

Kapitel 1

Differentialrechnung für Funktionen mehrerer Variabler

In diesem Kapitel behandeln wir die Differentialrechnung für Funktionen mehrerer Variabler. Des besseren Verständnisses wegen entwickeln wir alle Begriffe im \mathbb{R}^n , also mittels kartesischer Koordinaten.

Wir beginnen in Abschnitt 1.1 damit, die Begriffe *partielle Differenzierbarkeit*, *partielle Ableitung*, *Richtungsableitung*, *Gradient*, *Divergenz* und *Rotation* einzuführen und erste Eigenschaften partiell differenzierbarer Funktionen herzuleiten. In Abschnitt 1.2 wird der Begriff der *(totalen) Differenzierbarkeit* definiert und untersucht, inwieweit er mit der partiellen Differenzierbarkeit übereinstimmt; es zeigt sich, daß beide Begriffe auf der Klasse $C^1(\Omega, \mathbb{R}^N)$ der stetigen Funktionen $f : \Omega \rightarrow \mathbb{R}^N$ mit stetigen ersten partiellen Ableitungen übereinstimmen. Im Anschluß daran wird die *Tangentialebene* eines differenzierbaren Graphen definiert, und es werden die Begriffe *totales Differential*, *Differentialform ersten Grades (1-Form)* und *Kovektorfeld* eingeführt. Schließlich wird die allgemeine Form der Kettenregel angegeben.

In Abschnitt 1.3 wird die Differenzierbarkeit bestimmter Integrale nach einem Parameter untersucht. Das so gewonnene Ergebnis führt zu dem wichtigen Satz von H.A. Schwarz über die Vertauschbarkeit zweier partieller Ableitungen. Partielle Ableitungen höherer Ordnung werden in 1.5 studiert, und anschließend werden der Begriff des *Potentiales* eines Vektorfeldes definiert und notwendige sowie hinreichende Bedingungen für die Existenz eines Potentials aufgestellt.

In 1.6 wird die *Taylorsche Formel* für Funktionen mehrerer Variabler hergeleitet.

Ferner ist Abschnitt 1.7 der Untersuchung lokaler Extrema einer reellen Funktion $f : \Omega \rightarrow \mathbb{R}$ gewidmet, wobei die *Hessesche Matrix* $H_f = D^2f$ eine wesentliche Rolle spielt. Es folgen das Maximumprinzip für harmonische Funktionen und eine Untersuchung von Schwingungen um eine Gleichgewichtslage. Abschnitt 1.8 bietet eine kurze Einführung in die *Theorie der konvexen Mengen und konvexen Funktionen*. Einige wichtige Ungleichungen wie die von Hölder, Minkowski und Jensen beruhen auf diesem Konzept.

Abschnitt 1.9 behandelt den *Umkehrsatz*, mit dessen Hilfe man feststellen kann, ob eine C^1 -Abbildung $f : \Omega \rightarrow \mathbb{R}^n$, $\Omega \subset \mathbb{R}^n$, eine lokale oder globale Inverse der Klasse C^1 besitzt, ob f also ein lokaler oder globaler Diffeomorphismus ist. Als eine erste Anwendung des Umkehrsatzes wird in 1.10 die Legendretransformation untersucht, die in vielen Gebieten eine Rolle spielt, etwa in der konvexen Analysis, in der Variationsrechnung und in der Theorie der Differentialgleichungen, in der klassischen Mechanik (Hamilton-Jacobischer Formalismus), in der Thermodynamik und der Elastizitätstheorie.

In Abschnitt 1.11 beweisen wir den *Satz von Heine-Borel*, mit dessen Hilfe sich kompakte Mengen in ganz neuer Weise charakterisieren lassen. In allgemeinen topologischen Räumen ist dieser Satz der Ausgangspunkt zur Definition kompakter Mengen. Der Heine-Borelsche Satz besitzt vielfältige Anwendungen. Zur Illustration zeigen wir mit seiner Hilfe, wie man Lipschitzstetige Funktionen gewinnen kann, und ferner, daß kompakte Nullmengen im \mathbb{R}^n auch Mengen vom Inhalt Null sind. Dieses und ähnliche Resultate sind nützlich bei der Definition des mehrdimensionalen Riemannschen Integrales.

1 Partielle Ableitungen von Funktionen mehrerer Variabler

In diesem Abschnitt behandeln wir zuerst eine naheliegende Verallgemeinerung des Ableitungsbegriffs auf Funktionen mehrerer Variabler x_1, \dots, x_n .

Ist eine solche Funktion $f : \Omega \rightarrow \mathbb{R}$ gegeben, so denken wir uns für den Augenblick bloß eine der Variablen als unveränderlich, etwa x_j , und „frieren die übrigen Variablen ein“. Differenzieren wir f nach x_j , so entsteht die *partielle Ableitung* $D_j f(x)$, die wir auch mit $f_{x_j}(x)$ oder $\frac{\partial f}{\partial x_j}(x)$ bezeichnen.

Danach definieren wir die *Richtungsableitung* $\frac{\partial f}{\partial a}(x)$ einer Funktion $f : \Omega \rightarrow \mathbb{R}$ an der Stelle x in Richtung eines beliebigen Einheitsvektors a . Die partiellen Ableitungen $D_j f$ erweisen sich als die Richtungsableitungen von f in Richtung der Einheitsvektoren e_j in den Koordinatenrichtungen.

Mit $C^1(\Omega)$ bezeichnen wir die Klasse der Funktionen $f : \Omega \rightarrow \mathbb{R}$ auf einer offenen Menge des \mathbb{R}^n , für die in allen Punkten $x \in \Omega$ sämtliche partielle Ableitungen $D_1 f(x), \dots, D_n f(x)$ existieren und stetige Funktionen $D_j f : \Omega \rightarrow \mathbb{R}$ liefern. Setzen wir in eine solche Funktion f eine C^1 -Kurve $\varphi : I \rightarrow \Omega$ ein, so erhalten wir eine C^1 -Funktion $f \circ \varphi : I \rightarrow \mathbb{R}$, und es gilt die

Kettenregel II:

$$\frac{d}{dt} f(\varphi(t)) = \sum_{j=1}^n D_j f(\varphi_1(t), \dots, \varphi_n(t)) \dot{\varphi}_j(t).$$

