1.5 Der Umkehrsatz

Definition

Sei $B \subset \mathbb{R}^n$ offen. Eine Abbildung $\mathbf{f} = (f_1, \ldots, f_m) : B \to \mathbb{R}^m$ heißt in $\mathbf{a} \in B$ ***differenzierbar***, falls alle Komponentenfunktionen $f_1, \ldots, f_m$ in $\mathbf{a}$ differenzierbar sind.

Die durch $D\mathbf{f}(\mathbf{a})(\mathbf{v}) := (Df_1(\mathbf{a})(\mathbf{v}), \ldots, Df_m(\mathbf{a})(\mathbf{v}))$ gegebene lineare Abbildung

$$D\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m$$

heißt die ***Ableitung von f in a***. Die Matrix $J_\mathbf{f}(\mathbf{a}) \in M_{m,n}(\mathbb{R})$, die $D\mathbf{f}(\mathbf{a})$ (bezüglich der Standardbasen) beschreibt, nennt man die ***Funktionalmatrix*** oder ***Jacobi-Matrix*** von $\mathbf{f}$ in $\mathbf{a}$. (Es ist dann $D\mathbf{f}(\mathbf{a})(\mathbf{v}) = \mathbf{v} \cdot J_\mathbf{f}(\mathbf{a})^\top$).

Die j-te Spalte der Funktionalmatrix $J_\mathbf{f}(\mathbf{a})$ ist der Vektor

$$J_\mathbf{f}(\mathbf{a}) \cdot \mathbf{e}_j^\top = (D\mathbf{f}(\mathbf{a})(\mathbf{e}_j))^\top = \Big(\frac{\partial f_1}{\partial x_j}(\mathbf{a}), \ldots, \frac{\partial f_m}{\partial x_j}(\mathbf{a})\Big)^\top.$$

So erhält man:

5.1. Gestalt der Funktionalmatrix

$$J_\mathbf{f}(\mathbf{a}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} \nabla f_1(\mathbf{a}) \\ \vdots \\ \nabla f_m(\mathbf{a}) \end{pmatrix}.$$

Definition

Ist $n = m$, also $J_\mathbf{f}(\mathbf{x})$ eine quadratische Matrix, so heißt $\det J_\mathbf{f}(\mathbf{x})$ die ***Funktionaldeterminante*** oder ***Jacobi-Determinante*** von $\mathbf{f}$ in $\mathbf{x}$.

5.2. Beispiele

A. Ist $n = m = 1$, so ist $J_f(a) = f'(a)$ die gewöhnliche Ableitung.

B. Ist n beliebig und $m = 1$, so besitzt die skalare Funktion f nur eine Komponente. Also ist $J_f(\mathbf{a}) = \Big(\dfrac{\partial f}{\partial x_1}(\mathbf{a}), \cdots, \dfrac{\partial f}{\partial x_n}(\mathbf{a})\Big) = \nabla f(\mathbf{a})$.

C. Ist $n = 1$ und m beliebig, so ist $\mathbf{f} = (f_1, \dots, f_m)$ ein differenzierbarer Weg im $\mathbb{R}^n$, mit m Komponenten, der aber nur von einer Variablen abhängt. Weil die verschiedenen Komponenten in verschiedenen Zeilen der Jacobi-Matrix stehen müssen, ist zwar $J_{\mathbf{f}}(\mathbf{a}) = \mathbf{f}'(\mathbf{a})^\top$ die gewöhnliche Ableitung, aber als **Spaltenvektor** geschrieben!

D. Ist $\mathbf{a} \in \mathbb{R}^n$, so ist die ***Translation*** $\mathbf{T}_{\mathbf{a}} : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$ eine differenzierbare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$. Ist $\mathbf{a} = (a_1, \dots, a_n)$, so ist

$$\mathbf{T}_{\mathbf{a}}(x_1, \dots, x_n) = (x_1 + a_1, \dots, x_n + a_n)$$

und daher $J_{\mathbf{T}_{\mathbf{a}}}(\mathbf{x}) = E_n$ die Einheitsmatrix und $\det J_{\mathbf{T}_{\mathbf{a}}}(\mathbf{x}) = 1$ (beides unabhängig von $\mathbf{x}$).

E. Sei $A \in M_{m,n}(\mathbb{R})$ eine beliebige Matrix, $\mathbf{f}_A : \mathbb{R}^n \to \mathbb{R}^m$ die durch

$$\mathbf{f}_A(\mathbf{x}) := \mathbf{x} \cdot A^\top$$

definierte zugeordnete lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$.

Sind $\mathbf{a}_1, \dots, \mathbf{a}_n$ die Zeilen von A, so ist $\mathbf{f}_A(\mathbf{x}) = (\mathbf{x} \bullet \mathbf{a}_1, \dots, \mathbf{x} \bullet \mathbf{a}_n)$. Da die Ableitung einer Linearform mit eben dieser Linearform übereinstimmt, ist

$$D\mathbf{f}_A(\mathbf{x})(\mathbf{v}) = (\mathbf{v} \bullet \mathbf{a}_1, \dots, \mathbf{v} \bullet \mathbf{a}_n) = \mathbf{v} \cdot A^\top = \mathbf{f}_A(\mathbf{v}),$$

also $J_{\mathbf{f}_A}(\mathbf{x}) = A$, unabhängig von $\mathbf{x}$. Die Funktionaldeterminante kann natürlich nur gebildet werden, wenn $n = m$ ist.

F. Sei $\mathbf{f}(x, y) := (e^{kx} \cos y, e^{kx} \sin y)$. Dann gilt:

$$J_{\mathbf{f}}(x, y) = \begin{pmatrix} ke^{kx} \cos y & -e^{kx} \sin y \\ ke^{kx} \sin y & e^{kx} \cos y \end{pmatrix}$$

und

$$\det J_{\mathbf{f}}(x, y) = ke^{2kx} \cos^2 y + ke^{2kx} \sin^2 y = ke^{2kx}.$$

Wie bei den skalaren Funktionen steht auch für Abbildungen ein alternatives Kriterium für die Differenzierbarkeit zur Verfügung.

