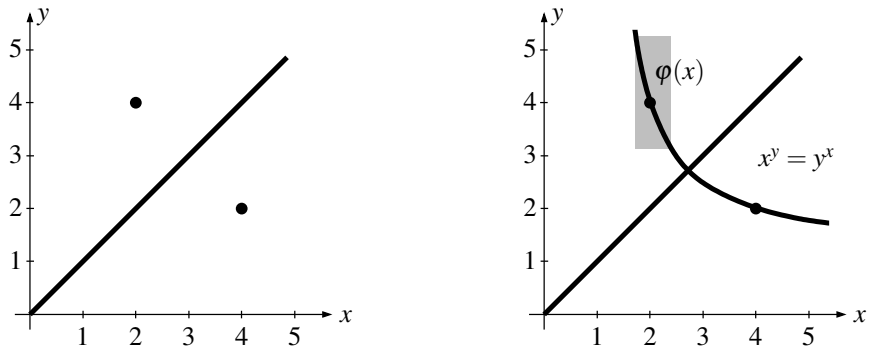
Bevor wir uns ab dem nächsten Kapitel der Integration widmen, wollen wir zum Abschluss unseres Studiums differenzierbarer Abbildungen noch auf das in der Praxis sehr wichtige Thema der sogenannten impliziten Funktionen eingehen, bei dem es um die Auflösbarkeit von Gleichungen nach bestimmten Variablen 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.

Diese Gleichung lässt sich mit den uns bekannten speziellen Funktionen aus Kapitel 9 nicht nach x oder y auflösen, da beide Variablen sowohl in der Basis als auch im Exponenten auftreten. Wir sehen allerdings schon einmal die Lösungen

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

also 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 27.13 (b)).

27.A Umkehrfunktionen

Wir beginnen unser Studium impliziter Funktionen mit einem wichtigen Spezialfall: Sind $D \subset \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}^n$ 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 $f\colon D \to \mathbb{R}$ 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 8.20 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 $x = f^{-1}(y)$.

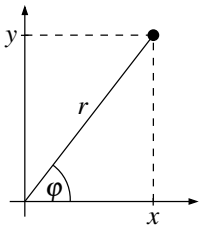
Bemerkung 27.3 (Ableitung einer Umkehrfunktion)**.** Die Formel für die Ableitung der Umkehrfunktion lässt sich sofort auf den mehrdimensionalen Fall verallgemeinern: Ist $f\colon D \to D'$ eine differenzierbare Funktion zwischen offenen Teilmengen D und D' von $\mathbb{R}^n$, und wissen wir bereits, dass f bijektiv ist und eine ebenfalls differenzierbare Umkehrfunktion $f^{-1}\colon D' \to D$ 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 $\mathbb{C} = \mathbb{R}^2$

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



(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

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

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 $f|_U\colon U \to V$ nun bijektiv ist und damit eine Umkehrabbildung $f^{-1}\colon V \to U$ besitzt. In der Tat können wir diese Umkehrabbildung auch sofort aus dem

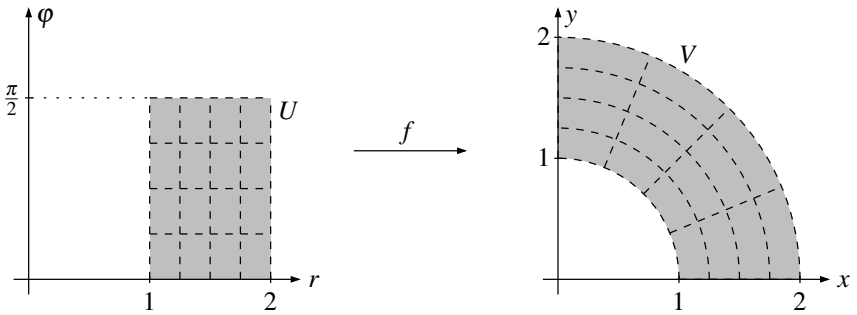
geometrischen Bild oben rechts ablesen: Man kann r und φ in diesem Winkelbereich offensichtlich mit den (stetig differenzierbaren) Formeln

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

aus x und y zurückgewinnen, so dass also

$$f^{-1}\colon V \to U,\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{x^2+y^2} \\ \arctan\frac{y}{x} \end{pmatrix} \tag{2}$$

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 $f\colon D \to \mathbb{R}^n$ 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 $f|_U\colon U \to V$ bijektiv ist. Darüber hinaus ist die dann existierende Umkehrfunktion $f^{-1}\colon V \to U$ 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 (c)).

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

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

für alle $x \in \mathbb{R}^n$ mit $\|x\| < \varepsilon$ gilt. Es sei nun $y \in \mathbb{R}^n$ mit $\|y\| < \frac{\varepsilon}{4}$ zunächst fest. Um Urbilder von y unter f zu suchen, betrachten wir die (von y abhängige) Hilfsfunktion

$$\varphi\colon D \to \mathbb{R}^n,\ x \mapsto x - f(x) + y,$$

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 $x, \tilde{x} \in K_{\frac{\varepsilon}{2}}(0)$

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

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

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

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

Da $K_{\frac{\varepsilon}{2}}(0)$ 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

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

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

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

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

Satz 27.6 (Satz über lokale Umkehrfunktionen). *Es seien $D \subset \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Weiterhin sei $a \in D$ ein Punkt, an dem* $\det f'(a) \neq 0$ *ist.*

Dann ist f lokal um den Punkt $(a, f(a))$ umkehrbar, d. h. es gibt offene Umgebungen $U \subset D$ von a und $V \subset \mathbb{R}^n$ von $f(a)$, so dass die Einschränkung $f|_U\colon U \to V$ bijektiv ist. Darüber hinaus ist die dann existierende Umkehrfunktion $f^{-1}\colon V \to U$ ebenfalls stetig differenzierbar mit Ableitung

$$(f^{-1})'(f(x)) = (f'(x))^{-1} \quad \textit{für alle } x \in U.$$

Beweis. Wir führen den Beweis in mehreren Reduktionsschritten auf Lemma 27.5 zurück.

(a) Indem wir D evtl. auf die nach Beispiel 24.18 offene Umgebung $\{x \in D : \det f'(x) \neq 0\}$ von a verkleinern, können wir annehmen, dass die Voraussetzung $\det f' \neq 0$ des Satzes nicht nur bei a, sondern sogar an jedem Punkt von D gilt.

(b) Es genügt zu zeigen, dass eine lokale Umkehrfunktion $f^{-1}\colon V \to U$ existiert und im Punkt $f(a)$ differenzierbar ist: Nach (a) können wir dieses Ergebnis dann auf jeden Punkt $x \in D$ anwenden und erhalten, dass f^{-1} sogar auf ganz V differenzierbar ist. Die Formel $(f^{-1})'(f(x)) = (f'(x))^{-1}$ für die Ableitung der Umkehrfunktion hatten wir dann bereits in Bemerkung 27.3 hergeleitet. Mit $y = f(x)$ ist also $(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$, und da dieser Ausdruck aus stetigen Funktionen zusammengesetzt ist (siehe Aufgabe 24.15 (b)), ist f^{-1} wie behauptet auch stetig differenzierbar.

