

10 Differentialrechnung für Funktionen mehrerer Veränderlicher

Reale Vorgänge hängen in der Regel von mehreren Einflussgrößen ab. Wir befassen uns daher in diesem Abschnitt mit der Differentialrechnung für Funktionen, die auf einer Teilmenge des \mathbb{R}^n definiert sind und in einen Raum \mathbb{R}^m abbilden.

10.1 Lineare Abbildungen und Stetigkeit

Wir bezeichnen wieder mit \mathbb{R}^n den linearen Raum aller Vektoren (x_1, \dots, x_n) mit den Operationen

$$\begin{aligned}(x_1, \dots, x_n) + (y_1, \dots, y_n) &:= (x_1 + y_1, \dots, x_n + y_n), \\ \alpha(x_1, \dots, x_n) &:= (\alpha x_1, \dots, \alpha x_n),\end{aligned}$$

wobei $\alpha \in \mathbb{R}$. Wir kennen bereits mehrere Normen auf \mathbb{R}^n :

$$\begin{aligned}\|(x_1, \dots, x_n)\|_\infty &:= \max_{1 \leq j \leq n} |x_j|, \\ \|(x_1, \dots, x_n)\|_2 &:= \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2}, \\ \|(x_1, \dots, x_n)\|_1 &:= \sum_{j=1}^n |x_j|.\end{aligned}$$

Allgemeiner wird für jedes $p \geq 1$ durch

$$\|(x_1, \dots, x_n)\|_p := \left(\sum_{j=1}^n |x_j|^p \right)^{1/p}$$

eine Norm auf \mathbb{R}^n definiert, die \mathbb{R}^n zu einem normierten linearen Raum macht (↗ Tutorium).

Zwei Normen $\|\cdot\|_A$ und $\|\cdot\|_B$ auf einem linearen Raum X heißen *äquivalent*, wenn es Konstanten $C_1, C_2 > 0$ so gibt, dass

$$C_1 \|x\|_A \leq \|x\|_B \leq C_2 \|x\|_A \quad \text{für alle } x \in X.$$

Satz 10.1 *Alle Normen auf \mathbb{R}^n sind untereinander äquivalent.*

Beweis Sei $\|\cdot\|$ eine Norm auf \mathbb{R}^n . Da die Äquivalenz von Normen eine Äquivalenzrelation ist (HA), genügt es zu zeigen, dass $\|\cdot\|$ zur Maximumnorm $\|\cdot\|_\infty$ äquivalent ist. Sei dazu $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ der j . Einheitsvektor im \mathbb{R}^n . Für $x = (x_1, \dots, x_n) = \sum_{j=1}^n x_j e_j$ ist dann

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|e_j\| \leq \|x\|_\infty \sum_{j=1}^n \|e_j\|,$$

woraus mit $C_2 := \sum_{j=1}^n \|e_j\|$ die Abschätzung

$$\|x\| \leq C_2 \|x\|_\infty \quad \text{für alle } x \in \mathbb{R}^n \quad (10.1)$$

folgt. Aus (10.1) erhalten wir für $x_1, x_2 \in \mathbb{R}^n$

$$\left| \|x_1\| - \|x_2\| \right| \leq \|x_1 - x_2\| \leq C_2 \|x_1 - x_2\|_\infty,$$

d.h. die Abbildung

$$(\mathbb{R}^n, \|\cdot\|_\infty) \rightarrow \mathbb{R}, \quad x \mapsto \|x\| \quad (10.2)$$

ist stetig (sogar Lipschitzstetig). Da $\{x \in \mathbb{R}^n : \|x\|_\infty = 1\}$ beschränkt und abgeschlossen, also kompakt ist, nimmt die Funktion (10.2) ihr Minimum C_1 auf dieser Menge an, und C_1 ist positiv (warum?). Es ist also

$$\|x\| \geq C_1 \quad \text{für alle } x \in \mathbb{R}^n \text{ mit } \|x\|_\infty = 1$$

bzw.

$$\|x\| \geq C_1 \|x\|_\infty \quad \text{für alle } x \in \mathbb{R}^n. \quad \blacksquare$$

Folgerung 10.2 (a) \mathbb{R}^n ist bezüglich jeder Norm vollständig.