5.3. Grauert-Kriterium für die Differenzierbarkeit

Sei $B \subset \mathbb{R}^n$ offen. Eine Abbildung $\mathbf{f} : B \to \mathbb{R}^m$ ist genau dann in $\mathbf{x}_0 \in B$ (total) differenzierbar, wenn es eine in $\mathbf{x}_0$ stetige Abbildung $\Delta : B \to M_{m,n}(\mathbb{R})$ gibt, so dass gilt:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \cdot \Delta(\mathbf{x})^\top.$$

Speziell ist dann $\Delta(\mathbf{x}_0) = J_{\mathbf{f}}(\mathbf{x}_0)$.

BEWEIS: Die Abbildung $\mathbf{f} = (f_1, \dots, f_m)$ ist genau dann in $\mathbf{x}_0$ differenzierbar, wenn alle Komponentenfunktionen f_μ es sind, wenn es also in $\mathbf{x}_0$ stetige Funktionen $\boldsymbol{\Delta}_\nu : B \to \mathbb{R}^n$ gibt, so dass $f_\mu(\mathbf{x}) = f_\mu(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \cdot \boldsymbol{\Delta}_\mu(\mathbf{x})^\top$ für $\mu = 1, \dots, m$ gilt. Wir definieren dann $\Delta(\mathbf{x})$ als die Matrix, deren Zeilen die Vektoren $\boldsymbol{\Delta}_\mu(\mathbf{x})$ sind. Es ist klar, dass auch umgekehrt aus dem Kriterium die Differenzierbarkeit folgt. ∎

5.4. Allgemeine Kettenregel

Sei $B \subset \mathbb{R}^n$ offen, $\mathbf{f} : B \to \mathbb{R}^m$ in $\mathbf{x}_0 \in B$ differenzierbar, $U \subset \mathbb{R}^m$ offen, $\mathbf{f}(B) \subset U$ und $\mathbf{g} : U \to \mathbb{R}^k$ in $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ differenzierbar. Dann ist $\mathbf{g} \circ \mathbf{f} : B \to \mathbb{R}^k$ in $\mathbf{x}_0$ differenzierbar und es gilt:

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \circ D\mathbf{f}(\mathbf{x}_0)$$

bzw.

$$J_{\mathbf{g} \circ \mathbf{f}}(\mathbf{x}_0) = J_{\mathbf{g}}(\mathbf{f}(\mathbf{x}_0)) \cdot J_{\mathbf{f}}(\mathbf{x}_0).$$

BEWEIS: Wir haben Darstellungen

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \mathbf{f}(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \cdot \boldsymbol{\Delta}(\mathbf{x})^\top \\ \text{und} \quad \mathbf{g}(\mathbf{y}) &= \mathbf{g}(\mathbf{y}_0) + (\mathbf{y} - \mathbf{y}_0) \cdot \boldsymbol{\Delta}^*(\mathbf{y})^\top, \end{aligned}$$

wobei jeweils $\boldsymbol{\Delta}$ in $\mathbf{x}_0$ und $\boldsymbol{\Delta}^*$ in $\mathbf{y}_0$ stetig ist. Setzt man die Gleichungen ineinander ein, so erhält man

$$\begin{aligned} \mathbf{g} \circ \mathbf{f}(\mathbf{x}) - \mathbf{g} \circ \mathbf{f}(\mathbf{x}_0) &= (\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)) \cdot \boldsymbol{\Delta}^*(\mathbf{f}(\mathbf{x}))^\top \\ &= (\mathbf{x} - \mathbf{x}_0) \cdot \boldsymbol{\Delta}(\mathbf{x})^\top \cdot \boldsymbol{\Delta}^*(\mathbf{f}(\mathbf{x}))^\top \\ &= (\mathbf{x} - \mathbf{x}_0) \cdot (\boldsymbol{\Delta}^*(\mathbf{f}(\mathbf{x})) \cdot \boldsymbol{\Delta}(\mathbf{x}))^\top, \end{aligned}$$

mit einer in $\mathbf{x}_0$ stetigen Funktion $\mathbf{x} \mapsto \boldsymbol{\Delta}^*(\mathbf{f}(\mathbf{x})) \cdot \boldsymbol{\Delta}(\mathbf{x})$. Das zeigt, dass $\mathbf{g} \circ \mathbf{f}$ in $\mathbf{x}_0$ differenzierbar ist. Weil $\boldsymbol{\Delta}(\mathbf{x}_0) = J_{\mathbf{f}}(\mathbf{x}_0)$ und $\boldsymbol{\Delta}^*(\mathbf{y}_0) = J_{\mathbf{g}}(\mathbf{y}_0)$ ist, folgt die Gleichung

$$J_{\mathbf{g} \circ \mathbf{f}}(\mathbf{x}_0) = \boldsymbol{\Delta}^*(\mathbf{f}(\mathbf{x}_0)) \cdot \boldsymbol{\Delta}(\mathbf{x}_0) = J_{\mathbf{g}}(\mathbf{f}(\mathbf{x}_0)) \cdot J_{\mathbf{f}}(\mathbf{x}_0).$$

Für die zugehörigen linearen Abbildungen gilt dann die analoge Beziehung

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \circ D\mathbf{g}(\mathbf{x}_0).$$

∎

5.5. Folgerung

$$\textit{Ist } n = m = k, \ \textit{so ist} \ \det J_{\mathbf{g} \circ \mathbf{f}}(\mathbf{x}) = \det J_{\mathbf{g}}(\mathbf{f}(\mathbf{x})) \cdot \det J_{\mathbf{f}}(\mathbf{x}).$$

Der BEWEIS ergibt sich unmittelbar aus dem Determinanten-Produktsatz.