(c) Nach geeigneter Koordinatentransformation im Start- und Zielraum können wir $a = f(a) = 0$ annehmen: Führen wir die neuen Koordinaten $\tilde{x} = x - a$ und $\tilde{y} = y - f(a)$ ein und finden wir von der entsprechend transformierten Funktion $\tilde{f}\colon \tilde{x} \mapsto \tilde{y} = f(x) - f(a) = f(\tilde{x} + a) - f(a)$ eine lokale Umkehrfunktion $\tilde{g}$ um $\tilde{x} = 0$, so dass also $\tilde{g}(\tilde{y}) = \tilde{x}$ gilt, so ist natürlich durch $g(y) := \tilde{g}(y - f(a)) + a = x$ eine lokale Umkehrfunktion von f um $x = a$ gegeben.

(d) Durch eine weitere Koordinatentransformation im Zielraum können wir zusätzlich $f'(0) = E$ annehmen: Mit $A := f'(0)$ betrachten wir die Funktion $\tilde{f}\colon x \mapsto \tilde{y} := A^{-1} \cdot y = A^{-1} \cdot f(x)$. Diese hat im Nullpunkt nach der Kettelregel die Ableitung $\tilde{f}'(0) = A^{-1} \cdot f'(0) = E$ – und wenn sie um 0 eine lokale Umkehrfunktion $\tilde{g}$ besitzt, so dass also $\tilde{g}(\tilde{y}) = \tilde{g}(A^{-1} \cdot f(x)) = x$ gilt, so ist damit $g(y) := \tilde{g}(A^{-1} \cdot y) = \tilde{g}(\tilde{y}) = x$ eine lokale Umkehrfunktion um 0 von f.

Die Behauptung des Satzes folgt damit aus Lemma 27.5. □

Aufgabe 27.7. Überprüfe, ob die Abbildung $f\colon \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$, $A \mapsto A^2$ lokal um $(A, f(A))$ mit

$$\text{(a) } A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad \text{(b) } A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

umkehrbar ist. Ist das Bild von f eine Umgebung der Einheitsmatrix $E \in \mathbb{R}^{2\times 2}$?

Aufgabe 27.8. Es seien $U \subset \mathbb{R}^n$ offen und beschränkt sowie $f\colon \overline{U} \to \mathbb{R}^n$ eine stetige Funktion, die auf U stetig differenzierbar ist. Ferner sei $y \in \mathbb{R}^n$ ein Punkt, so dass jeder Urbildpunkt $x \in f^{-1}(y)$ in U liegt und $\det f'(x) \neq 0$ erfüllt. Zeige, dass es dann nur endlich viele solcher Urbilder gibt.

Aufgabe 27.9. Es seien $D \subset \mathbb{R}^n$ eine offene Menge und $f\colon D \to \mathbb{R}^n$ eine stetig differenzierbare Funktion, so dass $f'(x)$ für jedes $x \in D$ invertierbar ist.

Zeige, dass die Funktion $g\colon D \to \mathbb{R}$, $x \mapsto \|f(x)\|$ für keine Wahl der verwendeten Norm ein lokales Maximum besitzt. 70

27.B Der Satz über implizite Funktionen

Nach dem Studium von Umkehrfunktionen kommen wir nun wieder auf die impliziten Funktionen wie in Beispiel 27.1 zurück. Im Sinne der mehrdimensionalen Analysis wollen wir dabei aber nicht nur *eine* Gleichung nach *einer* der darin vorkommenden Variablen auflösen können, sondern n gegebene Gleichungen nach n darin enthaltenen Variablen. Für den folgenden Satz fixieren wir dazu die folgenden Notationen: Auf einer offenen Teilmenge $D \subset \mathbb{R}^r \times \mathbb{R}^n$ mit Koordinaten $u = (u_1, \ldots, u_r)$ (von $\mathbb{R}^r$) und $v = (v_1, \ldots, v_n)$ (von $\mathbb{R}^n$) betrachten wir eine stetig differenzierbare Abbildung $f\colon D \to \mathbb{R}^n$, deren Ableitungsmatrix also die Form

$$f' = \left(\frac{\partial f}{\partial u_1} \,\middle|\, \cdots \,\middle|\, \frac{\partial f}{\partial u_r} \,\middle|\, \frac{\partial f}{\partial v_1} \,\middle|\, \cdots \,\middle|\, \frac{\partial f}{\partial v_n} \right) =: \left(\frac{\partial f}{\partial u} \,\middle|\, \frac{\partial f}{\partial v} \right)$$

hat, wobei die $n \times r$-Matrix $\frac{\partial f}{\partial u}$ die ersten r Spalten und die $n \times n$-Matrix $\frac{\partial f}{\partial v}$ die letzten n Spalten von f' bezeichnet. Wir wollen untersuchen, ob wir die n Gleichungen $f(u, v) = 0 \in \mathbb{R}^n$ in den $r + n$ Variablen u und v in einer Umgebung eines Punktes $(a, b) \in D$ mit $f(a, b) = 0$ nach den n Variablen v auflösen können. In Beispiel 27.1 wäre also $n = r = 1$ und $f\colon \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{R}$, $(u, v) \mapsto u^v - v^u$, und z. B. $(a, b) = (2, 4)$.

Der Satz über implizite Funktionen, den wir jetzt aus dem Satz über lokale Umkehrfunktionen herleiten wollen, besagt in dieser Situation gerade, dass eine solche lokale Auflösung nach v immer

möglich ist, wenn die Ableitungsmatrix $\frac{\partial f}{\partial v}$ an der betrachteten Stelle invertierbar ist. Als zusätzliche Aussage können wir in diesem Fall auch eine Formel für die Ableitung dieser Auflösungsfunktion angeben.

Satz 27.10 (Satz über implizite Funktionen). *Es seien $D \subset \mathbb{R}^r \times \mathbb{R}^n$ eine offene Teilmenge und $f\colon D \to \mathbb{R}^n$, $(u,v) \mapsto f(u,v)$ eine stetig differenzierbare Abbildung. Ferner sei $(a,b) \in D$ ein Punkt mit $f(a,b) = 0$ und $\det \frac{\partial f}{\partial v}(a,b) \neq 0$.*

Dann lässt sich die Gleichung $f(u,v) = 0$ lokal um (a,b) nach v auflösen, d. h. es gibt offene Umgebungen U von $a \in \mathbb{R}^r$ und V von $b \in \mathbb{R}^n$ mit $U \times V \subset D$ sowie eine stetig differenzierbare Funktion $\varphi\colon U \to V$ mit $\varphi(a) = b$, so dass für alle $u \in U$ und $v \in V$ gilt, dass

$$f(u,v) = 0 \quad \textit{genau dann, wenn} \quad v = \varphi(u).$$

Beweis. Wir betrachten die Hilfsfunktion $F\colon D \to \mathbb{R}^r \times \mathbb{R}^n$ mit $F(u,v) = (u, f(u,v))$. Wegen

