

Kapitel 6

Ausbau der Differentialrechnung

6.1 DIFFERENTIAL EINER TRANSFORMATION

Bisher haben wir Differenzierbarkeit von reellwertigen Funktionen in mehreren Variablen, sowie von Funktionen, die nur von einem Parameter abhängen, aber in einen höherdimensionalen Raum abbilden, untersucht. Nun wollen wir allgemeine Funktionen betrachten, bei denen sowohl der Ausgangsraum als auch der Bildraum höherdimensional sein kann. Eine solche Funktion ist differenzierbar, wenn sie sich lokal jeweils gut durch eine lineare Abbildung approximieren lässt. Die Definition der Differenzierbarkeit mithilfe der Dreigliedentwicklung lautet hier:

6.1.1 DEFINITION Sei $U \subset \mathbb{R}^n$ offen, $p \in U$ und $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Funktion. Sei weiter $\epsilon > 0$ so klein, dass die Kugel $K_\epsilon(p)$ vom Radius ϵ um p ganz im Definitionsbereich U liegt. Die Funktion f heisst an der Stelle $p \in U$ *differenzierbar*, wenn es eine lineare Abbildung $Df_p: \mathbb{R}^n \rightarrow \mathbb{R}^m$ und eine "Restfunktion" $R: K_\epsilon(0) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $\lim_{v \rightarrow 0} R(v) = 0$ gibt, so dass für alle $v \in \mathbb{R}^n$ mit $\|v\| < \epsilon$ gilt:

$$f(p + v) = f(p) + Df_p(v) + \|v\| \cdot R(v).$$

Die Funktion f lässt sich also entwickeln in einen konstanten Term $f(p)$, einen linearen Term $Df_p(v)$ und einen Restterm von höherer Ordnung. Die lineare Abbildung Df_p ist durch die Entwicklungsbedingung eindeutig festgelegt und wird als *Differential* von f an der Stelle p bezeichnet. Man nennt f auf U differenzierbar, wenn f an jeder Stelle von U differenzierbar ist.

6.1.2 BEISPIELE • Sei A eine $m \times n$ -Matrix, $w \in \mathbb{R}^m$ fest gegeben und $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ definiert durch $f(v) = Av + w$ für $v \in \mathbb{R}^n$. Dann ist f differenzierbar auf ganz \mathbb{R}^n und $Df_p(v) = Av$ für alle $p, v \in \mathbb{R}^n$. Hier ist jeweils die Restfunktion $R(v) = 0$.

- Sei $f(x, y) = \begin{pmatrix} 2x + y + xy \\ x - 3y + x^2 \end{pmatrix}$ für $x, y \in \mathbb{R}$. An der Stelle $p = 0$ ist $f(p) = 0$ und die Dreigliedentwicklung von f lautet hier: $f(x, y) = \begin{pmatrix} 2x + y \\ x - 3y \end{pmatrix} + \begin{pmatrix} xy \\ x^2 \end{pmatrix}$, wobei $Df_p(x, y) = \begin{pmatrix} 2x + y \\ x - 3y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

Überprüfen wir die Bedingung an die Restfunktion $R(x, y) = \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} xy \\ x^2 \end{pmatrix}$:

Für feste x, y und kleines $t \in \mathbb{R}$ gilt:

$$\lim_{t \rightarrow 0} R(tx, ty) = \lim_{t \rightarrow 0} \frac{1}{t\sqrt{x^2 + y^2}} \begin{pmatrix} t^2 xy \\ t^2 x^2 \end{pmatrix} = \lim_{t \rightarrow 0} \frac{1}{\sqrt{x^2 + y^2}} \begin{pmatrix} txy \\ tx^2 \end{pmatrix} = 0.$$

Also ist $\lim_{(x,y) \rightarrow 0} R(x, y) = 0$, wie verlangt.

6.1.3 BEMERKUNG Wie im eindimensionalen Fall kann man zeigen, dass eine in p differenzierbare Funktion an dieser Stelle auch stetig sein muss. Weiter lässt sich an der Definition sofort ablesen, dass eine Abbildung mit mehreren Komponenten $f: U \rightarrow \mathbb{R}^m$, $v \mapsto (f_1(v), \dots, f_m(v))$, genau dann in $a \in U$ differenzierbar ist, wenn jede der Komponentenfunktionen $f_i: U \rightarrow \mathbb{R}$ in p differenzierbar ist.

6.1.4 SATZ Sei jetzt $n > 1$, $m = 1$, $p \in U$ gegeben und $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion in n Variablen.

1. Ist f bei p differenzierbar, so gilt für $v \in \mathbb{R}^n$ mit $\|v\| < \epsilon$:

$$Df_p(v) = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t}.$$

Also gibt $Df_p(v)$ die Ableitung von f in Richtung von v an.

2. Ist f bei p differenzierbar, so existieren bei p auch sämtliche partiellen Ableitungen von f .

3. Ist f stetig partiell differenzierbar, so ist f auch differenzierbar.

Beweis. Um die erste Aussage zu zeigen, setzen wir in der Dreigliedentwicklung $p + tv$ ein und erhalten

$$f(p + tv) = f(p) + Df_p(tv) + \|tv\| \cdot R(tv) = f(p) + tDf_p(v) + |t|\|v\| \cdot R(tv),$$

weil Df_p linear ist. Daraus folgt

$$\lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t} = Df_p(v) \pm \lim_{t \rightarrow 0} \|v\| \cdot R(tv) = Df_p(v),$$

da nach Voraussetzung $\lim_{t \rightarrow 0} R(tv) = 0$.

Die erste Aussage liefert für die kanonischen Basisvektoren $v = e_j$:

$$Df_p(e_j) = \lim_{t \rightarrow 0} \frac{f(p + te_j) - f(p)}{t} = \partial_{x_j} f(p).$$

Insbesondere existieren also sämtliche partiellen Ableitungen bei p .

Nehmen wir jetzt an, dass f auf U stetig partiell differenzierbar ist. Für jeden Vektor $v \in \mathbb{R}^n$ mit $\|v\| < \epsilon$, ist dann die Funktion $g(t) = f(p + tv)$ für $|t| \leq 1$ definiert und bei $t = 0$ differenzierbar. Aus der Kettenregel folgt $g'(0) = \langle \nabla f(p), v \rangle$.

Setzen wir $t = 1$ ein in die Dreigliedentwicklung von g bei $t = 0$, erhalten wir folgende Dreigliedentwicklung von f :

$$f(p + v) = f(p) + \langle \nabla f(p), v \rangle + \|v\| \cdot R(v).$$

Also ist f differenzierbar bei p und wir lesen ab, dass $Df_p(v) = \langle \nabla f(p), v \rangle$. q.e.d.