5.6. Beispiele

A. Ist $k = 1$, also $g : U \to \mathbb{R}$ eine skalare Funktion, so ist auch $g \circ \mathbf{f}$ eine skalare Funktion und man erhält die Formel

$$\boxed{\nabla(g \circ \mathbf{f})(\mathbf{x}) = \nabla g(\mathbf{f}(\mathbf{x})) \cdot J_{\mathbf{f}}(\mathbf{x}).}$$

Mit $\mathbf{f} = (f_1, \ldots, f_m)$ folgt dann für die einzelnen Komponenten die Formel

$$\boxed{(g \circ \mathbf{f})_{x_\nu} = (g_{y_1} \circ \mathbf{f}) \cdot (f_1)_{x_\nu} + \cdots + (g_{y_m} \circ \mathbf{f}) \cdot (f_m)_{x_\nu}, \text{ für } \nu = 1, \ldots, n.}$$

Um es noch deutlicher zu machen, betrachten wir den Fall $n = m = 2$ und bezeichnen die Variablen, von denen g abhängt, mit x und y und die Variablen, von denen $\mathbf{f}$ abhängt, mit u und v. Dann schreibt sich die obige Formel wie folgt:

$$\begin{aligned} \frac{\partial(g \circ \mathbf{f})}{\partial u} &= \left(\frac{\partial g}{\partial x} \circ \mathbf{f}\right) \cdot \frac{\partial f_1}{\partial u} + \left(\frac{\partial g}{\partial y} \circ \mathbf{f}\right) \cdot \frac{\partial f_2}{\partial u} \\ \text{und} \quad \frac{\partial(g \circ \mathbf{f})}{\partial v} &= \left(\frac{\partial g}{\partial x} \circ \mathbf{f}\right) \cdot \frac{\partial f_1}{\partial v} + \left(\frac{\partial g}{\partial y} \circ \mathbf{f}\right) \cdot \frac{\partial f_2}{\partial v}. \end{aligned}$$

B. Sei $g(x, y) := 1/(x^2 - y^2)$ und $\mathbf{f}(r, t) := (r \cos t, r \sin t)$. Es ist

$$g_x = \frac{-2x}{(x^2 - y^2)^2} \quad \text{und} \quad g_y = \frac{2y}{(x^2 - y^2)^2}\,,$$

sowie

$$J_{\mathbf{f}}(r, t) = \begin{pmatrix} \cos t & -r \sin t \\ \sin t & r \cos t \end{pmatrix}.$$

Dann folgt:

$$\begin{aligned} (g \circ \mathbf{f})_r &= (g_x \circ \mathbf{f}) \cdot (f_1)_r + (g_y \circ \mathbf{f}) \cdot (f_2)_r \\ &= \frac{-2r \cos t}{r^4(\cos^2 t - \sin^2 t)^2} \cdot \cos t + \frac{2r \sin t}{r^4(\cos^2 t - \sin^2 t)^2} \cdot \sin t \\ &= \frac{-2}{r^3(\cos^2 t - \sin^2 t)} \\ \text{und} \quad (g \circ \mathbf{f})_t &= (g_x \circ \mathbf{f}) \cdot (f_1)_t + (g_y \circ \mathbf{f}) \cdot (f_2)_t \\ &= \frac{-2r \cos t \cdot (-r \sin t)}{r^4(\cos^2 t - \sin^2 t)^2} + \frac{2r \sin t \cdot (r \cos t)}{r^4(\cos^2 t - \sin^2 t)^2} \\ &= \frac{4 \sin t \cos t}{r^2(\cos^2 t - \sin^2 t)^2}\,. \end{aligned}$$

Bei solchen konkreten Aufgaben kann man natürlich auch $\mathbf{f}$ zuerst in g einsetzen und die dann entstandende Funktion von u und v direkt differenzieren. Welcher Weg einfacher ist, muss man von Fall zu Fall prüfen.

C. Sei $M \subset \mathbb{R}^n$ eine offene Teilmenge mit der Eigenschaft, dass mit $\mathbf{x} \in M$ und $\lambda \in \mathbb{R}$ auch $\lambda\mathbf{x}$ zu M gehört (man nennt eine solche Menge M auch eine „Kegelmenge“). Eine differenzierbare Funktion $f : M \to \mathbb{R}$ heißt ***homogen*** vom Grad p, falls $f(\lambda\mathbf{x}) = \lambda^p \cdot f(\mathbf{x})$ für jedes $\mathbf{x} \in M$ und jedes $\lambda \in \mathbb{R}$ gilt.

Wir betrachten die Funktion $g(\lambda) := f(\lambda\mathbf{x})$ für ein festes $\mathbf{x}$. Nach der Kettenregel ist

$$g'(1) = \nabla f(\mathbf{x}) \cdot \mathbf{x}^\top.$$

Wegen der Homogenität von f ist aber auch $g(\lambda) = \lambda^p \cdot f(\mathbf{x})$, also

$$g'(1) = \big(p \cdot \lambda^{p-1} \cdot f(\mathbf{x})\big)|_{\lambda=1} = p \cdot f(\mathbf{x}).$$

Zusammen ergibt das die ***Euler'sche Homogenitätsgleichung***

Ist f homogen vom Grad p, so ist $\quad \nabla f(\mathbf{x}) \bullet \mathbf{x} = p \cdot f(\mathbf{x})$.

Zum Beispiel ist $f(x, y) := x^4 + y^4 - 4x^2y^2$ eine homogene Funktion vom Grad 4 auf dem $\mathbb{R}^2$. Es ist $\nabla f(x, y) = (4x^3 - 8xy^2, 4y^3 - 8x^2y)$, also $\nabla f(x, y) \bullet (x, y) = 4x^4 + 4y^4 - 16x^2y^2 = 4\, f(x, y)$.

Definition

Es seien $G_1, G_2 \subset \mathbb{R}^n$ Gebiete und $\mathbf{f} : G_1 \to G_2$ eine differenzierbare Abbildung. $\mathbf{f}$ heißt ein ***Diffeomorphismus***, wenn $\mathbf{f}$ bijektiv und $\mathbf{f}^{-1} : G_2 \to G_1$ ebenfalls differenzierbar ist.