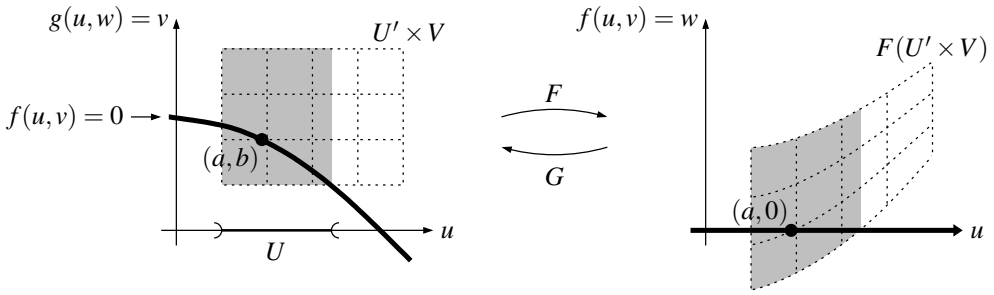
$$F' = \begin{pmatrix} E_r & 0 \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix}, \quad \text{also} \quad \det F'(a,b) = \det E_r \cdot \det \frac{\partial f}{\partial v}(a,b) \neq 0$$

können wir den Satz 27.6 über lokale Umkehrfunktionen auf F an der Stelle (a,b) anwenden und erhalten offene Umgebungen von (a,b) und $F(a,b) = (a,0)$, zwischen denen F bijektiv ist und eine stetig differenzierbare Umkehrfunktion G besitzt. Weil die Umkehrung der ersten Komponente von F, also der Identität, natürlich trivial ist, hat G notwendigerweise die Form $G(u,w) = (u, g(u,w))$ mit

$$g(u, f(u,v)) = v \quad \text{und} \quad f(u, g(u,w)) = w \tag{$*$}$$

für alle u, v, w in den gefundenen Umgebungen, wobei $G(a,0) = (a,b)$ und damit $g(a,0) = b$ gilt.

Durch evtl. Verkleinern können wir erreichen, dass die offene Umgebung von (a,b) von der Form $U' \times V$ mit offenen Umgebungen U' von a und V von b ist (wähle z. B. eine offene Kugel in der Maximumsnorm). Die entsprechende Situation ist dann im folgenden Bild dargestellt, wobei F und G die beiden gestrichelt umrandeten Bereiche bijektiv aufeinander abbilden und die dick eingezeichneten Kurven die gesuchten Stellen mit $f = 0$ markieren. Beachte, dass die erste Komponente von F und G die Identität ist und die vertikalen gestrichelten Linien damit wieder auf vertikale Linien an der gleichen Stelle u abgebildet werden.



Wie im Bild grau markiert verkleinern wir U' nun auf die Menge $U = \{u \in U' : (u,0) \in F(U' \times V)\}$, die wegen $(a,0) = F(a,b)$ immer noch den Punkt a enthält. Als Urbild der offenen Menge $F(U' \times V)$ unter der stetigen Abbildung $u \mapsto (u,0)$ ist U nach Satz 24.17 (b) außerdem offen, also eine offene Umgebung von a. Nach Konstruktion sind dann G und damit auch g auf den Punkten $(u,0)$ mit $u \in U$ definiert, und die Funktion

$$\varphi\colon U \to V,\ u \mapsto g(u,0)$$

leistet das Gewünschte, denn es ist $\varphi(a) = g(a,0) = b$, und nach $(*)$ außerdem $f(u,v) = 0$ genau dann, wenn $v = g(u,0)$. $\square$

Folgerung 27.11 (Ableitung der Auflösungsfunktion)**.** *In der Situation und mit den Notationen von Satz 27.10 ist die Ableitung der Auflösungsfunktion φ für alle $u \in U$ gegeben durch*

$$\varphi'(u) = -\left(\frac{\partial f}{\partial v}(u,v)\right)^{-1} \cdot \frac{\partial f}{\partial u}(u,v)$$

mit $v = \varphi(u)$.

Beweis. Nach Konstruktion der Auflösungsfunktion φ gilt $f(u,\varphi(u)) = 0$ für alle $u \in U$. Differenzieren dieser Gleichung nach u ergibt mit der Kettenregel aus Satz 25.30

$$0 = \left(\frac{\partial f}{\partial u} \,\middle|\, \frac{\partial f}{\partial v}\right) \cdot \begin{pmatrix} E_r \\ \varphi' \end{pmatrix} = \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v} \cdot \varphi',$$

woraus durch Auflösen nach φ' die behauptete Formel folgt. □

Bemerkung 27.12. In Satz 27.10 haben wir die Variablen $v_1, \ldots, v_n$, nach denen die gegebene Gleichung $f(u,v) = 0$ aufgelöst werden soll, mit anderen Buchstaben bezeichnet als die übrigen Variablen $u_1, \ldots, u_r$. Für die Formulierung und den Beweis des Satzes war dies sehr praktisch – in konkreten Anwendungen werden wir dies aber in der Regel nicht tun. Die Aussage von Satz 27.10 ist dann anschaulich formuliert: Haben wir n Gleichungen $f(x_1, \ldots, x_m) = 0$ in $m \geq n$ Variablen gegeben und können wir n Spalten aus der Ableitungsmatrix $f' \in \mathbb{R}^{n \times m}$ so auswählen, dass die resultierende quadratische Matrix an einem gegebenen Punkt der Lösungsmenge invertierbar ist, so lässt sich die Gleichung $f(x) = 0$ lokal um diesen Punkt nach den entsprechenden Variablen auflösen.

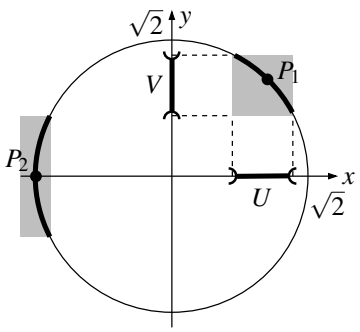
Beispiel 27.13.

(a) Zum besseren Verständnis beginnen wir mit einem Beispiel, bei dem wir die Auflösungsfunktion bereits kennen und uns Satz 27.10 daher nichts Neues sagt: Wir betrachten eine einfache Kreisgleichung $x^2 + y^2 = 2$, d. h. die Gleichung $f(x,y) = 0$ mit

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

und damit $f'(x,y) = (2x \quad 2y)$.

Der Punkt $P_1 = (1,1)$ erfüllt offensichtlich die Gleichung $f(x,y) = 0$, und es ist $\frac{\partial f}{\partial y}(P_1) = 2 \neq 0$. Nach Satz 27.10 bzw. Bemerkung 27.12 (hier haben wir lediglich *eine* Gleichung, daher ist die Determinante der 1×1-Matrix $\frac{\partial f}{\partial y}(P_1)$ einfach gleich dieser Zahl selbst) kann man die gegebene Gleichung also lokal um P_1 stetig differenzierbar nach y auflösen – was wir natürlich schon vorher wussten, denn in einer Umgebung $U \times V$ von P_1 ist $f(x,y) = 0$ ja einfach äquivalent zu $y = \sqrt{2 - x^2}$.