Aus der Existenz der partiellen Ableitungen folgt noch nicht die Differenzierbarkeit. Hierzu ein Beispiel:

6.1.5 BEISPIEL Sei $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{sonst} \end{cases}$. Die partiellen Ableitungen von f im Nullpunkt existieren. Denn nach Definition ist

$$\partial_x f(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = 0$$

und ebenso $\partial_y f(0, 0) = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t} = 0$. Aber f kann im Nullpunkt nicht differenzierbar sein, denn f ist bei $p = (0, 0)$ noch nicht einmal stetig. Beispielsweise ist $\lim_{t \rightarrow 0} f(t, t) = \frac{1}{2}$, aber $f(0, 0) = 0$. Die Grenzwertbildung ist also nicht vertauschbar mit der Auswertung von f .

Das Differential von f an der Stelle p ist bezogen auf die kanonischen Standardbasen von \mathbb{R}^n und \mathbb{R}^m durch eine Matrix gegeben, die sogenannte *Jacobimatrix* $Jf(p)$ an der Stelle p .

6.1.6 SATZ Ist f an der Stelle $p \in U \subset \mathbb{R}^n$ differenzierbar und besteht f aus den Komponenten f_1, \dots, f_m , so existieren auch alle partiellen Ableitungen $\frac{\partial f_k}{\partial x_j}(p)$ und die lineare Abbildung Df_p wird bezüglich der kanonischen Basen durch die Matrix beschrieben, die aus sämtlichen partiellen Ableitungen gebildet ist:

$$Jf(p) = \begin{pmatrix} \partial_{x_1} f_1(p), & \dots, & \partial_{x_n} f_1(p) \\ \partial_{x_1} f_2(p), & \dots, & \partial_{x_n} f_2(p) \\ \vdots & \vdots & \vdots \\ \partial_{x_1} f_m(p), & \dots, & \partial_{x_n} f_m(p) \end{pmatrix}.$$

Beweis. Dies ergibt sich sofort durch Vergleich mit den Dreigliedentwicklungen der Komponentenfunktionen von f . q.e.d.

6.1.7 BEISPIEL Betrachten wir die Umrechnung von Polarkoordinaten in kartesische Koordinaten der Ebene als eine Funktion in zwei Variablen mit zwei Komponenten, nämlich

$$f(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \quad \text{für } r \geq 0, \varphi \in \mathbb{R}.$$

Der Definitionsbereich U ist hier also eine Halbebene. Die Funktion f bildet die Parallelen zur positiven r -Achse auf Radialstrahlen ab, und die Parallelen zur φ -Achse werden unter f auf Kreise um den Nullpunkt abgebildet. Die Funktion f ist

überall differenzierbar und die Jacobimatrix zu f an einer Stelle $p = (r, \varphi)$ (mit $r > 0$) ist:

$$Jf(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Die bereits behandelten Spezialfälle ordnen sich hier folgendermassen ein:

$m = 1$. Hier besteht die Jacobimatrix nur aus einer Zeile und $Jf(p) = (\nabla f(p))^T$.

$n = 1$. Hier geht es sich eigentlich um einen parametrisierten Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$, und die Jacobimatrix an einer Stelle t ist jeweils der Geschwindigkeitsvektor $J\gamma(t) = \gamma'(t)$.

$n = m$. Ist f das Gradientenvektorfeld eines Potentials U , so stimmt die Jacobimatrix von $f = \nabla U$ mit der Hessematrix von U überein $Jf(p) = H_U(p)$.

6.1.8 BEMERKUNG Sei jetzt $n = m$. In diesem Fall gibt der Betrag der Determinante der Jacobimatrix den lokalen Expansions- oder Kontraktionsfaktor von f an. Genauer gilt:

$$|\det(Df_p)| = \lim_{\epsilon \rightarrow 0} \frac{\text{Vol}_n(f(K_\epsilon(p)))}{\text{Vol}_n(K_\epsilon(p))} \quad \forall p \in U.$$

Für lineare Abbildungen ist das klar (siehe Folgerung 5.4.6) und für beliebige Transformationen ist es schwieriger (wir verzichten hier auf den Beweis).

Im eben betrachteten Beispiel ist $\det Jf(p) = r > 0$, das heisst, im Bereich $0 < r < 1$ findet eine Kontraktion und im Bereich $r > 1$ eine Expansion statt.

Die Kettenregel lautet im Mehrdimensionalen folgendermassen:

6.1.9 SATZ Seien $g: U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^l$ und $f: V \rightarrow \mathbb{R}^m$ Funktionen und U, V jeweils offene Teilmengen. Ist g differenzierbar an der Stelle $p \in U$, und ist f differenzierbar an der Stelle $g(p) \in V$, so ist auch $f \circ g$ bei p differenzierbar. Für die Differentiale gilt:

$$D(f \circ g)_p = Df_{g(p)} \circ Dg_p.$$

Das bedeutet für die entsprechenden Jacobi-Matrizen:

$$J(f \circ g)(p) = Jf(g(p)) \cdot Jg(p).$$

6.1.10 BEISPIEL Sei $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch $g(x, y) = (x+2y, x \cdot y)$, und $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $f(u, v) = e^{2u}v$. Die Zusammensetzung lautet dann $(f \circ g)(x, y) = e^{2x+4y}xy$. Die zugehörigen Jacobimatrizen sind folgende:

$$Jf(u, v) = (2e^{2u}v, e^{2u}) , \quad Jg(x, y) = \begin{pmatrix} 1 & 2 \\ y & x \end{pmatrix}.$$

Durch Multiplikation erhalten wir

$$Jf(g(x, y)) \cdot Jg(x, y) = (2e^{2x+4y}xy, e^{2x+4y}) \cdot \begin{pmatrix} 1 & 2 \\ y & x \end{pmatrix} = e^{2x+4y}(2xy + y, 4xy + x).$$

Andererseits ist

$$\partial_x(f \circ g)(x, y) = ye^{2x+4y} + 2xye^{2x+4y} \quad \text{und} \quad \partial_y(f \circ g)(x, y) = xe^{2x+4y} + 4xye^{2x+4y},$$

Also bestätigt sich hier die Regel $J(f \circ g)(x, y) = Jf(g(x, y)) \cdot Jg(x, y)$.

6.1.11 FOLGERUNG Ist $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenzierbar und ist $\gamma: [a, b] \rightarrow U$ ein differenzierbarer Weg mit $\gamma(a) = p$, dann bildet f den Weg γ auf den Weg $\tilde{\gamma} = f \circ \gamma$ in \mathbb{R}^m ab und nach der Kettenregel ist $Df_p(\gamma'(a)) = (\tilde{\gamma}')(a)$. Das Differential Df_p bildet also die entsprechenden Geschwindigkeitsvektoren der Wege aufeinander ab.

6.1.12 SATZ Sei $D \subset \mathbb{R}^n$ offen und $f: D \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Sei weiter $p \in D$ ein Punkt mit $\det Df_p \neq 0$. Dann ist f in der Nähe von p lokal umkehrbar in folgendem Sinn: Es gibt eine offene Umgebung $U \subset D$ von p , so dass f , aufgefasst als Abbildung von U nach $V := f(U)$ bijektiv ist und die Umkehrabbildung $f^{-1}: V \rightarrow U$ ebenfalls differenzierbar ist. Außerdem gilt für alle $x \in U$

$$D(f^{-1})_{f(x)} = (Df_x)^{-1}.$$

Der Beweis des Umkehrsatzes ist schwierig. Man kann die Aussage auf den Satz über implizite Funktionen zurückführen, der im letzten Kapitel kurz vorgestellt wird.