Bemerkung: Ist $\mathbf{f} : G_1 \to G_2$ ein Diffeomorphismus, so ist einerseits $J_{\mathbf{f}^{-1}\circ\mathbf{f}}(\mathbf{x}) = J_{\text{id}}(\mathbf{x}) = E_n$ und andererseits $J_{\mathbf{f}^{-1}\circ\mathbf{f}}(\mathbf{x}) = J_{\mathbf{f}^{-1}}(\mathbf{f}(\mathbf{x})) \cdot J_{\mathbf{f}}(\mathbf{x})$, also

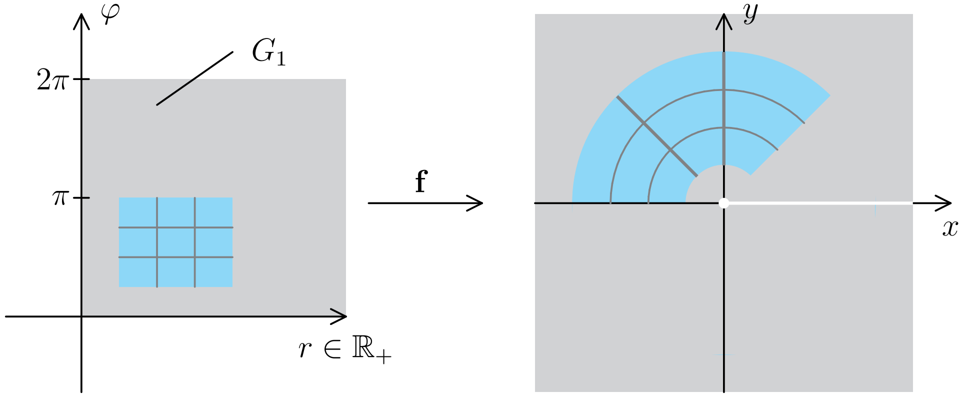
$$J_{\mathbf{f}^{-1}}(\mathbf{f}(\mathbf{x})) = J_{\mathbf{f}}(\mathbf{x})^{-1}.$$

5.7. Beispiel

Wir betrachten die ebenen ***Polarkoordinaten***

$$(x, y) = \mathbf{f}(r, \varphi) = (r\cos\varphi, r\sin\varphi).$$

Definitionsbereich ist $\mathbb{R}_+ \times \mathbb{R}$, die Bildmenge ist $\mathbb{R}^2 \setminus \{(0, 0)\}$. Leider ist $\mathbf{f}$ nicht injektiv, es ist ja $\mathbf{f}(r, \varphi) = \mathbf{f}(r, \varphi + 2\pi)$. Die Menge $\mathbb{R}_+ \times [0, 2\pi)$ ist kein Gebiet, weil sie nicht offen ist. Also benutzen wir als Definitionsbereich das Gebiet $G_1 := \mathbb{R}_+ \times (0, 2\pi)$.



Jetzt ist $\mathbf{f} : G_1 \to \mathbb{R}^2$ injektiv, aber was ist die Bildmenge? Nach wie vor kommt jeder Punkt $(x, y) \in \mathbb{R}^2$ als Bildpunkt vor, sofern er nicht auf der positiven x-Achse liegt. Also setzen wir $G_2 := \mathbb{R}^2 \setminus \{(x, y) \,:\, y = 0 \text{ und } x \geq 0\}$. Dann ist $\mathbf{f} : G_1 \to G_2$ eine bijektive differenzierbare Abbildung.

Ist $\mathbf{f}$ nun auch ein Diffeomorphismus? Wir versuchen, die Umkehrabbildung zu bestimmen, d.h. zu einem gegebenen Punkt $(x, y) \in G_2$ suchen wir ein $(r, \varphi) \in G_1$ mit

$$(x, y) = (r \cos\varphi, r \sin\varphi).$$

Dann ist auf jeden Fall $x^2 + y^2 = r^2$, also $r(x, y) = \sqrt{x^2 + y^2}$.

Ist $x \neq 0$, so ist $\tan\varphi = y/x$. Daraus folgt aber nicht, dass $\varphi = \arctan(y/x)$ ist, denn der Arcustangens nimmt nur Werte zwischen $-\pi/2$ und $+\pi/2$ an, während φ zwischen 0 und 2π liegen soll. Außerdem wird der Fall $x = 0$ dabei noch nicht berücksichtigt.

Wir müssen also etwas sorgfältiger vorgehen. Dazu führen wir die Halbebenen $H_+ = \{(x, y) \in \mathbb{R}^2 \,:\, y > 0\}$, $H_- = \{(x, y) \in \mathbb{R}^2 \,:\, y < 0\}$ und $H_0 = \{(x, y) \in \mathbb{R}^2 \,:\, x < 0\}$ ein. Sie sind offene Mengen, und es ist $H_0 \cup H_+ \cup H_- = G_2$.

Im folgenden verwenden wir einige Formeln aus der Trigonometrie:

- $\cos\left(\dfrac{\pi}{2} - t\right) = \sin t$ und $\sin\left(\dfrac{\pi}{2} - t\right) = \cos t$,
- $\arctan t = \arcsin \dfrac{t}{\sqrt{1 + t^2}} = \arccos \dfrac{1}{\sqrt{1 + t^2}}$,
- $\arctan(1/t) = \pm\dfrac{\pi}{2} - \arctan t$, je nachdem, ob $t > 0$ oder $t < 0$ ist.