Am Punkt $P_2 = (-\sqrt{2}, 0)$ hingegen ist $\frac{\partial f}{\partial y}(P_2) = 0$, hier macht der Satz über implizite Funktionen also keine Aussage über die lokale Auflösbarkeit der Gleichung $f(x,y) = 0$ nach y. In der Tat können wir y auch in keiner Umgebung von P_2 als Funktion von x ausdrücken: In jeder Umgebung gibt es für x in der Nähe von $-\sqrt{2}$ ja immer zwei oder gar keinen möglichen Wert von y.

Es ist jedoch $\frac{\partial f}{\partial x}(P_2) = -2\sqrt{2} \neq 0$, und daher können wir analog zum Fall von P_1 oben die Gleichung lokal um P_2 zumindest nach x auflösen, also x als Funktion von y schreiben – auch das wussten wir bereits vorher, die auflösende Gleichung ist hier $x = -\sqrt{2 - y^2}$.

(b) Wie in Beispiel 27.1 betrachten wir noch einmal die Gleichung $f(x,y) := x^y - y^x = 0$ für $x, y \in \mathbb{R}_{>0}$. Wir hatten dort bereits vermutet, dass sich diese Gleichung lokal um den Punkt

$(2,4)$ nach y auflösen lässt, also dass sich die Lösungsmenge als Graph einer (stetig differenzierbaren) Funktion $\varphi\colon x \mapsto y = \varphi(x)$ schreiben lässt. Dies können wir nun mit Satz 27.10 bestätigen, denn es ist

$$\frac{\partial f}{\partial y}(x,y) = x^y \log x - x y^{x-1}, \quad \text{und damit} \quad \frac{\partial f}{\partial y}(2,4) = 16\log 2 - 8 \neq 0.$$

Auch wenn wir die Funktion φ nicht explizit hinschreiben können, können wir mit Folgerung 27.11 noch ihre Ableitung an dieser Stelle berechnen: Wegen

$$\frac{\partial f}{\partial x}(x,y) = y x^{y-1} - y^x \log y, \quad \text{und damit} \quad \frac{\partial f}{\partial x}(2,4) = 32 - 16\log 4$$

ist die Steigung des Graphen am Punkt $(2,4)$

$$\varphi'(2) = -\left(\frac{\partial f}{\partial y}(2,4)\right)^{-1} \cdot \frac{\partial f}{\partial x}(2,4) = -\frac{32 - 16\log 4}{16\log 2 - 8} = -3{,}177\ldots\ .$$

Darüber hinaus haben wir am Bild in Beispiel 27.1 gesehen, dass die durch die Gleichung $f(x,y) = 0$ gegebene Punktmenge aus zwei Zweigen besteht, die sich in einem Punkt schneiden. Wiederum mit Hilfe von Satz 27.10 können wir nun auch die exakte Position des Kreuzungspunktes dieser beiden Zweige berechnen: Dort kann die Gleichung $f(x,y) = 0$ ja nicht lokal nach y auflösbar sein, und daher muss an diesem Punkt $\frac{\partial f}{\partial y} = 0$ sein. Da für diesen Punkt außerdem offensichtlich $y = x$ gilt, ergibt sich mit der obigen Formel für die Ableitung von f am Kreuzungspunkt also

$$x x^{x-1} - x^x \log x = 0, \quad \text{d. h.} \quad x^x(1 - \log x) = 0$$

und damit $x = \mathrm{e}$. Der Kreuzungspunkt liegt also bei (e,e).

(c) (Nullstellen von Polynomen) Für ein gegebenes $n \in \mathbb{N}$ sei

$$F\colon \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R},\ (u,x) \mapsto f_u(x) := u_n x^n + \cdots + u_1 x + u_0,$$

wobei $u_0, \ldots, u_n$ die Koordinaten von $u \in \mathbb{R}^{n+1}$ sind. Wir können F also als Polynomfunktion in x auffassen, wobei wir aber auch die Koeffizienten u des Polynoms als variabel ansehen.

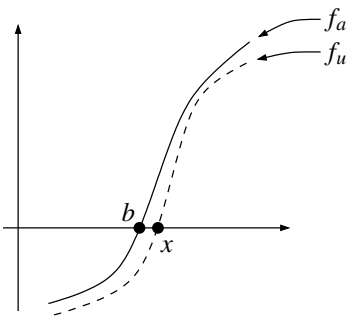
Natürlich gilt dann

$$\frac{\partial F}{\partial x}(u,x) = n u_n x^{n-1} + \cdots + u_1 = f_u'(x).$$

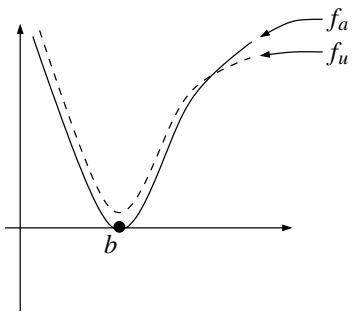
Ist also $(a,b) \in \mathbb{R}^{n+1} \times \mathbb{R}$ mit

$$F(a,b) = f_a(b) = 0 \quad \text{und} \quad \frac{\partial F}{\partial x}(a,b) = f_a'(b) \neq 0$$

gegeben, d. h. ist b eine *einfache* Nullstelle von f_a, so besagt Satz 27.10 gerade, dass es in einer Umgebung von (a,b) eine stetig differenzierbare Funktion $\varphi\colon u \mapsto x = \varphi(u)$ gibt, so dass x dort genau dann eine Nullstelle des Polynoms f_u ist, also $F(u,x) = f_u(x) = 0$ gilt, wenn $x = \varphi(u)$ gilt. Wir können φ für die betrachtete Polynomgleichung also lokal als eine Lösungsformel ansehen, die den Koeffizienten u eine Nullstelle von f_u zuordnet. Da diese Funktion stetig differenzierbar ist, sagt man auch: *Die (einfachen) Nullstellen eines Polynoms hängen stetig differenzierbar von seinen Koeffizienten ab.* Dies ist im Bild unten links dargestellt: Eine kleine Änderung im Polynom von f_a nach f_u hat auch nur eine kleine Änderung der Nullstelle von b nach x zur Folge.



b ist einfache Nullstelle von f_a
Nullstelle x variiert stetig differenzierbar mit u



b ist Nullstelle höherer Ordnung von f_a
keine Nullstelle von f_u in der Nähe von b

Ist hingegen $\frac{\partial F}{\partial x}(a,b) = f_a'(b) = 0$, also b eine Nullstelle höherer Ordnung von f_a, so ist der Satz über implizite Funktionen nicht wie oben anwendbar. In der Tat zeigt hier auch das Bild oben rechts, dass es in einer kleinen Umgebung von a, also für ein Polynom f_u mit u in der Nähe von a, in der Regel keine (eindeutige) Nullstelle von f_u in der Nähe von b gibt.