Hier zwei Beispiele.

6.1.13 BEISPIEL Sei $f(x, y) = \begin{pmatrix} xy \\ y \end{pmatrix}$ für $(x, y) \in \mathbb{R}^2$. Diese Funktion bildet die x -Achse auf den Nullpunkt ab, alle anderen Parallelen zur x -Achse gehen wieder in Parallelen zur x -Achse über. Dagegen werden die Parallelen zur y -Achse zu Geraden durch den Nullpunkt. Die Jacobimatrix von f lautet

$$Jf(x, y) = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad \det Jf(x, y) = y.$$

Also ist die Funktion f in der Nähe des Punktes $p = (1, 1)$ lokal umkehrbar. Setzen wir $u = xy$ und $v = y$, dann ist

$$f^{-1}(u, v) = \begin{pmatrix} u/v \\ v \end{pmatrix}.$$

Diese Umkehrfunktion ist definiert auf $V = \{(u, v) \mid v \neq 0\}$ und hier ist dann $U = \{(x, y) \mid y \neq 0\}$. Die Teilmenge $U \subset \mathbb{R}^2$ ist bereits die grösstmögliche Teilmenge, auf der f umkehrbar ist.

6.1.14 BEISPIEL Betrachten wir wiederum den Wechsel von Polarkoordinaten zu kartesischen Koordinaten der Ebene. Sei also

$$f(r, \varphi) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \quad \text{für } r \geq 0, \varphi \in \mathbb{R}.$$

Die Jacobimatrix an einer Stelle (r, φ) hatten wir bereits berechnet:

$$Jf(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix},$$

und $\det Jf(r, \varphi) = r$. Das Differential von f an der Stelle (r, φ) ist also invertierbar, falls $r \neq 0$. Aber es reicht hier nicht, die Punkte mit $r = 0$ aus dem Definitionsbereich herauszunehmen, weil f ausserdem bezogen auf den Winkel φ periodisch ist und bei der Umkehrung dann eine Mehrdeutigkeit entsteht. Dennoch können wir lokal umkehren.

Hier ist eine Beschreibung der Umkehrfunktion von f in der Nähe von $p = (1, 0)$:

$$f^{-1}(x, y) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan\left(\frac{y}{x}\right) \end{pmatrix} \quad \text{für } x > 0, y \in \mathbb{R}.$$

In diesem Fall ist $U = \{(r, \varphi) \mid r > 0, -\frac{\pi}{2} < \varphi < \frac{\pi}{2}\}$ und $V := \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$.

Wir überprüfen nun noch die Aussage über die Differentiale an diesem Beispiel. Einerseits ist

$$Jf(r, \varphi)^{-1} = \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix}.$$

Andererseits ist

$$J(f^{-1})(x, y) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}.$$

Setzen wir nun $x = r \cos \varphi$ und $y = r \sin \varphi$ ein, erhalten wir Übereinstimmung.

6.2 TRANSFORMATIONSREGEL

Nun können wir die Entsprechung der Substitutionsregel im Mehrdimensionalen formulieren.

Sei dazu $\Phi: S \rightarrow \Phi(S) \subset \mathbb{R}^n$ eine bijektive stetig differenzierbare Transformation, deren Umkehrung Φ^{-1} ebenfalls stetig differenzierbar sei. Die Verallgemeinerung der Aussage über lineare Transformationen lautet dann:

6.2.1 SATZ Ist $S \subset \mathbb{R}^n$ messbar und $f: \Phi(S) \rightarrow \mathbb{R}$ integrierbar, so gilt:

$$\int_{\Phi(S)} f(x) d^n(x) = \int_S f(\Phi(u)) |\det D\Phi_u| d^n u.$$

Beweis. Man kann die Überlegung für $n = 2$, die wir für lineare Transformationen angestellt hatten, folgendermassen anpassen. Sei Z eine Zerlegung eines Rechtecks $Q \supset S$, und sei jeweils u_k eine Stützstelle im Teilrechteck Q_k . Dann betrachten wir die Dreigliedentwicklung der Transformation Φ an der Stelle u_k

$$\Phi(u) = \Phi(u_k) + D\Phi_{u_k}(u - u_k) + R(u) \approx \Phi(u_k) + D\Phi_{u_k}(u - u_k).$$

In der Nähe des Punktes u_k wird die Transformation Φ also näherungsweise durch die affine Transformation $u \mapsto \Phi(u_k) + D\Phi_{u_k}(u - u_k)$ dargestellt. Die Transformation Φ bildet daher das Teilrechteck Q_k auf ein Gebiet $\Phi(Q_k)$ ab, für dessen Flächeninhalt gilt:

$$\text{Vol}_2(\Phi(Q_k)) \approx |\det D\Phi_{u_k}| \cdot \text{Vol}_2(Q_k).$$

Sei jetzt wieder $x_k := \Phi(u_k)$. Dann liefert der Vergleich der Riemannsummen:

$$\begin{aligned} \int_{\Phi(S)} f(x, y) d^2(x, y) &\approx \sum_k f(x_k) \text{Vol}_2(\Phi(Q_k)) \approx \\ \sum_k f(\Phi(u_k)) |\det D\Phi_{u_k}| \cdot \text{Vol}_2(Q_k) &\approx \int_S f(\Phi(u)) |\det D\Phi_u| d^2u. \end{aligned}$$

Die Details der Grenzübergänge lassen wir hier aus und beschränken uns auf diese Skizze der Ideen. q.e.d.

6.2.2 BEISPIELE

- Die schon genannte Transformation auf Polarkoordinaten ordnet sich hier folgendermassen ein. Die Zuordnung

$$\Phi: \mathbb{R}_{>0} \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}, \quad (r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi)$$

ist bijektiv und in beiden Richtungen differenzierbar und $|\det D\Phi_{(r,\varphi)}| = r$. Deshalb lautet die entsprechende Substitution $dx dy = r dr d\varphi$.