1. Fall: Ist $(x, y) \in H_+$, so setzen wir

$$\varphi_+(x, y) := \frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) \in (0, \pi).$$

Dann ist

$$\begin{aligned} \cos(\varphi_+(x,y)) &= \sin\Big(\arcsin\frac{x/y}{\sqrt{1+x^2/y^2}}\Big) = \frac{x}{\sqrt{x^2+y^2}} \\ \text{und} \quad \sin(\varphi_+(x,y)) &= \cos\Big(\arccos\frac{1}{\sqrt{1+x^2/y^2}}\Big) = \frac{y}{\sqrt{x^2+y^2}}\,. \end{aligned}$$

Offensichtlich ist $(x,y) \mapsto (r(x,y), \varphi_+(x,y))$ eine Umkehrung der Polarkoordinaten, denn es ist

$$\mathbf{f}\big(r(x,y), \varphi_+(x,y)\big) = (x,y), \quad r(r\cos\varphi, r\sin\varphi) = r$$

und

$$\varphi_+(r\cos\varphi, r\sin\varphi) = \frac{\pi}{2} - \arctan\Big(\frac{1}{\tan\varphi}\Big) = \frac{\pi}{2} - \Big(\frac{\pi}{2} - \arctan\tan\varphi\Big) = \varphi.$$

2. Fall: Ist $(x,y) \in H_-$, so setzen wir

$$\varphi_-(x,y) := \frac{3\pi}{2} - \arctan\Big(\frac{x}{y}\Big) \in (\pi, 2\pi).$$

3. Fall: Ist $(x,y) \in H_0$, so setzen wir

$$\varphi_0(x,y) := \pi + \arctan\Big(\frac{y}{x}\Big) \in \Big(\frac{\pi}{2}, \frac{3\pi}{2}\Big).$$

Auch in diesen beiden Fällen erhält man eine Umkehrung der Polarkoordinaten, und die Funktionen $\varphi_0, \varphi_+, \varphi_-$ sind differenzierbar.

Ist $y/x < 0$, so ist $\arctan(y/x) = -\pi/2 - \arctan(x/y)$. Auf $H_0 \cap H_+$ ist $x < 0$, $y > 0$ und deshalb $\varphi_0(x,y) = \varphi_+(x,y)$.

Ist $y/x > 0$, so ist $\arctan(y/x) = \pi/2 - \arctan(x/y)$. Auf $H_0 \cap H_-$ ist $x < 0$, $y < 0$ und deshalb $\varphi_0(x,y) = \varphi_-(x,y)$.

Zusammen ergibt das eine differenzierbare Umkehrabbildung $\mathbf{f}^{-1} : G_2 \to G_1$.

5.8. Lemma

Sei $M \subset \mathbb{R}^n$ offen, $\mathbf{F} : M \to \mathbb{R}^n$ in $\mathbf{x}_0$ differenzierbar und $\det J_{\mathbf{F}}(\mathbf{x}_0) \neq 0$. Außerdem habe $f := \|\mathbf{F}\|^2$ in $\mathbf{x}_0$ ein lokales Minimum. Dann ist $\mathbf{F}(\mathbf{x}_0) = \mathbf{0}$.

BEWEIS: Sei $\mathbf{v} \in \mathbb{R}^n$ und $\boldsymbol{\alpha}(t) := \mathbf{x}_0 + t\mathbf{v}$ für $-\varepsilon < t < \varepsilon$. Dabei sei $\varepsilon > 0$ so klein gewählt, dass die Spur von $\boldsymbol{\alpha}$ in M liegt. Weil f in $\mathbf{x}_0$ ein Minimum besitzt, ist

$$\begin{aligned} 0 = \nabla f(\mathbf{x}_0) \bullet \mathbf{v} &= (f \circ \boldsymbol{\alpha})'(0) = \frac{d}{dt}\Big|_0 \big((\mathbf{F} \circ \boldsymbol{\alpha}(t)) \bullet (\mathbf{F} \circ \boldsymbol{\alpha}(t))\big) \\ &= 2 \cdot \big(\mathbf{F} \circ \boldsymbol{\alpha}(0)\big) \bullet \big((\mathbf{F} \circ \boldsymbol{\alpha})'(0)\big) = 2 \cdot \mathbf{F}(\mathbf{x}_0) \bullet \big(J_{\mathbf{F}}(\mathbf{x}_0) \cdot \mathbf{v}^\top\big) \\ &= 2 \cdot \big(\mathbf{F}(\mathbf{x}_0)\big) \bullet \big(D\mathbf{F}(\mathbf{x}_0)(\mathbf{v})\big). \end{aligned}$$

Weil $D\mathbf{F}(\mathbf{x}_0)$ surjektiv ist, ist $\mathbf{F}(\mathbf{x}_0) \bullet \mathbf{w} = 0$ für alle $\mathbf{w} \in \mathbb{R}^n$, also $\mathbf{F}(\mathbf{x}_0) = \mathbf{0}$. ∎

5.9. Schrankensatz

Sei $B \subset \mathbb{R}^n$ offen, $K \subset B$ kompakt und konvex, $\mathbf{F} : B \to \mathbb{R}^m$ stetig differenzierbar und $C := \sup_K \|D\mathbf{F}(\mathbf{z})\|_{\text{op}}$. Dann gilt für $\mathbf{x}, \mathbf{y} \in K$:

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\| \le C \cdot \|\mathbf{x} - \mathbf{y}\|.$$

Beweis: Die Punkte auf der Verbindungsstrecke von $\mathbf{x}$ und $\mathbf{y}$ liegen in K. Nun sei $\mathbf{f} : [0,1] \to \mathbb{R}^m$ definiert durch $\mathbf{f}(t) := \mathbf{F}\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big)$. Dann ist

$$\begin{aligned}
\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\| &= \|\mathbf{f}(1) - \mathbf{f}(0)\| = \|\int_0^1 \mathbf{f}'(t)\,dt\| \\
&\le \int_0^1 \|\mathbf{f}'(t)\|\,dt = \int_0^1 \|D\mathbf{F}\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big)(\mathbf{x} - \mathbf{y})\|\,dt \\
&\le \|\mathbf{x} - \mathbf{y}\| \cdot \int_0^1 \|D\mathbf{F}\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big)\|_{\text{op}}\,dt \le C \cdot \|\mathbf{x} - \mathbf{y}\|.
\end{aligned}$$

■

5.10. Satz von der Umkehrabbildung

Sei $M \subset \mathbb{R}^n$ offen, $\mathbf{f} : M \to \mathbb{R}^n$ stetig differenzierbar. Ist $\mathbf{x}_0 \in M$, $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$ und $\det J_{\mathbf{f}}(\mathbf{x}_0) \ne 0$, so gibt es offene Umgebungen $U(\mathbf{x}_0) \subset M$ und $V(\mathbf{y}_0) \subset \mathbb{R}^n$, so dass gilt:

1. $\det J_{\mathbf{f}}(\mathbf{x}) \ne 0$ *für alle* $\mathbf{x} \in U$.
2. $\mathbf{f} : U \to V$ *ist bijektiv.*
3. $\mathbf{f}^{-1} : V \to U$ *ist wieder differenzierbar.*
4. *Für* $\mathbf{x} \in U$ *und* $\mathbf{y} = \mathbf{f}(\mathbf{x})$ *ist* $D\mathbf{f}^{-1}(\mathbf{y}) = (D\mathbf{f}(\mathbf{x}))^{-1}$.