Aufgabe 27.14. Zeige, dass das Gleichungssystem

$$\mathrm{e}^{xz} - x^2 + y^2 = xy^3 + x^2z + yz^2 = 1$$

in einer Umgebung des Punktes $x_0 = 1$, $y_0 = 1$, $z_0 = 0$ nach y und z aufgelöst, also als $\begin{pmatrix} y \\ z \end{pmatrix} = \varphi(x)$ für eine stetig differenzierbare Funktion φ geschrieben werden kann. Berechne auch die Ableitung $\varphi'(1)$ an diesem Punkt!

Aufgabe 27.15. Es seien $f\colon \mathbb{R}^3 \to \mathbb{R}$ eine stetig differenzierbare Funktion und $a \in \mathbb{R}^3$ ein Punkt mit $f(a) = 0$, an dem alle drei partiellen Ableitungen $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ und $\frac{\partial f}{\partial z}$ ungleich 0 sind.

Nach dem Satz über implizite Funktionen bestimmt die Gleichung $f(x,y,z) = 0$ dann in einer Umgebung von a jede der drei Variablen aus den beiden anderen. Wir können diese Gleichung dort also z. B. als $x = \varphi(y,z)$ für eine stetig differenzierbare Funktion φ schreiben. Gemäß der bekannten Differentialschreibweise bezeichnen wir die partielle Ableitung dieser Auflösungsfunktion φ nach y mit $\frac{\partial x}{\partial y}$, analog für die anderen Variablenkombinationen.

Zeige, dass dann die (etwas merkwürdig erscheinende) Gleichung $\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = -1$ gilt.

Aufgabe 27.16. Die *Lemniskate* L ist definiert als die Menge aller Punkte in $\mathbb{R}^2$, für die das Produkt der (euklidischen) Abstände zu den beiden Punkten $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ gleich 1 ist. Zeige, dass L in einer Umgebung des Punktes $\begin{pmatrix} \frac{2}{3} \\ \frac{1}{3}\sqrt{2} \end{pmatrix} \in L$ als Graph einer differenzierbaren Funktion $x_2 = \varphi(x_1)$ geschrieben werden kann.

Aufgabe 27.17. Berechne alle isolierten Punkte der Menge $M = \{(x,y) \in \mathbb{R}^2 : x^2 + x^3 + y^2 + y^3 = 0\}$.

Aufgabe 27.18. Für $a \in \mathbb{R}$ betrachten wir die Funktion

$$f_a\colon \mathbb{R}^2 \to \mathbb{R},\ f_a(x,y) = x^2 + y^2 + a\,\mathrm{e}^{x+2y}.$$

Für $a = 0$ hat diese Funktion offensichtlich ein (lokales und globales) Minimum im Nullpunkt.

(a) Zeige, dass es eine auf einer Umgebung $U \subset \mathbb{R}$ von 0 definierte und stetig differenzierbare Funktion $\varphi\colon U \to \mathbb{R}^2$ gibt, so dass für alle $a \in U$ die Funktion f_a ein lokales Minimum in $\varphi(a)$ hat.

(b) Berechne die Ableitung $\varphi'(0)$.

71

Wir haben nun also gesehen, dass man ein gegebenes Gleichungssystem $f(x) = 0$ mit einer stetig differenzierbaren Funktion $f\colon D \to \mathbb{R}^n$ auf $D \subset \mathbb{R}^m$ mit $m \geq n$ unter Umständen lokal nach n der Variablen $x_1, \ldots, x_m$ auflösen kann. Ob und nach welchen Variablen dies möglich ist, hängt dabei natürlich in der Regel wie in Beispiel 27.13 (a) nicht nur von f, sondern auch vom betrachteten Punkt ab. Ist eine solche Auflösung in *jedem* Punkt möglich – wenn auch je nach Punkt nach unterschiedlichen Variablen – so hat dies für die Lösungsmenge der Gleichung $f(x) = 0$ eine besondere geometrische Bedeutung, die wir jetzt kurz untersuchen wollen.

Definition 27.19 (Untermannigfaltigkeiten)**.** Es seien $r, n \in \mathbb{N}$ und $m := r + n$. Eine Teilmenge M von $\mathbb{R}^m$ heißt **r-dimensionale Untermannigfaltigkeit** von $\mathbb{R}^m$, wenn es zu jedem $c \in M$ eine offene Umgebung W und eine stetig differenzierbare Funktion $f\colon W \to \mathbb{R}^n$ gibt mit $\operatorname{rk} f'(c) = n$, so dass $M \cap W = \{x \in W : f(x) = 0\}$.

Beispiel 27.20. In Definition 27.19 darf die Funktion f von dem betrachteten Punkt $c \in M$ abhängen. Oft ist dies in der Praxis jedoch gar nicht der Fall, und es gibt eine (globale) stetig differenzierbare Funktion $f\colon \mathbb{R}^m \to \mathbb{R}^n$, so dass $M = \{x \in \mathbb{R}^m : f(x) = 0\}$ deren Nullstellenmenge ist und $\operatorname{rk} f'(c) = n$ an jedem Punkt $c \in M$ gilt.

So ist z. B. die Kreislinie $M = \{(x,y) \in \mathbb{R}^2 : f(x,y) = 0\}$ mit $f(x,y) = x^2 + y^2 - 2$ aus Beispiel 27.13 (a) eine 1-dimensionale Untermannigfaltigkeit von $\mathbb{R}^2$, denn die Ableitungsmatrix $f' = (2x \;\; 2y)$ hat an jedem Punkt $(x,y) \in M$ Rang 1 (beachte, dass der einzige Punkt $(x,y) = (0,0)$, an dem $\operatorname{rk} f' = 0$ gilt, nicht zu M gehört).

Lemma 27.21 (Untermannigfaltigkeiten als Graphen)**.** *Es seien $r, n \in \mathbb{N}$ und $m := r + n$. Für eine Teilmenge $M \subset \mathbb{R}^m$ sind äquivalent:*

(a) *M ist eine r-dimensionale Untermannigfaltigkeit.*

(b) *Zu jedem $c \in M$ gibt es*

- *eine Aufteilung der m Koordinaten von $\mathbb{R}^m$ in r Koordinaten $u_1, \ldots, u_r$ und n Koordinaten $v_1, \ldots, v_n$, mit entsprechender Aufteilung $c = (a,b)$,*
- *offene Umgebungen U von a in $\mathbb{R}^r$ und V von b in $\mathbb{R}^n$,*
- *eine stetig differenzierbare Funktion $\varphi\colon U \to V$ mit $M \cap (U \times V) = \{(u, \varphi(u)) : u \in U\}$ und $\varphi(a) = b$.*

(Mit anderen Worten lässt sich M also lokal als Graph einer stetig differenzierbaren Funktion in r Variablen schreiben, wobei um jeden Punkt von M unterschiedlich sein darf, welche der m Variablen als Variablen des Startraums und welche als Variable des Zielraums aufgefasst werden.)