- Die Einführung von Zylinderkoordinaten im dreidimensionalen Raum korrespondiert zu der Transformation

$$\Phi(r, \varphi, z) = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \\ z \end{pmatrix} \quad (r > 0, 0 \leq \varphi \leq 2\pi, z \in \mathbb{R}).$$

Wiederum ist Φ in beiden Richtungen differenzierbar und $|\det D\Phi_{(r,\varphi)}| = r$. Die entsprechende Substitution ist hier

$$dx dy dz = r dr d\varphi dz.$$

- Die Kugelkoordinaten sind gegeben durch

$$\Phi(r, \varphi, \theta) = \begin{pmatrix} r \cos(\varphi) \sin(\theta) \\ r \sin(\varphi) \sin(\theta) \\ r \cos \theta \end{pmatrix} \quad (r \geq 0, 0 \leq \varphi \leq 2\pi, 0 \leq \theta \leq \pi).$$

Man rechnet nach, dass $|\det D\Phi_{(r,\varphi,\theta)}| = r^2 \sin \theta$. Die entsprechende Substitution ist also

$$dx dy dz = r^2 \sin(\theta) dr d\varphi d\theta.$$

Im folgenden Beispiel werden Kugelkoordinaten verwendet:

6.2.3 BEISPIEL Das Gebiet $B = \{(x, y, z) \mid (x^2 + y^2 + z^2)^2 \leq y\}$ hat die Gestalt eines Tropfens. Seine Beschreibung in Kugelkoordinaten lautet:

$$A = \{(r, \varphi, \theta) \mid r^3 \leq \sin(\varphi) \sin(\theta), 0 \leq \varphi \leq \pi, 0 \leq \theta \leq \pi\}.$$

Das Volumen des Körpers B ist gegeben durch das Integral

$$\begin{aligned} \int_B 1 \, d^3(x, y, z) &= \int_A r^2 \sin(\theta) \, dr \, d\theta \, d\varphi = \int_0^\pi \int_0^\pi \int_0^{\sqrt[3]{\sin(\varphi) \sin(\theta)}} r^2 \sin(\theta) \, dr \, d\theta \, d\varphi = \\ &= \int_0^\pi \int_0^\pi \frac{1}{3} \sin(\varphi) \sin^2(\theta) \, d\theta \, d\varphi = \int_0^\pi \frac{\pi}{6} \sin(\varphi) \, d\varphi = \frac{\pi}{3}, \end{aligned}$$

Im letzten Beispiel verwenden wir eine nichtlineare Transformation, mit der die Beschreibung des Integrationsgebietes vereinfacht wird:

6.2.4 BEISPIEL Sei $B = \{(x, y) \in \mathbb{R}^2 \mid x, y > 0, 0 < xy < 3, x < y < 2x\}$. In den neuen Koordinaten $u = xy$ und $v = y/x$ lautet die Beschreibung des Gebietes $S = \{(u, v) \in \mathbb{R}^2 \mid 0 < u < 3, 1 < v < 2\}$. Die entsprechende Transformation ist

$$\Phi: (u, v) \mapsto \left(\sqrt{\frac{u}{v}}, \sqrt{uv} \right) \quad \text{und} \quad \det D\Phi_{(u,v)} = \begin{vmatrix} \frac{1}{2\sqrt{uv}} & -\frac{\sqrt{u}}{2\sqrt{v^3}} \\ \frac{1}{2}\sqrt{\frac{v}{u}} & \frac{1}{2}\sqrt{\frac{u}{v}} \end{vmatrix} = \frac{1}{2v}.$$

Der Flächeninhalt des Gebietes B beträgt also:

$$\text{Vol}_2(B) = \int_B 1 \, d^2(x, y) = \int_S \frac{1}{2v} \, d^2(u, v) = \int_0^3 \int_1^2 \frac{1}{2v} \, dv \, du = \frac{3}{2} \ln(2).$$

Das Integral über die Funktion $f(x, y) = y^2$ über B können wir mithilfe der Transformation folgendermassen berechnen:

$$\int_B y^2 \, d^2(x, y) = \int_S uv \left| \det D\Phi_{u,v} \right| \, d^2(u, v) = \int_0^3 \int_1^2 \frac{uv}{2v} \, dv \, du = \frac{9}{4}.$$

6.3 SATZ ÜBER IMPLIZITE FUNKTIONEN

Nehmen wir an, $f: D \rightarrow \mathbb{R}$ sei eine stetig differenzierbare Funktion in 2 Variablen, definiert auf einer offenen Teilmenge $D \subset \mathbb{R}^2$. Durch die Gleichung $f(x, y) = 0$ wird auf implizite Art und Weise eine Teilmenge des \mathbb{R}^2 beschrieben, nämlich die Nullstellenmenge der Funktion f . Um diese Nullstellenmenge explizit zu bestimmen, müsste man die Gleichung nach einer der Variablen auflösen. Hier dazu einige Beispiele:

6.3.1 BEISPIELE 1. Seien $a, b, c \in \mathbb{R}$ fest und $f(x, y) = ax + by - c$ für $x, y \in \mathbb{R}$.

Die Gleichung $f(x, y) = ax + by - c = 0$ ist eine Geradengleichung, falls a und b nicht beide gleichzeitig Null sind. Falls $b \neq 0$ ist, können wir nach y auflösen und erhalten $y = \frac{c-ax}{b}$ für alle x . Ist $b = 0$, beschreibt die entsprechende Geradengleichung eine Parallele zur y -Achse. In diesem Fall können wir nicht nach y auflösen, dafür aber nach x .

2. Seien $a, b > 0$ fest und $f(x, y) = ax^2 + by^2 - 1$ für $x, y \in \mathbb{R}$. Die Gleichung $ax^2 + by^2 = 1$ beschreibt eine Ellipse in \mathbb{R}^2 . Lösen wir diese Gleichung nach y auf, erhalten wir $y = \pm \frac{\sqrt{1-ax^2}}{b}$, für $x^2 \leq \frac{1}{a}$. Die Ellipse lässt sich also als Vereinigung von zwei Funktionsgraphen auffassen, allerdings sind die entsprechenden Funktionen bei $x = \pm \frac{1}{\sqrt{a}}$ nicht differenzierbar! Interessieren wir uns nur für einen Ausschnitt der Ellipse in der Nähe des Punktes $(x_0, y_0) = (0, \frac{1}{\sqrt{b}})$, so müssen wir den oberen Zweig, nämlich $y = +\frac{\sqrt{1-ax^2}}{b}$, wählen. Und wenn wir hier nur $x \in (-1, 1)$ zulassen, ist sichergestellt, dass die dadurch definierte Funktion sogar überall differenzierbar ist.

Wir werden nun zunächst ein Kriterium dafür formulieren, unter welchen Umständen sich die Gleichung $f(x, y) = 0$ nach y auflösen lässt, und zwar in der Nähe eines vorgegebenen Punktes (x_0, y_0) mit $f(x_0, y_0) = 0$. Gesucht sind genauer offene Umgebungen $x_0 \in U \subset \mathbb{R}$, $y_0 \in V \subset \mathbb{R}$ mit $U \times V \subset D$ und eine Funktion $g: U \rightarrow V$ mit $g(x_0) = y_0$ so dass

$$f(x, y) = 0 \iff y = g(x) \quad \text{für alle } (x, y) \in U \times V.$$

Das bedeutet, dass die Nullstellenmenge von f in $U \times V$ mit dem Graphen von g übereinstimmt. Gibt es eine solche Funktion g , so ist sie offenbar eindeutig durch f festgelegt. Man sagt auch, die Funktion g sei “implizit” durch die Gleichung $f(x, y) = 0$ definiert.