Dies ist die Essenz der Kettenregel für Funktionen mehrerer Variabler, aus der wir sofort den *Mittelwertsatz* gewinnen. Dieser liefert insbesondere die Inklusion $C^1(\Omega) \subset C^0(\Omega)$. Mit Hilfe des *Gradientenvektors* $\text{grad } f$ oder ∇f einer C^1 -Funktion $f: \Omega \rightarrow \mathbb{R}$, der durch $\nabla f = (D_1 f, \dots, D_n f)$ definiert ist, läßt sich die Richtungsableitung $\frac{\partial f}{\partial a}(x)$ als Skalarprodukt von $\nabla f(x)$ mit dem Richtungsvektor $a \in S^{n-1} \subset \mathbb{R}^n$ ausdrücken, also

$$\frac{\partial f}{\partial a}(x) = \langle \nabla f(x), a \rangle.$$

Hieraus folgt sofort die geometrische Deutung des Gradientenvektors $\nabla f(x)$ an der Stelle $x \in \Omega$. Er weist in die *Richtung des stärksten Anstiegs* von f , und seine Länge liefert die Größe des stärksten Anstiegs. In den Extrempunkten von $f: \Omega \rightarrow \mathbb{R}$ verschwindet ∇f , diese sind also *kritische Punkte* von f .

Bekanntlich sind die Nullstellen x_0 eines Vektorfeldes $F: \Omega \rightarrow \mathbb{R}^n$ auf einer offenen Menge $\Omega \subset \mathbb{R}^n$ die *Gleichgewichtspunkte* für die Differentialgleichung $\dot{X}(t) = F(X(t))$. Geht nämlich X durch einen solchen Punkt x_0 , so gilt $X(t) \equiv x_0$, falls F Lipschitzstetig ist. Entfernt von Gleichgewichtspunkten ist eine Strömung $X(t, x)$ nahezu eine Parallelströmung, während sie in der Nähe eines solchen Punktes, wie wir wissen, bereits für $n = 2$ äußerst unterschiedliche Gestalt haben kann. Dies gilt insbesondere für die Nullstellen von $F := \nabla f$ und erklärt die Bezeichnung „kritische Punkte“ für diese Stellen.

Eine offene Menge Ω des \mathbb{R}^n heißt *bogenweise zusammenhängend* oder *Gebiet*, wenn zwei beliebige Punkte aus Ω durch einen in Ω verlaufenden stetigen Bogen verbunden werden können. Eine C^1 -Funktion $u: \Omega \rightarrow \mathbb{R}$ auf einem Gebiet Ω ist genau dann konstant, wenn ihr Gradient ∇u auf Ω identisch verschwindet.

Der Begriff der partiellen Ableitung $D_j f$ überträgt sich ohne weiteres auf Abbildungen $f: \Omega \rightarrow \mathbb{R}^N$ mit den Komponenten f_1, \dots, f_N , die wir gewöhnlich in der Form $f(x) = (f_1(x), \dots, f_N(x))$ schreiben. Es ist $D_j f = (D_j f_1, \dots, D_j f_N)$. Der Mittelwertsatz gilt für jede Komponente f_k , läßt sich aber nicht auf f übertragen. Vielmehr gilt eine abgeschwächte Fassung in integrierter Form, die als *Hadamards Lemma* bekannt und in vielerlei Weise nützlich ist.

Für Abbildungen $f: \Omega \rightarrow \mathbb{R}^N$ definieren wir die *Jacobimatrix* $Df(x)$ als die Matrix $Df(x) = (D_j f_k(x))$ mit den Matricelementen $D_j f_k(x)$ und beweisen die *allgemeine Kettenregel*.

Besonders wichtig sind die für $n = N$ definierten *Diffeomorphismen* $f: \Omega \rightarrow \Omega^*$ von Ω auf $\Omega^* := f(\Omega)$. Dies sind Homöomorphismen der Klasse C^1 , deren Inverse $f^{-1}: \Omega^* \rightarrow \Omega$ ebenfalls von der Klasse C^1 sind. Für Diffeomorphismen f gilt

$$(Df)^{-1} = (Df^{-1}) \circ f$$

woraus für die *Jacobideterminante* $J_f := \det Df$ eines Diffeomorphismus f die Regel $J_{f^{-1}} \circ f = 1/J_f$ folgt. Für C^1 -Funktionen $f: \Omega \rightarrow \mathbb{R}^n$, $\Omega \subset \mathbb{R}^n$, definieren wir insbesondere die *Divergenz* $\text{div } f: \Omega \rightarrow \mathbb{R}$ durch

$$\text{div } f := D_1 f_1 + D_2 f_2 + \dots + D_n f_n,$$

und für den Spezialfall $n = 3$ den Vektor $\text{rot } f$, die *Rotation* von f . Für die wichtigen Differentialausdrücke $\text{grad } f$, $\text{div } f$ und $\text{rot } f$ beweisen wir verschiedene Rechenregeln. Weiterhin behandeln wir *positiv homogene Funktionen* $f: \Omega \rightarrow \mathbb{R}$ vom Grade α . C^1 -Funktionen dieser Art sind durch die *Eulersche Relation*

$$\alpha f(x) = \langle x, \nabla f(x) \rangle$$

charakterisiert.

Wenden wir uns nun den Einzelheiten zu. Im folgenden bezeichne Ω stets eine (nichtleere) offene Menge des \mathbb{R}^n . Wir betrachten Abbildungen

$$f : \Omega \rightarrow \mathbb{R}^N, f(x) = (f_1(x), \dots, f_N(x)), \quad x = (x_1, \dots, x_n).$$

Für $x \in \Omega$ und $1 \leq j \leq n$ bilden wir die Funktion

$$(1) \quad F_j(t) := f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_n),$$

die für $|t - x_j| \ll 1$ wohldefiniert ist. Mit anderen Worten: wir frieren alle Variablen von f ein bis auf die j -te Variable, die in t umbenannt wird. Die so entstehende Funktion $F_j(t)$ ist dann für t -Werte in der Nähe von x_j definiert.

Definition 1. Eine Funktion $f : \Omega \rightarrow \mathbb{R}^N$ heißt an der Stelle $x \in \Omega$ **partiell differenzierbar nach der j -ten Variablen**, wenn $F_j(t)$ an der Stelle $t = x_j$ differenzierbar ist. Man bezeichnet

$$(2) \quad D_j f(x) := \dot{F}_j(x_j)$$

als **partielle Ableitung von f nach der j -ten Variablen an der Stelle x** .