Beweis.

(a) $\Rightarrow$ (b): Ist c ein Punkt einer r-dimensionalen Untermannigfaltigkeit M von $\mathbb{R}^m$, so ist M in einer offenen Umgebung W von c die Menge aller Nullstellen einer stetig differenzierbaren Funktion $f\colon W \to \mathbb{R}^n$ mit $\operatorname{rk} f'(c) = n$. Nach Aufgabe 15.22 gibt es in der Ableitungsmatrix $f'(c) \in \mathbb{R}^{n \times m}$ also eine invertierbare $n \times n$-Teilmatrix. Bezeichnen wir die Koordinaten, die diesen Spalten entsprechen, mit $v_1, \ldots, v_n$, und die restlichen mit $u_1, \ldots, u_r$, so ergibt sich die Behauptung (b) direkt aus dem Satz 27.10 über implizite Funktionen.

(b) $\Rightarrow$ (a): Es sei $c \in M$ beliebig. Nach Voraussetzung gibt es dann bei geeigneter Aufteilung $u_1, \ldots, u_r, v_1, \ldots, v_n$ der Koordinaten von $\mathbb{R}^m$ eine offene Umgebung $U \times V$ von c, auf der M gegeben ist durch $v = \varphi(u)$ für eine stetig differenzierbare Funktion $\varphi\colon U \to V$, also durch die Gleichung $f(u,v) = 0$ mit

$$f\colon U \times V \to \mathbb{R}^n,\ (u,v) \mapsto v - \varphi(u).$$

Wegen $\frac{\partial f}{\partial v}(c) = E_n$ ist $\operatorname{rk} f'(c) = n$. Damit ist M nach Definition 27.19 eine r-dimensionale Untermannigfaltigkeit von $\mathbb{R}^m$. $\square$

Bemerkung 27.22 (Eine Untermannigfaltigkeit ist lokal ein gekrümmter $\mathbb{R}^r$)**.** Nach Lemma 27.21 sind die r-dimensionalen Untermannigfaltigkeiten genau die Teilmengen von $\mathbb{R}^m$, die sich lokal als Graph einer stetig differenzierbaren Funktion in r der Variablen von $\mathbb{R}^m$ schreiben lassen. Wie im Eindimensionalen bedeutet dies anschaulich, dass M als Graph solcher stetig differenzierbaren Funktionen „keine Knicke haben kann". Anschaulich können wir also sagen: *Eine r-dimensionale Untermannigfaltigkeit ist ein Raum, der lokal wie ein gekrümmter (aber nicht geknickter) $\mathbb{R}^r$ aussieht* – so wie z. B. die Kreislinie in Beispiel 27.13 (a) bzw. 27.20. Dementsprechend zeigen die ersten beiden Bilder unten 1-dimensionale Untermannigfaltigkeiten von $\mathbb{R}^2$, die anderen beiden wegen der grau markierten Kreuzungs- bzw. Knickpunkte jedoch nicht.

1-dimensionale Untermannigfaltigkeiten keine Untermannigfaltigkeiten

Aufgabe 27.23. Es sei $M = \{x \in \mathbb{R}^2 : x_1^3 = x_2^2\}$. Zeige, dass $M \setminus \{0\}$ eine 1-dimensionale Untermannigfaltigkeit von $\mathbb{R}^2$ ist, M jedoch nicht.

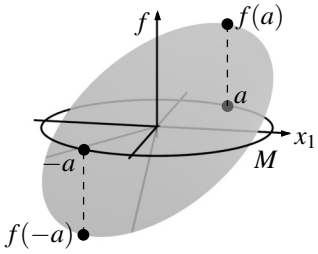
Skizziere auch die Menge M. Könnt ihr das Ergebnis der Aufgabe auch in dieser Skizze erkennen?

27.C Extrema unter Nebenbedingungen

Eine Anwendung von impliziten Funktionen und Untermannigfaltigkeiten, die wir zum Abschluss dieses Kapitels noch betrachten wollen, ist das bereits in Bemerkung 26.4 (b) beschriebene Problem der *Randextrema* bzw. der Extrema unter Nebenbedingungen. Haben wir eine differenzierbare Funktion $f\colon D \to \mathbb{R}$ auf einer offenen Teilmenge $D \subset \mathbb{R}^n$ und eine Untermannigfaltigkeit $M \subset D$, so geht es dabei wie im folgenden Beispiel darum, die (lokalen) Extrema der eingeschränkten Funktion $f|_M$ zu bestimmen.

Beispiel 27.24 (Randextrema durch Einsetzen der Nebenbedingungen)**.** Wir wollen die Extrema der unten rechts dargestellten Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$, $x \mapsto x_1 + 2x_2$ auf der abgeschlossenen Kreisscheibe $K_1(0) = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \le 1\}$ finden.

Da $K_1(0)$ als beschränkte und abgeschlossene Teilmenge von $\mathbb{R}^2$ nach Satz 23.51 (b) kompakt ist, nimmt f dort nach Folgerung 24.35 tatsächlich ein Maximum und Minimum an. Allerdings besitzt f als lineare Funktion keine kritischen Punkte, denn die Ableitung $f'(x,y) = (1\ \ 2)$ ist nirgends gleich Null – und damit kann f nach Lemma 26.2 keine Extrema im Inneren von $K_1(0)$ haben. Maximum und Minimum von f auf $K_1(0)$ werden also notwendigerweise auf dem Rand $M := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ angenommen.



Im Eindimensionalen, also wenn die betrachtete Definitionsmenge ein Intervall ist, besteht der Rand dieses Intervalls ja höchstens aus zwei Punkten, und daher konnten wir derartige Randextrema dann einfach durch Einsetzen aller Randpunkte bestimmen. In unserem momentanen Fall ist dies jedoch nicht mehr möglich, da der Rand M von $K_1(0)$ unendlich viele Punkte hat. Stattdessen müssen wir nun die Extrema der eingeschränkten Funktion $f|_M$, also die Extrema von f unter der Nebenbedingung $x_1^2 + x_2^2 = 1$ suchen.