Angenommen, die Funktion g existiert und ist differenzierbar auf U . Dann ergibt sich aus der Beziehung $f(x, g(x)) = 0$ mit der Kettenregel

$$0 = \frac{d}{dx} f(x, g(x)) = \partial_x f(x, g(x)) + g'(x) \partial_y f(x, g(x)).$$

Daraus folgt, falls $\partial_y f(x, g(x)) \neq 0$,

$$g'(x) = -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))} \quad \text{für alle } x \in U.$$

Tatsächlich reicht es, $\partial_y f(x, y) \neq 0$ für $x = x_0$ und $y = y_0$ vorauszusetzen, um die Existenz von g zu garantieren. Denn es gilt der folgende *Satz über implizite Funktionen*:

6.3.2 SATZ Sei $D \subset \mathbb{R}^2$ offen, $f: D \rightarrow \mathbb{R}$ stetig differenzierbar und $f(x_0, y_0) = 0$ für ein $(x_0, y_0) \in D$. Sei weiter $\partial_y f(x_0, y_0) \neq 0$. Dann gibt es offene Umgebungen $x_0 \in U \subset \mathbb{R}$, $y_0 \in V \subset \mathbb{R}$ mit $U \times V \subset D$ und eine eindeutig bestimmte, stetig differenzierbare Funktion $g: U \rightarrow V$ mit $g(x_0) = y_0$, so dass für alle $(x, y) \in U \times V$ gilt:

$$f(x, y) = 0 \iff y = g(x).$$

Weiter ist

$$g'(x) = -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))} \quad \text{für alle } x \in U.$$

Die Nullstellenmenge von f lässt sich also in einer passenden Umgebung von (x_0, y_0) als Graph einer stetig differenzierbaren Funktion darstellen.

Hier dazu noch ein Beispiel.

6.3.3 BEISPIEL Sei $a > 0$ fest gewählt und $f(x, y) := (x^2 + y^2)^2 - 2ax(x^2 + y^2) - a^2y^2$ für $x, y \in \mathbb{R}$. Die Lösungsmenge der Gleichung $f(x, y) = 0$ in \mathbb{R}^2 wird als Kardioide bezeichnet. Sie ist herzförmig und hat einen Knickstelle im Ursprung. Es gibt zwei Schnittpunkte mit der x -Achse, nämlich den Nullpunkt und den Punkt $(2a, 0)$. Ausserdem schneidet die Kurve die y -Achse bei $y = 0$ und $y = \pm a$. Wir haben hier

$$\partial_y f(x, y) = 2(x^2 + y^2)2y - 4axy - 2a^2y.$$

Die partielle Ableitung nach y hat vier Nullstellen auf der Kardioide.

Entfernt man diese Punkte, bleiben vier Kurvenabschnitte, die jeweils Funktionsgraphen sind. Im Nullpunkt dagegen ist $\partial_y f(0, 0) = 0$, und man kann die Gleichung in der Nähe des Nullpunktes nicht nach y auflösen. Denn für kleine $x < 0$ gibt es immer genau zwei y -Werte (nahe bei 0), die zusammen mit dem x -Wert die Gleichung lösen.

Beweis von Satz 6.3.2. Wir gehen in mehreren Schritten vor.

1. Schritt: Konstruktion von g : Dazu nehmen wir an, dass $\partial_y f(x_0, y_0) > 0$. (Im anderen Fall argumentiert man entsprechend.) Aus der Stetigkeit von $\partial_y f$ folgt, dass $\partial_y f(x, y) > 0$ für alle (x, y) aus einer passenden Umgebung \tilde{D} von (x_0, y_0) . Daraus können wir schliessen, dass f auf \tilde{D} bezüglich y streng monoton wachsend ist, das heisst $f(x, y_1) > f(x, y_2)$ für alle $(x, y_j) \in \tilde{D}$ mit $y_1 > y_2$. Nun wählen wir $\epsilon, \delta > 0$ mit $[x_0 - \delta, x_0 + \delta] \times [y_0 - \epsilon, y_0 + \epsilon] \in \tilde{D}$. Da $f(x_0, y_0) = 0$, ist insbesondere $f(x_0, y_0 - \epsilon) < 0$ und $f(x_0, y_0 + \epsilon) > 0$. Wegen der Stetigkeit von f können wir sicherstellen, dass $f(x, y_0 - \epsilon) < 0$ und $f(x, y_0 + \epsilon) > 0$ für alle $x \in [x_0 - \delta, x_0 + \delta]$, indem wir δ falls nötig noch verkleinern. Aus dem Zwischenwertsatz folgt nun, dass zu jedem x ein $y \in [y_0 - \epsilon, y_0 + \epsilon]$ existiert mit $f(x, y) = 0$. Dies y ist durch x eindeutig bestimmt, weil f bezüglich y streng monoton steigend ist. Wenn wir

$g(x) := y$ setzen, erhalten wir also die gesuchte Funktion g auf $U := (x_0 - \delta, x_0 + \delta)$, und $g(U) \subset V := (y_0 - \epsilon, y_0 + \epsilon)$.

Aus der Konstruktion ergibt sich auch sofort, dass g in x_0 stetig ist. Denn wenn wir ϵ verkleinern, so können wir δ immer entsprechend anpassen, so dass aus $|x - x_0| < \delta$ folgt $|g(x) - y_0| < \epsilon$. Jeder Punkt $x \in U$ erfüllt dieselben Voraussetzungen wie x_0 . Also kann man die Stetigkeit von g in x analog zeigen.

2. Schritt: Lipschitz-Stetigkeit von g : Nehmen wir ab jetzt an, dass $(x_0, y_0) = (0, 0)$ ist. (Dies können wir erreichen, indem wir anstelle von f die Funktion \tilde{f} , definiert durch $\tilde{f}(x, y) = f(x - x_0, y - y_0)$ betrachten.) Wir zeigen jetzt, dass es eine Konstante $K > 0$ und ein $\delta > 0$ gibt mit

$$|g(x)| \leq K|x| \quad \text{für alle } |x| < \delta.$$

Dazu verwenden wir die Dreigliedrung von f im Nullpunkt:

$$f(x, y) = Df_0\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) + R(x, y)\|(x, y)\|,$$

wobei $R: K_r(0) \rightarrow \mathbb{R}$ eine Funktion ist, für die gilt: $\lim_{(x, y) \rightarrow (0, 0)} R(x, y) = 0$. Setzen wir $a := \partial_x f(0, 0)$ und $b := \partial_y f(0, 0)$, so erhalten wir:

$$f(x, y) = ax + by + R(x, y)\sqrt{x^2 + y^2}.$$

Wenn wir $y = g(x)$ einsetzen, wird daraus:

$$0 = f(x, g(x)) = ax + bg(x) + R(x, g(x))\sqrt{x^2 + g(x)^2}.$$

Weil nach Voraussetzung $b \neq 0$ ist, können wir schliessen

$$g(x) = -\frac{a}{b}x - \frac{1}{b}R(x, g(x))\sqrt{x^2 + g(x)^2}.$$

Wegen der Stetigkeit von g bei 0 gilt $\lim_{x \rightarrow 0} g(x) = 0$ und $\lim_{x \rightarrow 0} R(x, g(x)) = 0$. Daher gibt es ein $\delta > 0$ mit $|R(x, g(x))| < \frac{|b|}{2}$ für alle $|x| < \delta$. Daraus folgt

$$|g(x)| \leq \left|\frac{a}{b}\right||x| + \frac{1}{2}\sqrt{x^2 + g(x)^2} \leq \left|\frac{a}{b}\right||x| + \frac{1}{2}(|x| + |g(x)|),$$

und schliesslich die gesuchte Abschätzung

$$|g(x)| \leq (2\left|\frac{a}{b}\right| + 1)|x| \quad \text{für alle } |x| < \delta.$$