- (b) Konvergiert eine Folge im \mathbb{R}^n bzgl. einer Norm, so konvergiert sie bzgl. jeder Norm.
- (c) Alle Normen auf \mathbb{R}^n liefern die gleichen offenen Mengen.
- (d) Die Stetigkeit einer Abbildung $f : X \rightarrow \mathbb{R}^n$ oder $g : \mathbb{R}^n \rightarrow Y$, wobei X und Y metrische Räume sind, hängt nicht von der Wahl der Norm auf \mathbb{R}^n ab.

Der Beweis ist Hausaufgabe. ■

Seien X, Y lineare Räume über \mathbb{R} . Eine Abbildung $A : X \rightarrow Y$ heißt *linear*, wenn

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \quad \forall x, y \in X, \alpha, \beta \in \mathbb{R}.$$

Sind X, Y normierte Räume, so heißt eine lineare Abbildung $A : X \rightarrow Y$ *beschränkt*, wenn sie die Einheitskugel von X in eine beschränkte Menge in Y überführt. Die Zahl

$$\|A\| := \sup \{\|Ax\|_Y : x \in X, \|x\|_X \leq 1\}$$

heißt die durch die Normen $\|\cdot\|_X, \|\cdot\|_Y$ induzierte *Operatornorm* von A . Die Norm $\|A\|$ ist also gleich dem Radius der kleinsten Kugel um $0 \in Y$, die das Bild der Einheitskugel von X unter der Abbildung A enthält. Sind beispielsweise \mathbb{R}^n

und \mathbb{R}^m mit der Maximumnorm versehen und ist (a_{ij}) die Matrixdarstellung von $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ bezüglich der kanonischen Basen, so ist

$$\|A\| = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|.$$

Die Menge der linearen beschränkten Abbildungen von X nach Y bezeichnen wir mit $L(X, Y)$. Versehen mit der Operatornorm wird $L(X, Y)$ zu einem normierten linearen Raum (HA).

Satz 10.3 *Für eine lineare Abbildung $A : X \rightarrow Y$ zwischen normierten Räumen sind folgende Aussagen äquivalent:*

- (a) A ist beschränkt.
- (b) Es gibt ein $C \geq 0$ so, dass $\|Ax\| \leq C\|x\|$ für alle $x \in X$.
- (c) A ist stetig.
- (d) A ist stetig in $0 \in X$.

Beweis (a) \implies (b): Ist $x \neq 0$, so ist $\left\| \frac{x}{\|x\|} \right\| = 1$ und daher $\|A \frac{x}{\|x\|}\| \leq \|A\|$. Also ist $\|Ax\| \leq \|A\| \|x\|$ für alle $x \in X$.

(b) \implies (c): Da A linear ist, ist

$$\|Ax - Ay\| = \|A(x - y)\| \leq C\|x - y\|$$

für beliebige $x, y \in X$. Also ist A sogar Lipschitzstetig.

(c) \implies (d): Trivial.

(d) \implies (a): Sei $\delta > 0$ so, dass $\|Ax - A0\| = \|Ax\| \leq 1$ falls $\|x\| \leq \delta$ (ε - δ -Definition der Stetigkeit mit $\varepsilon = 1$). Dann ist $\|A(\delta x)\| \leq 1$ für alle $\|x\| \leq 1$, d.h. $\|A\| \leq 1/\delta$. ■

Satz 10.4 *Jede lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m ist stetig.*

Beweis Sei $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear. Wegen Folgerung 10.2 (c) können wir annehmen, dass \mathbb{R}^n mit der Maximumnorm versehen ist. Für $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$ ist

$$\|Ax\| = \left\| \sum_{j=1}^n Ax_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|Ae_j\| \leq \|x\|_\infty \sum_{j=1}^n \|Ae_j\|.$$

Nach Satz 10.3, Implikation (b) \implies (c), ist A stetig. ■

Da jeder endlichdimensionale lineare Raum über \mathbb{R} zu einem Raum \mathbb{R}^n isomorph ist, gelten Sätze 10.1 und 10.4 entsprechend für beliebige endlichdimensionale

lineare Räume über \mathbb{R} . Auch für endlichdimensionale Räume über \mathbb{C} bleiben diese Sätze richtig.