Die einfachste Möglichkeit dafür ist hier vermutlich, die Nebenbedingung nach einer der Variablen aufzulösen und in die Zielfunktion f einzusetzen: Nehmen wir zunächst $x_2 \ge 0$ an, so ist M dort gegeben durch die aufgelöste Gleichung $x_2 = \sqrt{1 - x_1^2}$. Einsetzen in f zeigt also, dass wir – nun ohne Nebenbedingungen – eine Stelle $x_1 \in [-1, 1]$ suchen, an der die Funktion

$$h(x_1) := f\left(x_1, \sqrt{1 - x_1^2}\right) = x_1 + 2\sqrt{1 - x_1^2}$$

ein Extremum besitzt. Dieses können wir nun leicht mit Hilfe unserer eindimensionalen Techniken finden: Es ist

$$h'(x_1) = 1 - \frac{2x_1}{\sqrt{1-x_1^2}},$$

und die (einzige) Nullstelle dieser Funktion berechnet man schnell zu $a_1 = \frac{1}{\sqrt{5}}$, woraus durch Einsetzen in die Nebenbedingung $a_2 = \frac{2}{\sqrt{5}}$ folgt. Der entsprechende Punkt a ist im Bild rechts oben eingezeichnet, und man sieht leicht (z. B. mit Satz 11.18 durch Untersuchung der zweiten Ableitung von h in diesem Punkt), dass hier ein isoliertes lokales Maximum vorliegt. Untersuchen wir zusätzlich auch noch die Randpunkte von h (also $x_1 = \pm 1$) sowie die Punkte mit $x_2 < 0$, so sehen wir, dass dies sogar das globale Maximum von f auf M (und damit auf $K_1(0)$) ist. Entsprechend findet man das Minimum von f am gegenüberliegenden Punkt $-a$.

Diese Strategie, die Nebenbedingung nach einer der Variablen aufzulösen und in die Zielfunktion einzusetzen, führt aber zum einen wie oben schnell zu Fallunterscheidungen, und funktioniert zum anderen natürlich auch nur, wenn man diese Auflösungsfunktion explizit berechnen kann und das Einsetzen dieser Funktion nicht zu einer zu komplizierten Zielfunktion führt. Andernfalls muss man stattdessen mit dem Satz über implizite Funktionen arbeiten und erhält so das folgende Resultat.

Satz 27.25 (Extrema unter Nebenbedingungen). *Es seien $r,n \in \mathbb{N}$ und $f\colon D \to \mathbb{R}$ eine differenzierbare Funktion auf einer offenen Teilmenge $D \subset \mathbb{R}^m$ mit $m := r+n$. Ferner sei $g\colon D \to \mathbb{R}^n$ stetig differenzierbar mit $\operatorname{rk} g'(c) = n$ für alle $c \in M := \{x \in D : g(x) = 0\}$, so dass M also eine r-dimensionale Untermannigfaltigkeit von $\mathbb{R}^m$ ist.*

Hat dann die Einschränkung $f|_M$ ein lokales Extremum in einem Punkt $c \in M$, so gibt es ein $\lambda \in \mathbb{R}^n$ mit

$$f'(c) + \lambda^\mathsf{T} g'(c) = 0 \quad \in \mathbb{R}^{1\times m},$$

also mit $f'(c) + \lambda_1 g_1'(c) + \dots + \lambda_n g_n'(c) = 0$, wobei $\lambda_1,\dots,\lambda_n$ und $g_1,\dots,g_n$ die Komponenten von λ bzw. g sind.

*Man bezeichnet $\lambda_1,\dots,\lambda_n$ als **Lagrange-Multiplikatoren** und spricht bei der Anwendung dieses Satzes daher auch vom Verfahren der Lagrange-Multiplikatoren.*

Beweis. Da M nach Voraussetzung eine r-dimensionale Untermannigfaltigkeit ist, können wir die Koordinaten x von $D \subset \mathbb{R}^m$ nach Lemma 27.21 lokal um c so in $u_1,\dots,u_r$ und $v_1,\dots,v_n$ aufteilen, dass M dort durch eine aufgelöste Gleichung der Form $v = \varphi(u)$ gegeben ist. Für diese Auflösungsfunktion φ gilt außerdem nach Folgerung 27.11

$$\varphi' = -\left(\frac{\partial g}{\partial v}\right)^{-1} \cdot \frac{\partial g}{\partial u}. \tag{1}$$

Ist c nun ein lokales Extremum von $f|_M$, so hat die Funktion $h\colon u \mapsto f(u,\varphi(u))$ wie in Beispiel 27.24 ein lokales Extremum ohne Nebenbedingungen, und damit gilt dort nach Lemma 26.2 und der Kettenregel aus Satz 25.30

$$0 = h' = \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} E \\ \varphi' \end{pmatrix} = \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v} \cdot \varphi' \overset{(1)}{=} \frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} \cdot \left(\frac{\partial g}{\partial v}\right)^{-1} \cdot \frac{\partial g}{\partial u}. \tag{2}$$

Setzen wir nun

$$\lambda^\mathsf{T} := -\frac{\partial f}{\partial v}(c) \cdot \left(\frac{\partial g}{\partial v}(c)\right)^{-1} \quad \in \mathbb{R}^{1\times n}, \tag{3}$$

so ergibt sich an diesem Punkt c wie behauptet

$$\begin{aligned} f' + \lambda^\mathsf{T} g' &= \left(\frac{\partial f}{\partial u} + \lambda^\mathsf{T} \frac{\partial g}{\partial u} \quad \frac{\partial f}{\partial v} + \lambda^\mathsf{T} \frac{\partial g}{\partial v} \right) \\ &\overset{(3)}{=} \left(\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} \cdot \left(\frac{\partial g}{\partial v} \right)^{-1} \cdot \frac{\partial g}{\partial u} \quad \frac{\partial f}{\partial v} - \frac{\partial f}{\partial v} \cdot \left(\frac{\partial g}{\partial v} \right)^{-1} \cdot \frac{\partial g}{\partial v} \right) \\ &\overset{(2)}{=} (0 \;\; 0). \end{aligned}$$

□

Bemerkung 27.26.

(a) Man kann die Bedingung $f'(c) + \lambda^\mathsf{T} g'(c) = 0$ in Satz 27.25 auch so verstehen: Für einen beliebigen differenzierbaren Weg $\gamma\colon I \to M$ von einem offenen Intervall I nach M ist die Verkettung $g \circ \gamma\colon I \to \mathbb{R}^n$ ja die Nullfunktion, also insbesondere

$$0 = (g \circ \gamma)'(t) = g'(\gamma(t))\gamma'(t) \tag{$*$}$$

für alle $t \in I$. Ist nun $t \in I$ so, dass $f|_M$ in $\gamma(t) = c$ ein lokales Extremum hat, so hat $f \circ \gamma$ insbesondere einen kritischen Punkt in t, und damit gilt auch

$$0 = (f \circ \gamma)'(t) = f'(\gamma(t)) \cdot \gamma'(t).$$

Dies ist mit $(*)$ aber sichergestellt, wenn $f'(c) + \lambda^\mathsf{T} g'(c) = 0$, also $f'(\gamma(t)) = -\lambda^\mathsf{T} g'(\gamma(t))$ gilt.

(b) Die Lagrange-Multiplikatoren $\lambda^\mathsf{T} \in \mathbb{R}^n$ in Satz 27.25 sind n zusätzliche Variablen, deren Werte zunächst unbekannt sind. Mit dem Verfahren des Satzes müssen an einem lokalen Extremum also $m+n$ Variablen bestimmt werden. Da wir zusätzlich zu den m Gleichungen $f'(c) + \lambda^\mathsf{T} g'(c) = 0 \in \mathbb{R}^m$ aber auch noch die n Nebenbedingungen $g(c) = 0 \in \mathbb{R}^n$ haben, kann man dieses (in der Regel nichtlineare) Gleichungssystem oftmals lösen.