3. Schritt: Differenzierbarkeit von g : Um zu zeigen, dass g im Nullpunkt differenzierbar ist, reicht es zu zeigen, dass g im Nullpunkt eine Dreigliedrung besitzt. Wie eben gezeigt, gilt

$$g(x) = -\frac{a}{b}x - \frac{1}{b}R(x, g(x))\sqrt{x^2 + g(x)^2},$$

und falls die Ableitung von g an der Stelle $x = 0$ existiert, muss wegen der Kettenregel $g'(0) = -\frac{a}{b}$ sein. Der lineare Term in dieser Darstellung ist also bereits der Kandidat für die Ableitung. Jetzt müssen wir nur noch überprüfen, dass der zweite Term in der Entwicklung von g die Restbedingung erfüllt, das heisst

$$\lim_{x \rightarrow 0} \left(-\frac{1}{b} R(x, g(x)) \cdot \frac{\sqrt{x^2 + g(x)^2}}{|x|} \right) = 0.$$

Dazu verwenden wir die Abschätzung aus dem vorigen Schritt mit $K = 2|\frac{a}{b}| + 1$:

$$\frac{\sqrt{x^2 + g(x)^2}}{|x|} \leq \frac{|x| + |g(x)|}{|x|} \leq 1 + K \quad \text{für alle } |x| < \delta.$$

Da $\lim_{x \rightarrow 0} R(x, g(x)) = 0$, folgt nun die Behauptung. q.e.d.

Es gibt auch eine entsprechende Version des Satzes über implizite Funktionen für Gleichungen in mehr als nur zwei Variablen. Diese allgemeinere Version lautet:

6.3.4 SATZ Sei $D \subset \mathbb{R}^n \times \mathbb{R}$ offen, $f: D \rightarrow \mathbb{R}$ stetig differenzierbar und $f(x_0, y_0) = 0$ für ein $(x_0, y_0) \in D$. Sei weiter $\partial_{n+1} f(x_0, y_0) \neq 0$. Dann gibt es offene Umgebungen $x_0 \in U \subset \mathbb{R}^n$, $y_0 \in V \subset \mathbb{R}$ mit $U \times V \subset D$ und eine eindeutig bestimmte, stetig differenzierbare Funktion $g: U \rightarrow V$ mit $g(x_0) = y_0$, so dass für alle $(x, y) \in U \times V$ gilt:

$$f(x, y) = 0 \iff y = g(x).$$

Weiter ist

$$\partial_j g(x) = -\frac{\partial_j f(x, g(x))}{\partial_{n+1} f(x, g(x))} \quad \text{für alle } x \in U.$$

Die Nullstellenmenge von f lässt sich also in einer passenden Umgebung von (x_0, y_0) als Graph einer stetig differenzierbaren Funktion g in n Variablen darstellen.

Die Aussage über die partiellen Ableitungen von g ergibt sich wieder durch Anwendung der Kettenregel auf die Gleichung

$$f(x, g(x)) = f(x_1, \dots, x_n, g(x_1, \dots, x_n)) = 0 \quad \text{für alle } x \in U.$$

6.3.5 BEISPIEL Die Nullstellenmenge der Funktion $f(x, y, z) = x^2 + y^2 - z^2 - 1$ ist ein einschaliges Hyperboloid. Eine explizite Beschreibung dieser Fläche in der Nähe des Punktes $a = (2, 0, \sqrt{3})$ erhalten wir, indem wir nach z auflösen zu $z = \sqrt{x^2 + y^2 - 1}$ (falls $x^2 + y^2 > 1$). Für die Punkte $(x, y, z) \in \mathbb{R}^3$ mit $z > 0$ und $x^2 + y^2 > 1$ gilt:

$$f(x, y, z) = 0 \iff z = \sqrt{x^2 + y^2 - 1}.$$

Hier ist also $U := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 > 1\}$ und $V = \mathbb{R}_{>0}$ und $g(x, y) = \sqrt{x^2 + y^2 - 1}$ für alle $(x, y) \in U$. Der Graph von g ist sozusagen die obere Hälfte des Hyperboloids (wobei der Schnittkreis mit der x - y -Ebene nicht mitzählt). Überprüfen wir hier noch die Aussage über die partiellen Ableitungen von g :

$$\partial_x g(x, y) = \frac{2x}{2\sqrt{x^2 + y^2 - 1}} = \frac{2x}{2g(x, y)} = \frac{\partial_x f(x, y, g(x, y))}{\partial_z f(x, y, g(x, y))}.$$

$$\partial_y g(x, y) = \frac{2y}{2\sqrt{x^2 + y^2 - 1}} = \frac{2y}{2g(x, y)} = \frac{\partial_y f(x, y, g(x, y))}{\partial_z f(x, y, g(x, y))}.$$

Wie bereits betont, liefert der Satz über implizite Funktionen keine Formel für die lokale Funktion g , aber schon aus der theoretischen Existenz der Funktion g ergeben sich wichtige qualitative Konsequenzen.

6.3.6 FOLGERUNG *Ist $f: D \subset \mathbb{R}^d \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion in $d > 1$ Variablen, $c \in \mathbb{R}$ fest gewählt und gilt $\nabla f(a) \neq 0$ für alle $a \in D$ mit $f(a) = c$, so ist die Niveaumenge $N = N_c(f) = \{a \in D \mid f(a) = c\}$ von f lokal Graph einer stetig differenzierbaren Funktion g . Das heisst, zu jedem Punkt $p \in N_c$ gibt es eine Umgebung $p \in U \subset \mathbb{R}^d$ und eine reellwertige, stetig differenzierbare Funktion g in $d - 1$ Variablen, so dass $N_c \cap U$ mit dem Graphen von g übereinstimmt. Ist $d = 2$, können wir daraus schliessen, dass N eine Kurve ist, die in jedem Punkt eine eindeutige Tangente besitzt und keine Spitzen, Selbstübergreifungen oder Selbstberührungen hat. Ist $d = 3$, so ist N eine Fläche, die in jedem Punkt eine eindeutige Tangentialebene hat, wobei wiederum keine Selbstdurchdringungen oder Falten auftreten.*

Beweis. Sei $p \in N_c$ vorgegeben. Nach Voraussetzung ist $\nabla f(p) \neq 0$, es gibt also einen Index j mit $\partial_j f(p) \neq 0$. Wir ändern nun die Numerierung der Variablen so, dass wir $\partial_d f(p) \neq 0$ erhalten. Nun können wir den Satz über implizite Funktionen auf die Funktion $h(x) := f(x) - c$ anwenden. q.e.d.