Als Ergänzung zur Stetigkeit schauen wir uns noch die Vertauschbarkeit von Grenzübergängen an. Genauer: wir fragen, wann für $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gilt

$$\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y) = \lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y)?$$

Beispiel 1 Auf $(0, 1] \times (0, 1]$ sei $f(x, y) = x^y$. Dann ist

$$\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} x^y = \lim_{y \rightarrow 0} 0 = 0, \quad \lim_{x \rightarrow 0} \lim_{y \rightarrow 0} x^y = \lim_{x \rightarrow 0} 1 = 1. \quad \blacksquare$$

Beispiel 2 Auf $(0, 1] \times (0, 1]$ sei

$$f(x, y) = \frac{x \sin \frac{1}{x} + y}{x + y}.$$

Dann ist $\lim_{x \rightarrow 0} f(x, y) = 1$ und $\lim_{y \rightarrow 0} f(x, y) = \sin \frac{1}{x}$. Folglich ist

$$\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} f(x, y) = 1,$$

aber $\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y)$ existiert nicht. \blacksquare

Satz 10.5 Sei $f : (a, b) \times (c, d) \rightarrow \mathbb{R}$ und $(x_0, y_0) \in [a, b] \times [c, d]$. Der Grenzwert

$$A := \lim_{(x, y) \rightarrow (x_0, y_0)} f(x, y) \quad (10.3)$$

soll existieren, und für jedes $y \in (c, d)$ existiere der Grenzwert

$$\varphi(y) := \lim_{x \rightarrow x_0} f(x, y). \quad (10.4)$$

Dann existiert auch der iterierte Grenzwert $\lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y)$ und ist gleich A . Eine analoge Aussage gilt, wenn für jedes $x \in (a, b)$ der Grenzwert

$$\psi(x) := \lim_{y \rightarrow y_0} f(x, y) \quad (10.5)$$

existiert. Existieren also alle Grenzwerte (10.3) – (10.5), so ist

$$\lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y) = \lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y) = A.$$

Beweis Wir zeigen nur die erste Aussage. Existenz des Grenzwertes (10.3) bedeutet:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall (x, y) \neq (x_0, y_0) \quad \text{mit } \|(x, y) - (x_0, y_0)\|_\infty < \delta : \\ |f(x, y) - A| < \varepsilon.$$

Mit der Definition der Maximumnorm ist also

$$\forall (x, y) \neq (x_0, y_0) \quad \text{mit } |x - x_0| < \delta, |y - y_0| < \delta : |f(x, y) - A| < \varepsilon.$$

Wir fixieren ein $y \neq y_0$ mit $|y - y_0| < \delta$ und lassen $x \rightarrow x_0$ streben. Dann folgt

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall y \neq y_0 \quad \text{mit } |y - y_0| < \delta : |\varphi(y) - A| \leq \varepsilon.$$

Also existiert $\lim_{y \rightarrow y_0} \varphi(y)$ und ist gleich A . \blacksquare

10.2 Partielle Ableitungen

In diesem Abschnitt sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$. Ist $x = (x_1, \dots, x_n) \in U$, so schreiben wir statt $f(x)$ auch $f(x_1, \dots, x_n)$.

Definition 10.6 Die Funktion f heißt in $x \in U$ partiell differenzierbar bzgl. x_i , falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h} \quad (10.6)$$

existiert. Dieser Grenzwert heißt die partielle Ableitung von f bzgl. x_i an der Stelle x und wird mit $(D_i f)(x)$, $\frac{\partial f}{\partial x_i}(x)$ oder $f_{x_i}(x)$ bezeichnet. Die Funktion f heißt partiell differenzierbar in x (auf U), wenn alle partiellen Ableitungen von f in x (in jedem Punkt von U) existieren. Ist außerdem jede dieser partiellen Ableitungen in x (bzw. auf U) stetig, so heißt f stetig partiell differenzierbar in x (bzw. auf U).