Noch Euler hat $\frac{d}{dx_j} f$ für die partielle Ableitung von f nach x_j geschrieben; seit Jacobi schreibt man stattdessen $\frac{\partial f}{\partial x_j}$ oder f_{x_j} , während Cauchy die Schreibweise $D_j f$ eingeführt hat.

Bezeichne e_1, \dots, e_n die kanonische Basis von \mathbb{R}^n , und sei

$$(3) \quad \Delta_h f(x) := \frac{1}{h} [f(x + h e_j) - f(x)]$$

der **j -te Differenzenquotient** von f an der Stelle x zur Schrittweite $h \in \mathbb{R}$ mit $0 < |h| \ll 1$, also

$$\Delta_h f(x) = \frac{1}{h} [f(x_1, \dots, x_j + h, \dots, x_n) - f(x_1, \dots, x_j, \dots, x_n)].$$

Falls f an der Stelle x partiell nach x_j differenzierbar ist, gilt

$$(4) \quad D_j f(x) = \lim_{h \rightarrow 0} \Delta_h f(x),$$

und umgekehrt folgt aus der Existenz von $\lim_{h \rightarrow 0} \Delta_h f(x)$, daß f an der Stelle x partiell nach x_j differenzierbar ist und die partielle Ableitung $D_j f(x)$ durch (4) gegeben wird.

Bemerkung 1. Eine Abbildung $f = (f_1, \dots, f_N)$ ist genau dann an der Stelle x nach x_j partiell differenzierbar, wenn ihre Komponenten f_1, \dots, f_N dies sind, und es gilt $D_j f = (D_j f_1, \dots, D_j f_N)$.

Definition 2. Eine Abbildung $f : \Omega \rightarrow \mathbb{R}^N$ heißt **von der Klasse C^1** , wenn die partiellen Ableitungen $D_1 f(x), \dots, D_n f(x)$ in allen Punkten $x \in \Omega$ existieren und stetig von x abhängen; wir schreiben hierfür: $f \in C^1(\Omega, \mathbb{R}^N)$.

Wegen (2) übertragen sich die Differentiationsregeln von 3.1 aus Band 1 für d/dt sofort auf die Operatoren D_1, \dots, D_n . Insbesondere sehen wir, daß $C^1(\Omega, \mathbb{R}^N)$ ein linearer Raum über \mathbb{R} ist.

Für $N = 1$ schreiben wir $C^1(\Omega) := C^1(\Omega, \mathbb{R})$, und jeder skalaren Funktion $f \in C^1(\Omega)$ ordnen wir ein Vektorfeld $\text{grad } f : \Omega \rightarrow \mathbb{R}^n$ zu, das **Gradientenfeld** von f ; es ist durch

$$(5) \quad \text{grad } f := (D_1 f, D_2 f, \dots, D_n f)$$

definiert. Dem irischen Mathematiker W.R. Hamilton folgend führen wir den „**Nabla-Vektor**“

$$(6) \quad \nabla = (D_1, D_2, \dots, D_n)$$

mit den „Komponenten“ D_1, D_2, \dots, D_n ein. Dann ist

$$(7) \quad \text{grad } f = \nabla f,$$

und man interpretiert ∇f formal als „Produkt“ von ∇ mit dem Skalar f . Die geometrische Bedeutung des Gradientenfeldes werden wir in Kürze erkennen.

Ist $f : \Omega \rightarrow \mathbb{R}^N$ nach allen Variablen an der Stelle x partiell differenzierbar, so können wir die **Jacobimatrix** $Df(x)$ von f definieren:

$$(8) \quad Df(x) = \begin{pmatrix} D_1 f_1(x) & \dots & D_n f_1(x) \\ \vdots & & \vdots \\ D_1 f_N(x) & \dots & D_n f_N(x) \end{pmatrix} = \begin{pmatrix} \text{grad } f_1 \\ \vdots \\ \text{grad } f_N \end{pmatrix}.$$

Statt Df schreiben wir auch $\frac{\partial f}{\partial x}$ oder f_x .

Falls $n = N$ ist, können wir auch die **Jacobideterminante** oder **Funktionaldeterminante** bilden:

$$(9) \quad J_f := \det Df = \begin{vmatrix} D_1 f_1 & \dots & D_n f_1 \\ \vdots & & \vdots \\ D_1 f_n & \dots & D_n f_n \end{vmatrix}.$$

1 Die *Abstandsfunktion* $r : \mathbb{R}^n \rightarrow \mathbb{R}$, $r(x) := |x| = \sqrt{x_1^2 + \dots + x_n^2}$ hat für $x \neq 0$ die partiellen Ableitungen

$$\frac{\partial r}{\partial x_j}(x) = \frac{x_j}{|x|},$$

d.h. r ist von der Klasse C^1 auf $\Omega \setminus \{0\}$. Das Vektorfeld $\nu(x) = |x|^{-1}x = \text{grad } r(x)$ ist ein „radiales Feld“ von Einheitsvektoren.

2 Transformation auf Polarkoordinaten um den Ursprung. Die kartesischen Koordinaten x, y und die Polarkoordinaten r, φ um den Ursprung sind durch die Transformationsformeln $x = r \cos \varphi$, $y = r \sin \varphi$ verbunden. Die Jacobimatrix dieser Transformation ist

$$\begin{pmatrix} x_r & x_\varphi \\ y_r & y_\varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix},$$

und die zugehörige Jacobideterminante

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

Satz 1. (Kettenregel II). Sei $f \in C^1(\Omega)$, und bezeichne $\varphi : I \rightarrow \mathbb{R}^n$ eine im Punkte $t \in I$ differenzierbare Kurve mit $\varphi(I) \subset \Omega$. Dann ist die Funktion $f \circ \varphi : I \rightarrow \mathbb{R}^n$ an der Stelle t differenzierbar und es gilt

$$(10) \quad \frac{d}{dt} f(\varphi(t)) = \sum_{j=1}^n f_{x_j}(\varphi(t)) \dot{\varphi}_j(t),$$

d.h.

$$(11) \quad \frac{d}{dt} f(\varphi(t)) = \nabla f(\varphi(t)) \cdot \dot{\varphi}(t) = \langle \nabla f(\varphi(t)), \dot{\varphi}(t) \rangle.$$