6.3.7 BEISPIELE • Die Niveaulinien N_c der Funktion $f(x, y) = x^2 + y^2$ sind

konzentrische Kreise (für $c > 0$). Ist $c = 0$, entartet die Niveaumenge zu einem Punkt, für $c < 0$ hat die entsprechende Gleichung jeweils keine Lösung, die zugehörigen Niveaumengen sind also leer. Ist $c > 0$ fest gewählt, können wir die entsprechende Kreislinie jeweils lokal als Graphen einer differenzierbaren Funktion von x oder von y schreiben, nämlich $x = \pm\sqrt{c - y^2}$ (für $|y| < \sqrt{c}$) oder $y = \pm\sqrt{c - x^2}$ (für $|x| < \sqrt{c}$). Mit weniger als vier lokalen Graphen kommt man aber nicht aus, wenn man den ganzen Kreis überdecken möchte.

- Betrachten wir jetzt die Funktion $f(x, y) = y^2 - x^3$. Die Gleichung $f(x, y) = 0$ beschreibt die sogenannte *Neillsche Parabel*. Es handelt sich um eine Kurve in der Halbebene rechts von der y -Achse mit einer Spitze im Nullpunkt. Entfernt man den Nullpunkt, bleibt noch eine Vereinigung von zwei Funktionsgraphen übrig, nämlich der Funktionen $y = \pm\sqrt{x^3}$ (jeweils für $x > 0$). Aber im Nullpunkt gibt es keine lokale Beschreibung der Kurve als Graphen einer Funktion $y = g(x)$, denn in jeder Umgebung der Form $K_\epsilon(0)$ gibt es zu jedem positiven x -Wert zwei entsprechende Punkte auf der Kurve.

Ist $c < 0$, können wir die entsprechende Niveaulinie von f als Graph der Funktion $x = \sqrt[3]{y^2 + |c|}$ (für $y \in \mathbb{R}$) darstellen. Ist $c > 0$, so benötigen wir mindestens drei lokale Graphen, um die Niveaulinie zu überdecken, etwa $y = \pm\sqrt{x^3 + c}$ (jeweils für $x > -\sqrt[3]{c}$) und $x = \sqrt[3]{y^2 - c}$ (für $|y| < \sqrt{c}$).

Und hier schliesslich noch zwei Beispiele.

6.3.8 BEISPIELE

- Die Funktion $f(x, y) = 1 - (x^2 - 1)^2 - y^2$ hat zwei isolierte lokale Maxima bei $(\pm 1, 0)$. Die entsprechende Niveaumenge N_1 besteht nur aus diesen beiden Punkten. Die Niveaumenge N_0 ist eine Kurve mit einer Selbstüberkreuzung im Nullpunkt, die ansonsten glatt ist. Im Nullpunkt liegt ein Sattelpunkt von f vor. Für alle übrigen Werte $c < 1$, $c \neq 0$, sind die Niveaumengen $N_c(f)$ glatte Kurven, denn auf diesen Mengen liegen keine kritischen Punkte von f .
- Sei $a > 0$ vorgegeben und $f(x, y) = (x^2 + y^2)^2 - 2ax(x^2 + y^2) - a^2y^2$. Die Funktion f hat zwei kritische Stellen, nämlich $p = (0, 0)$ und $q = (3a/2, 0)$. Im Nullpunkt p ist ein Sattelpunkt und im Punkt q ein isoliertes Minimum. Der Graph von f erinnert an eine Sitzbadewanne. Die Niveaulinie N_0 durch p ist die oben beschriebene Kardioide (siehe Beispiel 6.3.3). Sie hat im Nullpunkt eine Spitze und ist ansonsten glatt. Alle übrigen nichtleeren Niveaumengen sind glatte Kurven.

Noch allgemeiner kann man von einem System von Gleichungen ausgehen, das in eine explizite Form gebracht werden soll, indem man nach möglichst vielen Variablen auflöst.

6.3.9 BEISPIEL Sei $f_1(x, y, z) = x - y$ und $f_2(x, y, z) = x^2 + y^2 - z^2 - 1$. Die Gleichung $f_1(x, y, z) = 0$ beschreibt eine Ebene E im \mathbb{R}^3 , die die z -Achse enthält, und die Gleichung $f_2(x, y, z) = 0$, wie eben untersucht, ein Hyperboloid. Die Lösungsmenge des Systems aus beiden Gleichungen ist also die Schnittmenge des Hyperboloids mit der Ebene E . Es handelt sich dabei um eine Hyperbel in der Ebene E , die um 90° gedreht erscheint. Interessieren wir uns nur für den Ausschnitt der Hyperbel in der Nähe des Punktes $(1, 1, 1)$, so können wir die Gleichungen folgendermassen nach y bzw. z auflösen: Ein Punkt (x, y, z) mit $x > \frac{\sqrt{2}}{2}$ und $z > 0$ erfüllt die Gleichungen $f_1(x, y, z) = 0$ und $f_2(x, y, z) = 0$ genau dann, wenn $y = x$ und $z = \sqrt{2x^2 - 1}$.

Es geht also um folgendes: Ein System aus m Gleichungen in $n + m$ Variablen der Form

$$\begin{aligned} f_1(x_1, \dots, x_n, y_1, \dots, y_m) &= 0 \\ f_2(x_1, \dots, x_n, y_1, \dots, y_m) &= 0 \\ &\vdots \\ f_m(x_1, \dots, x_n, y_1, \dots, y_m) &= 0 \end{aligned}$$

beschreibt wiederum auf implizite Art eine Teilmenge des \mathbb{R}^{n+m} . Wir erhalten eine explizite Beschreibung eines Ausschnitts der Lösungsmenge, wenn es gelingt, die Gleichungen nach y_1, \dots, y_m aufzulösen, so dass das System nun die folgende Gestalt annimmt:

$$\begin{aligned} y_1 &= g_1(x_1, \dots, x_n) \\ y_2 &= g_2(x_1, \dots, x_n) \\ &\vdots \quad \vdots \end{aligned}$$

$$y_m = g_m(x_1, \dots, x_n)$$

Wir können die Komponenten f_1, \dots, f_m zu einer Funktion $F: D \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ und die Komponenten g_1, \dots, g_m zu einer Funktion $G: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ zusammenfassen. Verwenden wir ausserdem die Notation $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_m)$, so lautet jetzt das Problem, zu gegebenem F ein passendes G zu finden, so dass lokal gilt:

$$F(x, y) = 0 \Leftrightarrow y = G(x).$$

Es geht es also wiederum darum, die Nullstellenmenge von F lokal als Graph darzustellen.