(c) Mit den Bezeichnungen von Satz 27.25 können die $m+n$ Gleichungen aus (b) auch gut über die sogenannte *Lagrange-Funktion*

$$L\colon D \times \mathbb{R}^n \to \mathbb{R},\ (x, \lambda) \mapsto f(x) + \lambda^\mathsf{T} g(x)$$

zusammengefasst werden: Ist $c \in M$ ein lokales Extremum von $f|_M$, so gibt es ein $\lambda \in \mathbb{R}^n$ mit

$$L'(c,\lambda) = \left(\frac{\partial L}{\partial x}(c,\lambda) \quad \frac{\partial L}{\partial \lambda}(c,\lambda) \right) = \left(f'(c) + \lambda^\mathsf{T} g'(c) \quad g(c) \right) \overset{27.26}{=} (0 \;\; 0) \quad \in \mathbb{R}^{1 \times (m+n)},$$

d. h. so dass L dort einen kritischen Punkt im Sinne von Lemma 26.2 hat.

(d) Allgemeine *hinreichende* Kriterien für Extrema unter Nebenbedingungen analog zu Satz 26.20 sind aufwändiger zu beweisen und sollen deshalb hier nicht behandelt werden.

72

Beispiel 27.27 (Randextrema durch Lagrange-Multiplikatoren)**.** Wir wollen noch einmal die Extremwerte der Funktion $f(x) = x_1 + 2x_2$ aus Beispiel 27.24 auf der abgeschlossenen Kreisscheibe $K_1(0) = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$ finden – diesmal aber mit dem Verfahren der Lagrange-Multiplikatoren.

Wir hatten schon gesehen, dass wir dazu nur noch die Extrema von f auf dem Rand der Kreisscheibe

$$M = \{x \in \mathbb{R}^2 : g(x) = 0\} \quad \text{mit} \quad g(x) := x_1^2 + x_2^2 - 1$$

finden müssen. Wegen $g'(x) = (2x_1 \;\; 2x_2) \neq (0 \;\; 0)$, also $\operatorname{rk} g'(x) = 1$ für alle $x \in M$, ist Satz 27.25 anwendbar und liefert uns, dass an jedem solchen Extremum die Gleichung $f' + \lambda\, g' = 0$ für ein $\lambda \in \mathbb{R}$ gelten muss. Mit $f' = (1 \;\; 2)$ bedeutet dies

$$1 + \lambda \cdot 2x_1 = 0 \quad \text{und} \quad 2 + \lambda \cdot 2x_2 = 0, \quad \text{zusammen mit der Nebenbedingung} \quad x_1^2 + x_2^2 - 1 = 0.$$

Dieses Gleichungssystem mit drei Variablen und Gleichungen können wir nun leicht lösen: Lösen wir die ersten beiden Gleichungen nach x_1 bzw. x_2 auf und setzen dies in die Nebenbedingung ein, so erhalten wir

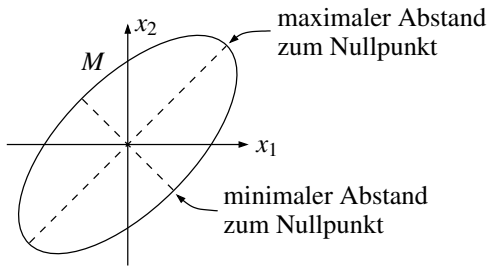
$$\frac{1}{4\lambda^2}+\frac{1}{\lambda^2}=1 \quad \text{und damit} \quad \lambda=\pm\frac{\sqrt{5}}{2},$$

woraus sich mit den obigen Gleichungen sofort $x_1=\mp\frac{1}{\sqrt{5}}$ und $x_2=\mp\frac{2}{\sqrt{5}}$ ergibt – also das gleiche Ergebnis, das wir auch in Beispiel 27.24 herausbekommen hatten.

Beispiel 27.28 (Hauptachsentransformation in der Analysis)**.** Es sei $A\in\mathbb{R}^{m\times m}$ eine symmetrische, positiv definite Matrix. Wir betrachten die Menge

$$M=\{x\in\mathbb{R}^m : g(x)=0\} \text{ mit } g(x)=x^\mathsf{T}Ax-1,$$

die nach der Hauptachsentransformation aus Konstruktion 22.37 der Rand eines (in der Regel gedrehten) Ellipsoids ist. Wir wollen auf M die Punkte mit maximalem und minimalem euklidischen Abstand zum Nullpunkt bestimmen – also die Punkte, an denen die Funktion $f\colon x\mapsto \|x\|_2^2=x^\mathsf{T}x$ wie im Bild rechts ihre Extrema annimmt.



Dazu berechnen wir zunächst wieder die Ableitungen: Nach Aufgabe 25.22 (a) ist

$$f'(x)=2x^\mathsf{T} \quad \text{und} \quad g'(x)=2x^\mathsf{T}A.$$

Beachte, dass $g'(x)\neq 0$ und damit auch $\operatorname{rk} g'(x)=1$ für alle $x\in M$ gilt: Andernfalls wäre nämlich $0=g'(x)\cdot x=2x^\mathsf{T}Ax$, woraus aus der positiven Definitheit von A sofort $x=0$ und damit $x\notin M$ folgt. Also ist die Rangbedingung an g' aus Satz 27.25 erfüllt, d. h. M ist eine $(m-1)$-dimensionale Untermannigfaltigkeit von $\mathbb{R}^m$. Damit können Extremstellen von f auf M nach diesem Satz nur dort vorliegen, wo es ein $\lambda\in\mathbb{R}$ gibt mit $f'(x)+\lambda\, g'(x)=0$. Durch Einsetzen der Ableitungen und Transponieren folgt daraus die Gleichung

$$2x+\lambda\cdot 2Ax=0, \quad \text{und damit} \quad Ax=-\frac{1}{\lambda}x$$

(da $\lambda=0$ unmöglich ist, weil daraus $x=0$ folgen würde). Der maximale und minimale Abstand von M zum Nullpunkt tritt damit nur an Eigenvektoren auf – was mit dem Ergebnis von Konstruktion 22.37 übereinstimmt. Die Lagrange-Multiplikatoren sind hier also durch die Eigenwerte von A bestimmt.

Aufgabe 27.29. Es sei noch einmal $M=\{x\in\mathbb{R}^2 : x_1^3=x_2^2\}$ wie in Aufgabe 27.23. Berechne alle lokalen Extrema der Funktion $f\colon M\to\mathbb{R}$, $x\mapsto x_1+x_2$.

Aufgabe 27.30. Finde mit Hilfe von Lagrange-Multiplikatoren eine Matrix $A\in\mathbb{R}^{2\times 2}$ mit $\det A=0$, deren Abstand zur Matrix $\begin{pmatrix}1&0\\0&2\end{pmatrix}$ in der euklidischen Norm minimal ist.