Beweis. Sei $x := \varphi(t)$. Dann gibt es ein $r > 0$, so daß der abgeschlossene Würfel $W_r(x) = \{\xi \in \mathbb{R}^n : |\xi - x|_* \leq r\}$ in Ω liegt. Ferner existiert ein $\delta > 0$, so daß

$$\varphi(t+h) \in W_r(x) \text{ für jedes } h \in [-\delta, \delta]$$

gilt. Wir fixieren ein $h \in \mathbb{R}$ mit $|h| < \delta$ und setzen $\xi := \varphi(t+h)$, $\Delta x := \xi - x$, also $\xi = x + \Delta x$. Sei $x = (x_1, \dots, x_n)$, $\Delta x = (\Delta x_1, \dots, \Delta x_n)$. Dann liegen die Punkte

$$P_0 := \xi = x + \Delta x, \quad P_1 := (x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n),$$

$$P_2 := (x_1, x_2, x_3 + \Delta x_3, \dots, x_n + \Delta x_n), \dots, P_n := x$$

und auch die Verbindungsstrecken $[P_{j-1}, P_j] = \{\lambda P_{j-1} + (1-\lambda)P_j : 0 \leq \lambda \leq 1\}$ für $1 \leq j \leq n$ im Würfel $W_r(x)$. Es folgt

$$\begin{aligned} f(\xi) - f(x) &= f(P_0) - f(P_n) \\ &= [f(P_0) - f(P_1)] + [f(P_1) - f(P_2)] + \dots + [f(P_{n-1}) - f(P_n)]. \end{aligned}$$

Auf jede der eckigen Klammern kann man den gewöhnlichen Mittelwertsatz anwenden und erhält so $f(P_{j-1}) - f(P_j) = D_j f(Q_j) \Delta x_j$ mit einem geeigneten Punkte $Q_j \in [P_{j-1}, P_j]$. Hieraus folgt

$$f(\xi) - f(x) = \sum_{j=1}^n D_j f(Q_j) \Delta x_j$$

und daher

$$(12) \quad \frac{1}{h} [f(\varphi(t+h)) - f(\varphi(t))] = \sum_{j=1}^n D_j f(Q_j) \frac{1}{h} [\varphi_j(t+h) - \varphi_j(t)] .$$

Mit $h \rightarrow 0$ streben die Punkte P_j gegen x , und damit streben auch die Q_j gegen x . Folglich ergibt sich aus (12) die Behauptung. \square

Ist beispielsweise f eine Funktion von $(x, y, z) \in \Omega \subset \mathbb{R}^3$ und bezeichnet $t \mapsto (\varphi(t), \psi(t), \chi(t))$ eine C^1 -Kurve in Ω , so gilt

$$\begin{aligned} \frac{d}{dt} f(\varphi(t), \psi(t), \chi(t)) \\ = f_x(\varphi(t), \psi(t), \chi(t))\dot{\varphi}(t) + f_y(\varphi(t), \psi(t), \chi(t))\dot{\psi}(t) + f_z(\varphi(t), \psi(t), \chi(t))\dot{\chi}(t) . \end{aligned}$$

Satz 2. (Mittelwertsatz). Sei $f \in C^1(\Omega)$ und seien x, y Punkte aus Ω , deren Verbindungsstrecke $[x, y]$ in Ω liegt. Dann gibt es ein $\vartheta \in (0, 1)$ mit

$$(13) \quad f(y) - f(x) = \nabla f(x + \vartheta(y-x)) \cdot (y-x) .$$

Beweis. Für $0 \leq t \leq 1$ setzen wir $g(t) := f(x + t(y-x))$.

Nach der Kettenregel II ist $g : [0, 1] \rightarrow \mathbb{R}$ von der Klasse C^1 und es gilt

$$\dot{g}(t) = \nabla f(x + t(y-x)) \cdot (y-x) .$$

Der (eindimensionale) Mittelwertsatz liefert

$$f(y) - f(x) = g(1) - g(0) = \dot{g}(\vartheta) \text{ für ein } \vartheta \in (0, 1) ,$$

woraus sich (13) ergibt. \square

Korollar 1. Es gilt $C^1(\Omega, \mathbb{R}^N) \subset C^0(\Omega, \mathbb{R}^N)$.

Beweis. Ist $f \in C^1(\Omega, \mathbb{R}^N)$, so folgt $f_j \in C^1(\Omega)$ für jede Komponente von f , und Satz 2 liefert $f_j \in C^0(\Omega)$, also $f \in C^0(\Omega, \mathbb{R}^N)$. \square

Definition 3. Seien $f : \Omega \rightarrow \mathbb{R}$, $x \in \Omega$ und $a \in \mathbb{R}^n$ mit $|a| = 1$ gegeben. Dann nennt man den folgenden Grenzwert – falls er existiert – die **Richtungsableitung von f an der Stelle x in Richtung von a :**

$$(14) \quad \frac{\partial f}{\partial a}(x) := \lim_{t \rightarrow 0} \frac{1}{t} [f(x + ta) - f(x)] .$$

Bemerkung 2. Partielle Ableitungen sind spezielle Richtungsableitungen. Es gilt nämlich, wenn e_1, \dots, e_n die kanonische Basis von \mathbb{R}^n bezeichnet,

$$\frac{\partial f}{\partial x_j}(x) = \frac{\partial f}{\partial e_j}(x) .$$

Satz 3. Für $f \in C^1(\Omega)$ existiert $\frac{\partial f}{\partial a}(x)$ für jedes $x \in \Omega$ und jedes $a \in \mathbb{R}^n$ mit $|a| = 1$, und es gilt

$$(15) \quad \frac{\partial f}{\partial a}(x) = \langle \nabla f(x), a \rangle = \nabla f(x) \cdot a ,$$

d.h.

$$(16) \quad \frac{\partial f}{\partial a}(x) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) a_j .$$

Beweis. Wir setzen $\varphi(t) := x + ta$. Dann gilt $\varphi(t) \in \Omega$ für $|t| \ll 1$, und nach der Kettenregel II folgt

$$\frac{d}{dt}f(x + ta) = \nabla f(x + ta) \cdot a$$

und insbesondere

$$\frac{d}{dt}f(x + ta)|_{t=0} = \nabla f(x) \cdot a .$$

Andererseits ist

$$\frac{d}{dt}f(x + ta)|_{t=0} = \lim_{t \rightarrow 0} \frac{1}{t} [f(x + ta) - f(x)] .$$

Damit ergibt sich die Formel (15). □

Korollar 2. Für $f \in C^1(\Omega)$ und $x \in \Omega$ liefert $\text{grad } f(x)$ die Richtung des stärksten Anstieges von f an der Stelle x , und $-\text{grad } f(x)$ hat die Richtung stärksten Gefälles. Wenn $\nabla f(x) \neq 0$ ist, sind diese beiden Richtungen eindeutig bestimmt.