Beweis: Ist $\det J_{\mathbf{f}}(\mathbf{x}) \ne 0$, so nennt man $\mathbf{f}$ ***regulär*** in $\mathbf{x}$. Die Funktionalmatrix $J_{\mathbf{f}}(\mathbf{x})$ ist dann eine invertierbare Matrix, und man kann die Umkehrmatrix $(J_{\mathbf{f}}(\mathbf{x}))^{-1}$ bilden.

Sei $L := D\mathbf{f}(\mathbf{x}_0)$, $k := \|L^{-1}\|_{\text{op}} > 0$ und $0 < c < 1/k$. Weil $\mathbf{f}$ stetig differenzierbar ist, gibt es ein $\varepsilon > 0$, so dass gilt:

$$\det J_{\mathbf{f}}(\mathbf{x}) \ne 0 \quad \text{und} \quad \|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{x}_0)\|_{\text{op}} < c \ \text{ für } \ \mathbf{x} \in B_\varepsilon(\mathbf{x}_0).$$

Wir setzen $U := B_\varepsilon(\mathbf{x}_0)$. Das ist eine (i.a. sehr kleine) offene, konvexe Umgebung von $\mathbf{x}_0$ in M, und die Aussage (1) gilt auf U.

1. Schritt: Die fundamentale Ungleichung

Für $\mathbf{w} \in \mathbb{R}^n$ ist $\|L^{-1}\mathbf{w}\| \le \|L^{-1}\|_{\text{op}} \cdot \|\mathbf{w}\| = k \cdot \|\mathbf{w}\|$. Da jeder Vektor $\mathbf{w} \in \mathbb{R}^n$ die Gestalt $\mathbf{w} = L\mathbf{v}$ besitzt, folgt für beliebige Vektoren $\mathbf{v} \in \mathbb{R}^n$ die Ungleichung

$$\|L\mathbf{v}\| \ge \frac{1}{k}\|\mathbf{v}\|.$$

Die Abbildung $\mathbf{h} := \mathbf{f} - L : M \to \mathbb{R}^n$ ist ebenfalls stetig differenzierbar, und es ist

$$\sup_{\mathbf{x}\in U}\|D\mathbf{h}(\mathbf{x})\|_{\text{op}} = \sup_{\mathbf{x}\in U}\|D\mathbf{f}(\mathbf{x}) - L\|_{\text{op}} = \sup_{\mathbf{x}\in U}\|D\mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x}_0)\|_{\text{op}} \le c.$$

Nach dem Schrankensatz ist nun

$$\|\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})\| \le c\|\mathbf{x} - \mathbf{y}\|, \text{ für alle } \mathbf{x}, \mathbf{y} \in U.$$

Weil $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) = \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y}) + L(\mathbf{x} - \mathbf{y})$ ist, folgt:

$$\begin{aligned} \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| &\ge \|L(\mathbf{x} - \mathbf{y})\| - \|\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})\| \\ &\ge \frac{1}{k}\|\mathbf{x} - \mathbf{y}\| - c\|\mathbf{x} - \mathbf{y}\| \\ &= \frac{1 - ck}{k}\|\mathbf{x} - \mathbf{y}\| \text{ für } \mathbf{x}, \mathbf{y} \in U. \end{aligned}$$

Dabei ist $0 < ck < 1$, also auch $0 < 1 - ck < 1$. Insbesondere ist $(1 - ck)/k > 0$.

2. Schritt: Die Injektivität von f auf U

Ist $\mathbf{x} \ne \mathbf{y}$, so ist $\|\mathbf{x} - \mathbf{y}\| > 0$, nach der fundamentalen Ungleichung also auch $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| > 0$, d.h. $\mathbf{f}(\mathbf{x}) \ne \mathbf{f}(\mathbf{y})$.

3. Schritt: Die Stetigkeit der Umkehrabbildung

Sei $V := \mathbf{f}(U)$, $\mathbf{v}_0 = \mathbf{f}(\mathbf{u}_0)$ ein Punkt in V, $\varepsilon > 0$ und $0 < \delta < (1 - ck)\varepsilon/k$. Ist $\mathbf{v} = \mathbf{f}(\mathbf{u}) \in V$ mit $\|\mathbf{v} - \mathbf{v}_0\| < \delta$, so folgt aus der fundamentalen Ungleichung:

$$\|\mathbf{f}^{-1}(\mathbf{v}) - \mathbf{f}^{-1}(\mathbf{v}_0)\| = \|\mathbf{u} - \mathbf{u}_0\| \le \frac{k}{1 - ck}\|\mathbf{v} - \mathbf{v}_0\| < \frac{k}{1 - ck}\,\delta < \varepsilon.$$

4. Schritt: Die Offenheit von $V = \mathbf{f}(U)$

Sei $\alpha := \dfrac{1 - ck}{2k}$. Dann ist $\dfrac{1 - ck}{k} - \alpha = 2\alpha - \alpha = \alpha$.

Sei $\mathbf{b}_0 = \mathbf{f}(\mathbf{a}_0) \in V$ ein beliebiger Punkt und $r > 0$ so gewählt, dass $B_r(\mathbf{a}_0) \subset\subset U$ ist. Wir setzen $\varepsilon := \alpha r$ und wollen zeigen, dass $B_\varepsilon(\mathbf{b}_0) \subset \mathbf{f}\big(B_r(\mathbf{a}_0)\big) \subset V$ ist.