Ist F differenzierbar, so wird das Differential $DF(x, y)$ der Funktion F an einer Stelle $(x, y) \in D$ durch eine Matrix vom Typ $m \times (n + m)$ beschrieben. Wir können diese Matrix also in zwei Blöcke unterteilen, eine Teilmatrix vom Typ $m \times n$, die wir mit $D_x F(x, y)$ bezeichnen, und eine quadratische Teilmatrix vom Typ $m \times m$, die wir mit $D_y F(x, y)$ bezeichnen. $D_x F(x, y)$ ist das Differential derjenigen Funktion, die wir aus F erhalten, indem wir y konstant halten, und $D_y F(x, y)$ ist das Differential der Funktion, die durch F definiert wird, indem wir x fest halten. Nehmen wir an, dass $F(x, G(x)) = 0$ für alle x , und sind F und G differenzierbar, so folgt aus der Kettenregel

$$D_x F(x, G(x)) + D_y F(x, G(x)) \circ DG(x) = 0 \quad \text{für alle } x.$$

Daraus ergibt sich, falls $D_y F(x, G(x))$ invertierbar ist, die Beziehung

$$DG(x) = -(D_y F(x, G(x))^{-1} \circ D_x F(x, G(x))).$$

Der Satz über implizite Funktionen in allgemeiner Form lautet mit diesen Bezeichnungen:

6.3.10 SATZ Sei $F: D \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig differenzierbar, D offen, $F(p, q) = 0$ für ein $(p, q) \in D$. Sei weiter $D_y F(p, q)$ invertierbar. Dann gibt es offene Umgebungen $p \in U \subset \mathbb{R}^n$ und $q \in V \subset \mathbb{R}^m$ mit $U \times V \subset D$ und eine eindeutig bestimmte, stetig differenzierbare Funktion $G: U \rightarrow V$ mit $G(p) = q$, so dass für alle $(x, y) \in U \times V$ gilt

$$F(x, y) = 0 \Leftrightarrow y = G(x).$$

Ausserdem ist

$$DG(x) = -(D_y F(x, G(x)))^{-1} \cdot D_x F(x, G(x)) \quad \text{für alle } x \in U.$$

Auf den schwierigen Beweis dieses Satzes müssen wir hier verzichten.

Im Spezialfall $m = 1$ stimmt die Aussage mit der früheren überein. Denn in diesem Fall sind F und G reellwertige Funktionen. Also ist $DF(x, y) = \text{grad } F(x, y)$, $D_x F(x, y) = (\partial_x F(x, y), \dots, \partial_n F(x, y))$ und $D_y F(p, q) = \partial_{n+1} F(p, q)$. Die Bedingung an $D_y F(p, q)$ lautet in diesem Fall also einfach $\partial_{n+1} F(p, q) \neq 0$. Und für G erhalten wir die Aussage:

$$\text{grad } G(x) = -\frac{1}{\partial_{n+1} F(x, G(x))} \cdot (\partial_x F(x, y), \dots, \partial_n F(x, y)).$$

Schauen wir uns ausserdem unser Beispiel noch einmal an.

6.3.11 BEISPIEL Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, definiert durch $f_1(x, y, z) = x - y$ und $f_2(x, y, z) = x^2 + y^2 - z^2 - 1$. Das Differential von F ist gegeben durch die Jacobimatrix

$$JF(x, y, z) = \begin{pmatrix} 1 & -1 & 0 \\ 2x & 2y & -2z \end{pmatrix}.$$

Hier ist jetzt $m = 2$ und $n = 1$, wir wollen nach y und z auflösen in der Nähe des Punktes $a = (1, 1, 1)$. Das Kriterium des Satzes verlangt dazu, dass die Teilmatrix $D_{(y,z)}F(a) = \begin{pmatrix} -1 & 0 \\ 2 & -2 \end{pmatrix}$ invertierbar ist. Dies ist gewährleistet, denn die Determinante der Matrix beträgt 2. Die gesuchte Funktion $G: U \rightarrow V$ ist gegeben durch

$$G(x) = \begin{pmatrix} x \\ \sqrt{2x^2 - 1} \end{pmatrix},$$

wobei wir $U = (\frac{\sqrt{2}}{2}, \infty) \subset \mathbb{R}$ und $V = \{(y, z) \in \mathbb{R}^2 \mid z > 0\}$ wählen können. Das Differential von G lautet $DG(x) = \begin{pmatrix} 1 \\ \frac{2x}{\sqrt{2x^2 - 1}} \end{pmatrix}$ für $x \in U$. Vergleichen wir mit

$$\begin{aligned} -(D_{(y,z)}F(x, G(x)))^{-1} \cdot D_x F(x, G(x)) &= - \begin{pmatrix} -1 & 0 \\ 2g_1(x) & -2g_2(x) \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \\ 2x \end{pmatrix} = \\ \frac{1}{2g_2(x)} \begin{pmatrix} 2g_2(x) & 0 \\ 2g_1(x) & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2x \end{pmatrix} &= \begin{pmatrix} 1 \\ \frac{2g_1(x) + 2x}{2g_2(x)} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{4x}{2g_2(x)} \end{pmatrix}. \end{aligned}$$

Wieder stellen wir Übereinstimmung fest.

Der im vorigen Abschnitt formulierte Umkehrsatz kann mithilfe des Satzes über implizite Funktionen jetzt bewiesen werden.

6.3.12 FOLGERUNG Sei $D \subset \mathbb{R}^n$ offen und $f: D \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Sei weiter $q \in D$ ein Punkt mit $\det Df_q \neq 0$. Dann gibt es eine offene Umgebung $U \subset D$ von q , so dass f , aufgefasst als Abbildung von U nach $V := f(U)$ bijektiv ist und die Umkehrabbildung $f^{-1}: V \rightarrow U$ ebenfalls differenzierbar ist.

Beweis. Wir definieren die Funktion $F: \mathbb{R}^n \times D \rightarrow \mathbb{R}^n$ durch $F(x, y) = x - f(y)$ für alle $y \in D$, $x \in \mathbb{R}^n$. Sei $p := f(q)$. Offenbar ist $F(p, q) = 0$. Ausserdem ist $(D_y F)_{(p,q)} = -Df_q$ nach unserer Annahme invertierbar. Damit sind die Voraussetzungen des Satzes über implizite Funktionen erfüllt, und es gibt daher offene Teilmengen $p \in U \subset \mathbb{R}^n$ und $q \in D \subset \mathbb{R}^n$ und eine eindeutig bestimmte, stetig differenzierbare Funktion $G: U \rightarrow V$ mit $G(p) = q$, so dass für alle $(x, y) \in U \times V$ gilt

$$F(x, y) = x - f(y) = 0 \iff y = G(x).$$

Die Funktion G ist gerade die Umkehrfunktion von f . q.e.d.