Beweis. Aus (15) folgt wegen $|a| = 1$ aus der Schwarzschen Ungleichung

$$\left| \frac{\partial f}{\partial a}(x) \right| \leq |\nabla f(x)| ,$$

wobei das Gleichheitszeichen genau dann vorliegt, wenn a und $\nabla f(x)$ linear abhängig sind. Ist $\nabla f(x) \neq 0$ und bezeichnet ν den Einheitsvektor $\nu := |\nabla f(x)|^{-1} \nabla f(x)$, so gilt

$$\frac{\partial f}{\partial \nu}(x) = |\nabla f(x)| , \quad \frac{\partial f}{\partial (-\nu)}(x) = -|\nabla f(x)|$$

und

$$-|\nabla f(x)| < \frac{\partial f}{\partial a}(x) < |\nabla f(x)| \quad \text{für } a \neq \pm \nu.$$

Weiterhin haben wir $\frac{\partial f}{\partial a}(x) = 0$ für alle $a \in S^{n-1}$, falls $\nabla f(x) = 0$ ist. □

Bemerkung 3. Das Korollar 2 erklärt, warum das Vektorfeld $x \mapsto \nabla f(x)$ als das *Gradientenfeld* der Funktion $f \in C^1(\Omega)$ bezeichnet wird.

Definition 4. Die Nullstellen des Gradientenfeldes ∇f einer Funktion $f \in C^1(\Omega)$ werden als **kritische Punkte** von f bezeichnet.

Definition 5. Wir sagen, $f : \Omega \rightarrow \mathbb{R}$ habe im Punkte $x_0 \in \Omega$ ein **lokales Minimum** (bzw. ein **lokales Maximum**), wenn es eine Kugel $B_r(x_0)$ in Ω gibt, so daß

$$f(x_0) \leq f(x) \quad (\text{bzw. } f(x_0) \geq f(x))$$

für alle $x \in B_r(x_0)$ gilt, und wir bezeichnen x_0 als einen **lokalen Minimierer** (**Maximierer**) von f .

Wie in Band 1 bezeichnen wir lokale Minima und Maxima als (lokale) **Extrema**.

Satz 4. Besitzt $f \in C^1(\Omega)$ im Punkte $x \in \Omega$ ein lokales Extremum, so gilt

$$\nabla f(x) = 0,$$

d.h. x ist ein kritischer Punkt von f .

Beweis. Für $1 \leq j \leq n$ hat die Funktion

$$F_j(t) := f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_n)$$

an der Stelle $t = x_j$ ein lokales Extremum. Nach 3.2 von Band 1 gilt $\dot{F}_j(x_j) = 0$, und dies bedeutet $D_j f(x) = 0$ für $1 \leq j \leq n$, d.h. $\nabla f(x_0) = 0$. □

Definition 6. (i) Eine Menge M des \mathbb{R}^n heißt (**bogenweise**) **zusammenhängend**, wenn es zu jedem Paar $x, y \in M$ eine Kurve $\varphi \in C^1(I, \mathbb{R}^n)$ mit $I = [0, 1]$ und $\varphi(I) \subset M$ gibt, so daß $\varphi(0) = x$ und $\varphi(1) = y$ gilt.

(ii) Eine zusammenhängende offene Menge Ω des \mathbb{R}^n wird **Gebiet** in \mathbb{R}^n genannt. Gebiete wollen wir gewöhnlich mit G bezeichnen.

Satz 5. Sei G ein Gebiet in \mathbb{R}^n , $f \in C^1(G, \mathbb{R}^N)$, und es gelte $Df(x) \equiv 0$ auf G . Dann folgt $f(x) \equiv \text{const}$ auf G .

Beweis. (i) Wir betrachten zunächst den Fall $N = 1$. Wir fixieren einen Punkt $x_0 \in G$ und setzen $c := f(x_0)$. Dann gibt es zu jedem $x \in G$ eine Kurve $\varphi \in C^1(I, \mathbb{R}^N)$ mit $I = [0, 1]$ und $\varphi(I) \subset G$, so daß $\varphi(0) = x_0$ und $\varphi(1) = x$ ist. Es folgt

$$f(x) - f(x_0) = f(\varphi(1)) - f(\varphi(0)) = \int_0^1 \frac{d}{dt} f(\varphi(t)) dt = \int_0^1 \nabla f(\varphi(t)) \cdot \dot{\varphi}(t) dt ,$$

und wegen $\nabla f(x) \equiv 0$ verschwindet das letzte Integral. Somit ergibt sich

$$f(x) = c \text{ für jedes } x \in G .$$

(ii) Ist $N > 1$ und $f = (f_1, \dots, f_N)$, so gibt es wegen (i) Konstanten c_1, \dots, c_N , so daß $f_1(x) \equiv c_1, \dots, f_N(x) \equiv c_N$ auf G ist. Mit $c = (c_1, \dots, c_N)$ folgt dann $f(x) \equiv c$ auf G . □

Die soeben benutzte Schlußweise formulieren wir als

Hadamards Lemma. *Ist $f \in C^1(\Omega, \mathbb{R}^N)$ und sind x, y zwei Punkte aus Ω , deren Verbindungsstrecke $[x, y] := \{x + t(y - x) : 0 \leq t \leq 1\}$ in Ω liegt, so gilt*

$$(17) \quad f(y) - f(x) = A \cdot (y - x) ,$$

wobei $f(y) - f(x)$ sowie $y - x$ als Spaltenvektoren aufzufassen sind und A die $N \times n$ -Matrix

$$(18) \quad A := \int_0^1 Df(x + t(y - x)) dt$$

bezeichnet.