Dazu sei $\mathbf{y}^* \in B_\varepsilon(\mathbf{b}_0)$ fest, aber beliebig, und $g(\mathbf{x}) := \|\mathbf{f}(\mathbf{x}) - \mathbf{y}^*\|$. Die stetige Funktion g nimmt auf $\overline{B_r(\mathbf{a}_0)}$ ihr Minimum an.

1. Es ist $g(\mathbf{a}_0) = \|\mathbf{f}(\mathbf{a}_0) - \mathbf{y}^*\| = \|\mathbf{b}_0 - \mathbf{y}^*\| < \varepsilon = \alpha r$.

2. Für $\mathbf{x} \in \partial B_r(\mathbf{a}_0)$ ist $\|\mathbf{x} - \mathbf{a}_0\| = r$, also

$$\begin{aligned} g(x) &\geq \|\mathbf{f}(\mathbf{x}) - \mathbf{b}_0\| - \|\mathbf{b}_0 - \mathbf{y}^*\| \\ &> \frac{1-ck}{k}\|\mathbf{x} - \mathbf{a}_0\| - \varepsilon \\ &= \frac{1-ck}{k} r - \alpha r \;=\; \alpha r. \end{aligned}$$

Also nimmt g sein Minimum in einem Punkt $\mathbf{x}^*$ im Innern der Kugel $B_r(\mathbf{a}_0)$ an. Dort wird auch $\|\mathbf{f}(\mathbf{x}) - \mathbf{y}^*\|^2$ minimal, und nach dem Lemma muss dann $\mathbf{f}(\mathbf{x}^*) = \mathbf{y}^*$ sein. Das bedeutet, dass $\mathbf{y}^*$ in $\mathbf{f}(B_r(\mathbf{a}_0))$ liegt, was zu beweisen war.

5. Schritt: Die Differenzierbarkeit der Umkehrabbildung

Jetzt beweisen wir, dass $\mathbf{f}^{-1}$ in jedem Punkt $\mathbf{y}_1 = \mathbf{f}(\mathbf{x}_1) \in V$ differenzierbar ist.

Da $\mathbf{f}$ in $\mathbf{x}_1$ differenzierbar ist, gibt es eine Darstellung

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_1) + (\mathbf{x} - \mathbf{x}_1) \cdot \Delta(\mathbf{x})^\top,$$

mit einer in $\mathbf{x}_1$ stetigen Abbildung $\Delta : U \to M_n(\mathbb{R})$. Dann ist $d(\mathbf{x}) := \det \Delta(\mathbf{x})$ in $\mathbf{x}_1$ stetig und $d(\mathbf{x}_1) \neq 0$. Es gibt also eine offene Umgebung von $\mathbf{x}_1$, auf der $d(\mathbf{x}) \neq 0$ ist. Das bedeutet, dass $\Delta(\mathbf{x})$ dort invertierbar ist.

Die Menge $G := \mathrm{GL}_n(\mathbb{R}) := \{A \in M_n(\mathbb{R}) \,:\, \det(A) \neq 0\}$ ist eine offene Teilmenge von $M_n(\mathbb{R})$ und die Abbildung $i : G \to G$ mit $i(A) := A^{-1}$ ist stetig, denn die Koeffizienten von A^{-1} sind rationale Funktionen der Koeffizienten von A (Cramer'sche Regel). Also ist auch $\Delta^*(\mathbf{y}) := i(\Delta(\mathbf{f}^{-1}(\mathbf{y}))) = \Delta(\mathbf{f}^{-1}(\mathbf{y}))^{-1}$ stetig in $\mathbf{y}_1$. Aus der Gleichung $(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_1)) \cdot (\Delta(\mathbf{x})^\top)^{-1} = \mathbf{x} - \mathbf{x}_1$ folgt nun:

$$\mathbf{f}^{-1}(\mathbf{y}) = \mathbf{f}^{-1}(\mathbf{y}_1) + (\mathbf{y} - \mathbf{y}_1) \cdot \Delta^*(\mathbf{y})^\top.$$

Das liefert die Differenzierbarkeit von $\mathbf{f}^{-1}$ in $\mathbf{y}_1$ und die Formel

$$D\mathbf{f}^{-1}(\mathbf{y}_1) = (D\mathbf{f}(\mathbf{x}_1))^{-1}.$$

Damit ist der Umkehrsatz bewiesen. ∎

5.11. Beispiele

A. Die Polarkoordinaten $(x, y) = \mathbf{f}(r, \varphi) = (r\cos\varphi, r\sin\varphi)$ haben wir schon an früherer Stelle betrachtet. Es ist

$$\det J_{\mathbf{f}}(r, \varphi) = \det \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} = r.$$

In jedem Punkt (r, φ) mit $r > 0$ und $\varphi \in \mathbb{R}$ ist $\mathbf{f}$ also lokal umkehrbar.

$$\mathbf{f} : \mathbb{R}_+ \times [0, 2\pi) \to \mathbb{R}^2 \setminus \{(0,0)\}$$

ist sogar global umkehrbar.

B. Sei $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch $\mathbf{f}(x, y) := (x^2 - y^2, 2xy)$. Dann gilt:

$$J_{\mathbf{f}}(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}, \text{ also } \det J_{\mathbf{f}}(x, y) = 4(x^2 + y^2).$$

Damit ist $\det J_{\mathbf{f}}(x, y) \neq 0$ für $(x, y) \neq (0, 0)$ und $\mathbf{f}$ überall außerhalb des Nullpunktes lokal umkehrbar.

$\mathbf{f}$ ist aber nicht global umkehrbar, denn es ist z.B. $\mathbf{f}(-x, -y) = \mathbf{f}(x, y)$.