Ist e_i der i . Einheitsvektor von \mathbb{R}^n , so ist also

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \rightarrow 0} \frac{f(x + he_i) - f(x)}{h}.$$

Beispiel 1 Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben durch

$$f(x_1, \dots, x_n) = \sqrt{x_1^2 + \dots + x_n^2} \quad (= \|x\|_2). \quad (10.7)$$

Beim partiellen Differenzieren betrachten wir nur die i . der Variablen x_1, \dots, x_n als veränderlich und die übrigen als fixiert. Wir können daher die bekannten Differentiationsregeln für Funktionen einer reellen Veränderlichen anwenden und erhalten: Die Funktion (10.7) ist auf $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar, und

$$\frac{\partial f}{\partial x_i}(x) = \frac{\partial f}{\partial x_i}(x_1^2 + \dots + x_i^2 + \dots + x_n^2)^{1/2} = \frac{1}{2}(x_1^2 + \dots + x_n^2)^{-1/2} \cdot 2x_i = \frac{x_i}{\|x\|_2}.$$

Offenbar ist die Funktion (10.7) auf $\mathbb{R}^n \setminus \{0\}$ sogar *stetig* partiell differenzierbar. ■

Beispiel 2 Bei Funktionen auf \mathbb{R}^2 oder \mathbb{R}^3 schreibt man oft $f(x, y)$ bzw. $f(x, y, z)$ statt $f(x_1, x_2)$ bzw. $f(x_1, x_2, x_3)$. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ erklärt durch $f(0, 0) = 0$ und

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2} \quad \text{falls } (x, y) \neq (0, 0).$$

Für $(x, y) \neq (0, 0)$ finden wir die partiellen Ableitungen sofort:

$$f_x(x, y) = \frac{y}{(x^2 + y^2)^2} - 4 \frac{x^2 y}{(x^2 + y^2)^3}, \quad f_y(x, y) = \frac{x}{(x^2 + y^2)^2} - 4 \frac{xy^2}{(x^2 + y^2)^3}.$$

An der Stelle $(x, y) = (0, 0)$ arbeiten wir mit Definition 10.6:

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0.$$

Analog ist $f_y(0, 0) = 0$. Also ist f auf ganz \mathbb{R}^2 partiell differenzierbar. Man beachte, dass f in $(0, 0)$ nicht stetig ist! Es ist nämlich

$$f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{(1/n^2)}{(2/n^2)^2} = \frac{n^2}{4} \rightarrow \infty \quad \text{für } n \rightarrow \infty.$$

Dieses Beispiel zeigt, dass aus der partiellen Differenzierbarkeit *nicht* die Stetigkeit folgt. Später werden wir sehen, dass dagegen aus der *stetigen* partiellen Differenzierbarkeit die Stetigkeit folgt. ■

Ist $f : U \rightarrow \mathbb{R}$ partiell differenzierbar und sind alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} = D_i f : U \rightarrow \mathbb{R}$ wieder partiell differenzierbar, so heißt f *zweimal partiell differenzierbar*, und wir schreiben $\frac{\partial^2 f}{\partial x_j \partial x_i} = D_j D_i f = f_{x_i x_j}$ für die partielle Ableitung von $\frac{\partial f}{\partial x_i}$ nach x_j . Allgemein heißt f *k-mal partiell differenzierbar* ($k \geq 2$) wenn f $(k-1)$ -mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $k-1$ partiell differenzierbar sind. Schließlich heißt f *k-mal stetig partiell differenzierbar*, wenn f k -mal partiell differenzierbar ist und alle partiellen Ableitungen bis zur k . Ordnung stetig sind.