Beweis. Wir bilden die C^1 -Kurve $\varphi(t) := x + t(y - x)$, $0 \leq t \leq 1$. Sie erfüllt $\varphi(0) = x$, $\varphi(1) = y$ und $\varphi(t) \in \Omega$ für $0 \leq t \leq 1$. Ist nun $f = (f_1, \dots, f_N)$, so folgt mit $g(t) := f_j(\varphi(t)) = f_j(x + t(y - x))$ die Relation

$$f_j(y) - f_j(x) = g(1) - g(0) = \int_0^1 \dot{g}(t) dt = \int_0^1 \nabla f_j(x + t(y - x)) dt \cdot (y - x) ,$$

und dies ist äquivalent zu (17) & (18). □

Wir können Hadamards Lemma als eine verallgemeinerte Fassung des Mittelwertsatzes ansehen, die für vielerlei Zwecke nützlich ist.

Satz 6. (Kettenregel III). *Sind $f \in C^1(\Omega, \mathbb{R}^N)$, $\varphi \in C^1(\Omega^*, \mathbb{R}^n)$, wobei Ω und Ω^* offene Mengen des \mathbb{R}^n bzw. \mathbb{R}^m bezeichnen, und gilt $\varphi(\Omega^*) \subset \Omega$, so folgt $g := f \circ \varphi \in C^1(\Omega^*, \mathbb{R}^N)$, und Dg berechnet sich aus*

$$(19) \quad Dg(z) = Df(\varphi(z)) \cdot D\varphi(z) \quad \text{für } z \in \Omega^* .$$

Beweis. Sei $f = (f_1, \dots, f_N)$ und $g = (g_1, \dots, g_N)$. Dann gilt $g_j = f_j \circ \varphi$, und nach Satz 1 folgt

$$\frac{\partial}{\partial z_k} g_j(z) = \nabla f_j(\varphi(z)) \cdot \frac{\partial}{\partial z_k} \varphi(z).$$

Fixieren wir k und fassen $\frac{\partial}{\partial z_k} g$ und $\frac{\partial}{\partial z_k} \varphi$ als Spaltenvektoren auf, so erhält diese Formel die Gestalt

$$\frac{\partial g}{\partial z_k}(z) = Df(\varphi(z)) \cdot \frac{\partial \varphi}{\partial z_k}(z).$$

Schreiben wir nun Dg und $D\varphi$ mittels ihrer Spalten als

$$Dg = \left(\frac{\partial g}{\partial z_1}, \dots, \frac{\partial g}{\partial z_m} \right), \quad D\varphi = \left(\frac{\partial \varphi}{\partial z_1}, \dots, \frac{\partial \varphi}{\partial z_m} \right),$$

so ergibt sich (19). □

Bemerkung 4. Als mnemotechnische Hilfe ist es gelegentlich nützlich, an D als Index diejenigen Variablen anzuhängen, nach denen abgeleitet wird. Beispielsweise würde (19) in dieser Schreibweise lauten:

$$(20) \quad D_z g(z) = D_x f(\varphi(z)) \cdot D_z \varphi(z).$$

Auf diese Weise erinnert man sich daran, daß $f(x)$ eine Funktion der Variablen x ist. Dann bedeutet $D_x f(\varphi(z))$, daß man zuerst die Jacobimatrix $(D_x f)(x)$ von f bilden soll und anschließend x durch $\varphi(z)$ ersetzt. Mit anderen Worten: $D_x f(\varphi(z))$ steht für $((Df) \circ \varphi)(z) = (Df)(\varphi(z))$. Diese Bezeichnungen und die Formel

$$(D(f \circ \varphi))(z) = ((Df) \circ \varphi)(z) \cdot (D\varphi)(z)$$

statt (19) sind unmißverständlich, aber schwer zu lesen und schlecht zu behalten. Besser ist es, sich (19) und (20) samt der Interpretation von $D_x f(\varphi(z))$ als $(D_x f)(\varphi(z))$ zu merken.

Noch einfacher ist es, sich auf die bewährte *Leibnizsche Symbolik* zu verlassen. Man denkt sich drei Typen von Variablen, y, x, z , die untereinander verbunden sind durch $y = y(x)$, $x = x(z)$, und damit $y = y(x(z)) = y(z)$. Diese schlampige, aber „intuitive“ Schreibweise bedeutet

$$y = f(x), \quad x = \varphi(z) \quad \text{und damit} \quad y = f(\varphi(z)).$$

Nunmehr schreibt man $\frac{\partial}{\partial x} = D_x$, $\frac{\partial}{\partial z} = D_z$ und benutzt „formal“ die Regeln der „Bruchrechnung“:

$$(21) \quad \frac{\partial y}{\partial z} = \frac{\partial y}{\partial x} \cdot \frac{\partial x}{\partial z}.$$

Anschließend wird (21) als

$$(22) \quad \frac{\partial y_j}{\partial z_k} = \sum_{l=1}^n \frac{\partial y_j}{\partial x_l} \frac{\partial x_l}{\partial z_k}$$

mit $1 \leq j \leq N$, $1 \leq k \leq m$ interpretiert und schließlich noch x in $\frac{\partial y_j}{\partial x_l}(x)$ durch $\varphi(z)$ ersetzt.

Noch besser kann man sich (22) merken, wenn man das Summenzeichen wegläßt und anstelle von (22) einfach

$$(23) \quad \frac{\partial y_j}{\partial z_k} = \frac{\partial y_j}{\partial x_l} \frac{\partial x_l}{\partial z_k}$$

schreibt, wobei vereinbart wird, daß über den doppelt auftretenden Index l von 1 bis n zu summieren ist. Dies ist die *Einsteinsche Summenkonvention*, die beispielsweise in der Differentialgeometrie und in der allgemeinen Relativitätstheorie häufig benutzt wird. Auch in (23) funktioniert die Leibnizsche Symbolik auf das Beste.