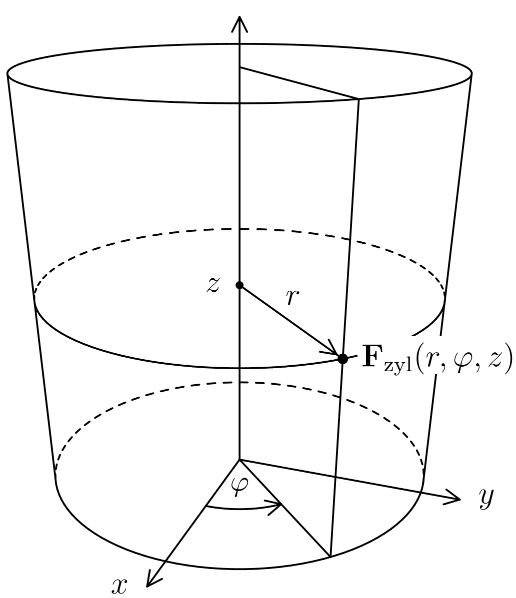
C. Zylinderkoordinaten:

Sei $G := \{(r, \varphi, z) \in \mathbb{R}^3 \,:\, r > 0,\ 0 < \varphi < 2\pi \text{ und } z \text{ beliebig}\}$ und

$$\boxed{\mathbf{F}_{\text{zyl}}(r, \varphi, z) := (r \cos \varphi, r \sin \varphi, z).}$$

Dann ist $J_{\mathbf{F}_{\text{zyl}}}(r, \varphi, z) = \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$ und $\det J_{\mathbf{F}_{\text{zyl}}}(r, \varphi, z) = r$.

Also ist $\mathbf{F}_{\text{zyl}}$ außerhalb der z-Achse ein lokaler Diffeomorphismus.

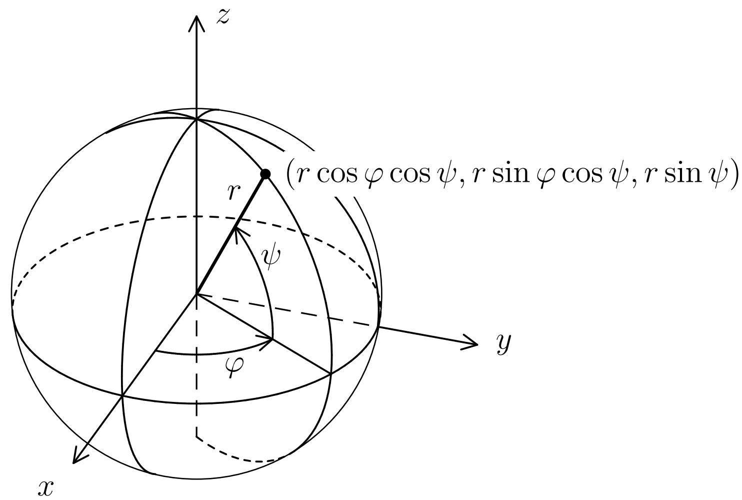


D. Räumliche Polarkoordinaten (Kugelkoordinaten):

Sei $G := \{(r, \varphi, \psi) \,:\, r > 0,\ 0 < \varphi < 2\pi \text{ und } -\frac{\pi}{2} < \psi < \frac{\pi}{2}\}$, und $\mathbf{F}_{\text{sph}} : G \to \mathbb{R}^3$ definiert durch

$$\boxed{\mathbf{F}_{\text{sph}}(r, \varphi, \psi) := (r \cos \varphi \cos \psi, r \sin \varphi \cos \psi, r \sin \psi).}$$

Dann ist r der Abstand vom Nullpunkt, φ der Winkel in der x-y-Ebene gegen die positive x-Achse (also die geographische Länge) und ψ der Winkel des Radiusvektors gegen die x-y-Ebene (also die geographische Breite).



Leider gibt es in der Literatur verschiedene Definitionen der Kugelkoordinaten. Ebenso gebräuchlich wie die obige Version ist die Abbildung

$$\widetilde{\mathbf{F}}_{\text{sph}}(r,\theta,\varphi) = (r\cos\varphi\sin\theta, r\sin\varphi\sin\theta, r\cos\theta)$$

mit $r > 0$, $0 < \theta < \pi$ und $0 < \varphi < 2\pi$. Dabei ist θ der Winkel gegen die positive z-Achse, also $\theta + \psi = \pi/2$, $\cos\theta = \cos(\pi/2 - \psi) = -\sin(-\psi) = \sin\psi$ und $\sin\theta = \sin(\pi/2 - \psi) = \cos(-\psi) = \cos\psi$. Die Reihenfolge der Koordinaten ist wichtig aus Gründen, die erst später erklärt werden können. Bei beiden Versionen ist die Funktionaldeterminante positiv. Man findet in der Literatur auch Variationen der Kugelkoordinaten, bei denen dies nicht der Fall ist, was relativ unsinnig ist.

Im Falle der Abbildung $\mathbf{F} = \mathbf{F}_{\text{sph}}$ ist $\theta_1 = \psi$ und

$$\begin{aligned}
\det J_{\mathbf{F}}(r,\varphi,\psi) &= \det \begin{pmatrix} \cos\varphi\cos\psi & -r\sin\varphi\cos\psi & -r\cos\varphi\sin\psi \\ \sin\varphi\cos\psi & r\cos\varphi\cos\psi & -r\sin\varphi\sin\psi \\ \sin\psi & 0 & r\cos\psi \end{pmatrix} \\
&= \sin\psi \cdot r^2(\sin^2\varphi\sin\psi\cos\psi + \cos^2\varphi\sin\psi\cos\psi) \\
&\quad + r\cos\psi \cdot r(\cos^2\varphi\cos^2\psi + \sin^2\varphi\cos^2\psi) \\
&= r^2\sin^2\psi\cos\psi + r^2\cos^3\psi \;=\; r^2\cos\psi.
\end{aligned}$$

Für $r > 0$ und $-\pi/2 < \psi < \pi/2$ ist tatsächlich $\det J_{\mathbf{F}}(r,\varphi,\psi) > 0$ und $\mathbf{F}_{\text{sph}}$ ein lokaler Diffeomorphismus.