Beispiel 3 Wir erklären $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ durch $f(0, 0) = 0$ und

$$f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2} \quad \text{für } (x, y) \neq (0, 0).$$

Für alle Punkte $(x, y) \neq (0, 0)$ ist

$$f_x(x, y) = y \frac{x^4 - y^4 + 4x^2 y^2}{(x^2 + y^2)^2}, \quad f_y(x, y) = x \frac{x^4 - y^4 - 4x^2 y^2}{(x^2 + y^2)^2},$$

und für $(x, y) = (0, 0)$ ist

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0 \quad \text{und} \quad f_y(0, 0) = 0.$$

Für die gemischten zweiten Ableitungen in $(0, 0)$ finden wir schließlich

$$\begin{aligned} f_{xy}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_x(0, h) - f_x(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{-h - 0}{h} = -1, \\ f_{yx}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h - 0}{h} = 1. \end{aligned} \tag{10.8}$$

Die Reihenfolge der partiellen Ableitungen darf also i. Allg. nicht vertauscht werden. Der folgende Satz gibt Bedingungen an, die dieses Vertauschen erlauben. ■

Satz 10.7 (H.A. Schwarz) Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ und $x^* \in U$. Alle partiellen Ableitungen erster Ordnung von f sollen auf U existieren. Weiter existiere die zweite Ableitung $\frac{\partial^2 f}{\partial x_i \partial x_j}$ auf U , und diese sei in x^* stetig. Dann existiert auch $\frac{\partial^2 f}{\partial x_j \partial x_i}$ in x^* , und es gilt

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x^*).$$

Beweis Wir beschränken uns auf den Fall $n = 2$, schreiben (x, y) statt (x_1, x_2) und setzen voraus, dass $\frac{\partial^2 f}{\partial y \partial x}$ auf U existiert und in x^* stetig ist. Weiter nehmen wir an, dass $(0, 0) \in U$ und $x^* = (0, 0)$ (andernfalls verschieben wir U geeignet). Schließlich wählen wir $\delta > 0$ so, dass $(-\delta, \delta) \times (-\delta, \delta) \subseteq U$ und arbeiten im weiteren ausschließlich auf diesem Quadrat. Wir beginnen mit einer Vorüberlegung. Seien $h, k \in (-\delta, \delta) \setminus \{0\}$ fixiert. Der Mittelwertsatz, angewandt auf die Funktion

$$F : (-\delta, \delta) \rightarrow \mathbb{R}, \quad s \mapsto f(s, k) - f(s, 0)$$

liefert die Existenz eines $\xi = \xi(h, k)$ zwischen 0 und h so, dass

$$F(h) - F(0) = hF'(\xi) = h\left(\frac{\partial f}{\partial x}(\xi, k) - \frac{\partial f}{\partial x}(\xi, 0)\right).$$

Erneute Anwendung des Mittelwertsatzes, nun auf die Funktion

$$G : (-\delta, \delta) \rightarrow \mathbb{R}, \quad t \mapsto \frac{\partial f}{\partial x}(\xi, t),$$

liefert die Existenz eines $\eta = \eta(h, k)$ zwischen 0 und k so, dass

$$\frac{\partial f}{\partial x}(\xi, k) - \frac{\partial f}{\partial x}(\xi, 0) = G(k) - G(0) = kG'(\eta) = k\frac{\partial^2 f}{\partial y \partial x}(\xi, \eta).$$

Es ist also

$$F(h) - F(0) = f(h, k) - f(h, 0) - f(0, k) + f(0, 0) = hk\frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)$$

mit gewissen $\xi = \xi(h, k)$, $\eta = \eta(h, k)$. Damit wird

$$\begin{aligned} \frac{f_y(h, 0) - f_y(0, 0)}{h} &= \lim_{k \rightarrow 0} \frac{f(h, k) - f(h, 0) - f(0, k) + f(0, 0)}{hk} \\ &= \lim_{k \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \end{aligned} \quad (10.9)$$

mit gewissen Zahlen ξ, η , die von h und k abhängen und für die $|\xi| \leq |h|$ und $|\eta| \leq |k|$. Für jedes fixierte h existiert also der Grenzwert (10.9). Außerdem

existiert wegen der Stetigkeit von $\frac{\partial^2 f}{\partial y \partial x}$ in $(0, 0)$ und wegen $(\xi, \eta) \rightarrow (0, 0)$ für $(h, k) \rightarrow (0, 0)$ der Grenzwert