[3] Betrachten wir eine reellwertige Funktion $u(x, y)$ in der x, y -Ebene, die durch $x = r \cos \varphi$, $y = r \sin \varphi$ auf Polarkoordinaten r, φ um den Ursprung transformiert werden soll. Dann bilden wir $v(r, \varphi) := u(x(r, \varphi), y(r, \varphi))$ und erhalten nach (23)

$$v_r = u_x x_r + u_y y_r, \quad v_\varphi = u_x x_\varphi + u_y y_\varphi.$$

Diese Formeln liefern also für $v(r, \varphi) := u(r \cos \varphi, r \sin \varphi)$ die Ableitungen

$$\begin{aligned} v_r(r, \varphi) &= u_x(r \cos \varphi, r \sin \varphi) \cos \varphi + u_y(r \cos \varphi, r \sin \varphi) \sin \varphi, \\ v_\varphi(r, \varphi) &= -u_x(r \cos \varphi, r \sin \varphi) r \sin \varphi + u_y(r \cos \varphi, r \sin \varphi) r \cos \varphi, \end{aligned}$$

was wir nach (21) auch als Matrixgleichung

$$(v_r, v_\varphi) = (u_x, u_y) \cdot \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

schreiben können.

Korollar 3. *Ist $m = n = N$, so folgt unter den Voraussetzungen von Satz 6*

$$(24) \quad J_g(z) = J_f(\varphi(z)) \cdot J_\varphi(z) \quad \text{für } z \in \Omega^*.$$

Beweis. Aus Formel (19) ergibt sich

$$\det Dg(z) = \det[Df(\varphi(z)) \cdot D\varphi(z)] = [\det Df(\varphi(z))] \cdot [\det D\varphi(z)].$$

□

Bemerkung 5. Auch hier bewährt sich die Leibnizsche Symbolik, denn aus (21) folgt

$$\det \frac{\partial y}{\partial z} = \det \frac{\partial y}{\partial x} \cdot \det \frac{\partial x}{\partial z}.$$

Definition 7. *Sind Ω und Ω^* nichtleere offene Mengen des \mathbb{R}^n , so nennt man eine Abbildung $f : \Omega \rightarrow \mathbb{R}^n$ einen **Diffeomorphismus (der Klasse C^1) von Ω auf Ω^*** , wenn folgendes gilt:*

- (i) f bildet Ω bijektiv auf Ω^* ab;
- (ii) $f \in C^1(\Omega, \mathbb{R}^n)$ und $f^{-1} \in C^1(\Omega^*, \mathbb{R}^n)$.

Korollar 4. *Ist f ein C^1 -Diffeomorphismus von Ω auf Ω^* mit der Inversen $g = f^{-1}$, so gelten für $y = f(x)$ mit $x \in \Omega$ (bzw. für $x = g(y)$ mit $y \in \Omega^*$) die Beziehungen*

$$(25) \quad Dg(y) = [Df(x)]^{-1}$$

und

$$(26) \quad J_g(y) = 1/J_f(x).$$

Beweis. Aus $x = g(f(x))$ für alle $x \in \Omega$ folgt nach der Kettenregel

$$E = Dg(y) \cdot Df(x) ,$$

wobei E die Einheitsmatrix in $M(n)$ bezeichnet. Korollar 3 liefert nunmehr

$$1 = J_g(y) \cdot J_f(x) ,$$

d.h. $J_g(y) \neq 0$ und $J_f(x) \neq 0$. Somit sind $Df(x)$ und $Dg(y)$ invertierbar, und wir erhalten (25) und (26). □

Ein Monom $f(x)$ der Form

$$f(x) = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \quad , \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n ,$$

hat die Eigenschaft, daß sein Wert sich mit t^k multipliziert, $k := i_1 + i_2 + \dots + i_n$, wenn man das Argument x durch das Vielfache tx , $t \in \mathbb{R}$, ersetzt; es gilt also

$$(27) \quad f(tx) = t^k f(x) .$$

Die gleiche Beziehung gilt, wenn man für $f(x)$ eine Form k -ten Grades, also eine Linearkombination von Monomen k -ten Grades wählt:

$$f(x) = \sum_{i_1 + \dots + i_n = k} a_{i_1 i_2 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} , \quad a_{i_1 i_2 \dots i_n} \in \mathbb{R} .$$

Um dies zu verallgemeinern, führen wir die *positiv homogenen* Funktionen ein.

Definition 8. (i) Eine offene Menge Ω des \mathbb{R}^n mit $0 \notin \Omega$ heißt Strahlenmenge, wenn mit $x \in \Omega$ auch der ganze Strahl $\{tx : 0 < t < \infty\}$ in Ω liegt.

(ii) Eine auf einer Strahlenmenge Ω definierte Funktion $f : \Omega \rightarrow \mathbb{R}$ heißt **positiv homogen vom Grad** $\alpha \in \mathbb{R}$, wenn für jedes $x \in \Omega$ gilt:

$$(28) \quad f(tx) = t^\alpha f(x) \quad \text{für alle } t > 0 .$$

[4] Die „typische“ Strahlenmenge Ω , die wir gewöhnlich im Sinn haben, ist $\Omega = \mathbb{R}^n \setminus \{0\}$. Eine positiv homogene Funktion vom Grade $-\alpha$ mit $\alpha > 0$ ist

$$f(x) := |x|^{-\alpha} .$$

[5] Der Cosinus des Winkels zwischen zwei Vektoren $x, \xi \in \mathbb{R}^n \setminus \{0\}$

$$\cos \sphericalangle(x, \xi) := \frac{\langle x, \xi \rangle}{|x| |\xi|} , \quad x \neq 0 , \quad \xi \neq 0 ,$$

ist eine positiv homogene Funktion nullten Grades auf

$$\Omega := \{(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n : x \neq 0 \text{ und } \xi \neq 0\} .$$

☐ Auf $\Omega = \{(x, y) \in \mathbb{R}^2, y \neq 0\}$ definiert $f(x, y) := x^2 \sin \frac{x}{y} + y\sqrt{x^2 + y^2}$ eine positiv homogene Funktion zweiten Grades.

Wie Euler 1755 in seinen *Institutiones calculi differentialis* (Kap. 7, Nr. 222-225) bemerkt hat, gilt für homogene Funktionen der folgende

Satz 7. (Eulersche Relation). *Eine auf einer Strahlenmenge Ω des \mathbb{R}^n definierte Funktion $f : \Omega \rightarrow \mathbb{R}$ der Klasse C^1 ist genau dann positiv homogen vom Grade α , wenn für alle $x \in \Omega$ die partielle Differentialgleichung*

$$(29) \quad x \cdot \nabla f(x) = \alpha f(x)$$

gilt, d.h. wenn die „Eulersche Relation“

$$(30) \quad \sum_{j=1}^n x_j f_{x_j}(x) = \alpha f(x)$$

erfüllt ist.

Beweis. (i) Differenzieren wir die Gleichung $f(tx) = t^\alpha f(x)$ nach t , so folgt

$$\nabla f(tx) \cdot x = \alpha t^{\alpha-1} f(x) \quad \text{für } t > 0.$$

Mit $t = 1$ ergibt sich (29).

(ii) Ist umgekehrt (29) erfüllt, so fixieren wir ein beliebiges $x \in \Omega$ und bilden

$$\varphi(t) := t^\alpha f(x) - f(tx), \quad t > 0.$$

Die Funktion $\varphi : (0, \infty) \rightarrow \mathbb{R}$ ist von der Klasse C^1 und erfüllt

$$\dot{\varphi}(t) = \alpha t^{\alpha-1} f(x) - x \cdot f_x(tx) \stackrel{(29)}{=} \alpha t^{\alpha-1} f(x) - \alpha t^{-1} f(tx) = \alpha t^{-1} \varphi(t).$$

Somit befriedigt $\varphi(t)$ die homogene lineare Differentialgleichung

$$\dot{\varphi}(t) = \alpha t^{-1} \varphi(t) \quad \text{auf } (0, \infty)$$

und genügt der Anfangsbedingung $\varphi(1) = 0$. Nach dem Prinzip „Einmal Null, immer Null“ folgt $\varphi(t) \equiv 0$ auf $(0, \infty)$. □

Abschließend führen wir noch einige vielgebrauchte Bezeichnungen ein.

Ist $f = (f_1, \dots, f_n)$ ein Vektorfeld der Klasse $C^1(\Omega, \mathbb{R}^n)$ auf $\Omega \subset \mathbb{R}^n$, so nennt man den Ausdruck

$$(31) \quad \operatorname{div} f := \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \dots + \frac{\partial f_n}{\partial x_n}$$

die *Divergenz* von f . Offenbar ist $\operatorname{div} f$ die Spur der Jacobimatrix $Df(x)$,

$$(32) \quad \operatorname{div} f = \operatorname{spur} (Df).$$

Die Physikerschreibweise hierfür ist

$$(33) \quad \operatorname{div} f = \nabla \cdot f = D_1 f_1 + \dots + D_n f_n ,$$

d.h. man faßt $\nabla \cdot f$ formal als Skalarprodukt zwischen $\nabla = (D_1, \dots, D_n)$ und $f = (f_1, \dots, f_n)$ auf.

Für $n = 3$ und $x = (x_1, x_2, x_3)$ führt man noch das Vektorfeld

$$(34) \quad \operatorname{rot} f := \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) = \nabla \times f$$

ein, die **Rotation des Vektorfeldes** $f : \Omega \rightarrow \mathbb{R}^3$ mit $\Omega \subset \mathbb{R}^3$.

Wir geben einige *Rechenregeln* für das Operieren mit ∇ an, die ohne Mühe nachgeprüft werden können. Dazu bezeichnen wir im folgenden mit λ, μ skalare Funktionen $\Omega \rightarrow \mathbb{R}$ und mit f, g Vektorfunktionen $\Omega \rightarrow \mathbb{R}^3$, wobei $\Omega \subset \mathbb{R}^3$ ist:

$$(35) \quad \begin{aligned} \nabla(\lambda\mu) &= \mu \nabla \lambda + \lambda \nabla \mu , \\ \nabla \cdot (\lambda f) &= (\nabla \lambda) \cdot f + \lambda \nabla \cdot f , \\ \nabla \times (\lambda f) &= (\nabla \lambda) \times f + \lambda \nabla \times f , \\ \nabla \cdot (f \times g) &= g \cdot (\nabla \times f) - f \cdot (\nabla \times g) , \\ \nabla \times (f \times g) &= g \cdot \nabla f - f \cdot \nabla g + f(\nabla \cdot g) - g(\nabla \cdot f) . \end{aligned}$$

In der letzten Formel ist $g \cdot \nabla f$ zu lesen als $(g \cdot \nabla)f$, und entsprechend $f \cdot \nabla g$ als $(f \cdot \nabla)g$.

Warnung. In der älteren Literatur wird für ein Vektorfeld $f : \Omega \rightarrow \mathbb{R}^3$ die Divergenz oft als $\operatorname{div} f = \nabla f$ (statt $\nabla \cdot f$) geschrieben, und $\nabla \cdot f$ steht dort für die Jacobimatrix Df bzw. für deren Transponierte $(Df)^T$, je nach Schreibweise. Deshalb ist es ratsam, sich über die Notation eines Autors zu informieren, bevor man seine Formeln übernimmt.

Aufgaben.

1. Man bestimme die partiellen Ableitungen der folgenden Funktionen $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$:
 $f(x, y) := ax^2 + 2bxy + cy^2$, $g(x, y) := \exp(x^2 + y^2)$, $h(x, y) := \sin(xy) + \cos(xy) + \exp(x^2 y^2)$.
2. Für $\varphi \in C^1(\mathbb{R}^+)$, $\mathbb{R}^+ := (0, \infty)$, und fest gewähltes $x_0 \in \mathbb{R}^n$ berechne man $\nabla f(x)$ der durch $f(x) := \varphi(|x - x_0|)$ definierten Funktion $f : \mathbb{R}^n \setminus \{x_0\} \rightarrow \mathbb{R}$.
3. Fixiere zwei Punkte $a, b \in \mathbb{R}^3$ mit $a \neq b$ und setze

$$f(x) := \frac{1}{|x - a|} + \frac{1}{|x - b|} \quad \text{für } x \in \mathbb{R}^3 \setminus \{a, b\} .$$

Was sind die kritischen Punkte von f ?

4. Für fest gewählte $x_0 \in \mathbb{R}^n$ und $\epsilon > 0$ sei $k : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert durch $k(x) := 0$ für $|x - x_0| \geq \epsilon$ und $k(x) := \exp\left(\frac{1}{|x - x_0| - \epsilon}\right)$ für $|x - x_0| < \epsilon$. Zu zeigen ist: Die partiellen Ableitungen existieren überall auf \mathbb{R}^n und sind stetig.
5. Sei $f(r, \theta, \varphi) := (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta)$ definiert für $(r, \theta, \varphi) \in (0, \infty) \times (0, \pi) \times (0, 2\pi)$. Man berechne die Jacobimatrix Df und die Jacobideterminante J_f . Liefert f einen C^1 -Diffeomorphismus?