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \quad \text{und ist gleich} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Nach Satz 10.5 existiert dann auch der iterierte Grenzwert

$$\lim_{h \rightarrow 0} \left(\lim_{k \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \right) \quad \text{und ist gleich} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Wegen (10.9) bedeutet dies, dass der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \frac{\partial^2 f}{\partial x \partial y}(0, 0) \quad \text{existiert und gleich} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) \quad \text{ist.}$$

Das ist die Behauptung. ■

Entsprechend gilt natürlich $f_{xxy} = f_{yyx} = f_{xyx}$, falls nur eine dieser Ableitungen existiert und stetig ist.

Eine Funktion $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}^m$, gegeben durch

$$f(x) = f(x_1, \dots, x_n) = \left(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n) \right),$$

heißt *partiell differenzierbar* bzw. *stetig partiell differenzierbar*, wenn jede ihrer Komponenten diese Eigenschaft besitzt.

Beispiel 4 Sei $U \subseteq \mathbb{R}^n$ offen. Ist $f : U \rightarrow \mathbb{R}$ partiell differenzierbar, so heißt der Vektor

$$(\text{grad } f)(x) := \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \in \mathbb{R}^n$$

der *Gradient* von f in x . Ist $f : U \rightarrow \mathbb{R}^n$ partiell differenzierbar, so heißt die Zahl

$$(\text{div } f)(x) := \frac{\partial f_1}{\partial x_1}(x) + \dots + \frac{\partial f_n}{\partial x_n}(x) \in \mathbb{R}$$

die *Divergenz* von f in x . Ist schließlich $n = 3$ und $f : U \rightarrow \mathbb{R}^3$ partiell differenzierbar, so heißt der Vektor

$$(\text{rot } f)(x) := \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) \in \mathbb{R}^3$$

die *Rotation* von f in x .

Aus dem Satz von Schwarz folgt: Ist $n = 3$ und $f : U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar, so gilt

$$\text{rot grad } f = 0. \tag{10.10}$$

Die erste Komponente des Vektors $\operatorname{rot} \operatorname{grad} f$ ist nämlich

$$\frac{\partial}{\partial x_2} \frac{\partial f}{\partial x_3} - \frac{\partial}{\partial x_3} \frac{\partial f}{\partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2} = 0.$$

Ebenso sind die übrigen Komponenten gleich 0. Ähnlich zeigt man, dass für jede zweimal stetig partiell differenzierbare Funktion $f : U \rightarrow \mathbb{R}^3$ gilt:

$$\operatorname{div} \operatorname{rot} f = 0. \quad (10.11)$$

10.3 Differenzierbarkeit

In diesem Abschnitt sehen wir uns an, wie sich das Konzept der Differentiation auf Funktionen mehrerer Veränderlicher übertragen lässt. Dazu erinnern wir an die Zerlegungsformel für differenzierbare Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$:

$$f(x+h) = f(x) + \alpha h + r(h) \quad \text{mit } \alpha \in \mathbb{R} \text{ und } \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

Wir können diese Formel und damit die Grundidee der Differentiation wie folgt interpretieren: Die Funktion f wird im Punkt x lokal durch die lineare Funktion $h \mapsto \alpha h$ approximiert. Im Weiteren sei $\|\cdot\|$ die Euklidische Norm.

Definition 10.8 Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}^m$ heißt differenzierbar in $x \in U$, wenn es eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sowie eine in einer Umgebung W von $0 \in \mathbb{R}^n$ definierte Funktion r mit Werten in \mathbb{R}^m und mit

$$\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$$

so gibt, dass

$$f(x+h) = f(x) + Ah + r(h) \quad \text{für alle } h \in W. \quad (10.12)$$

Man beachte, dass A und r von x abhängen und dass A eindeutig bestimmt ist (HA). Ist f in jedem Punkt $x \in U$ differenzierbar, so heißt f differenzierbar auf U . In der Literatur spricht man auch von *Fréchet-Differenzierbarkeit* oder *totaler Differenzierbarkeit*.

Man beachte auch, dass die Räume $\mathbb{R}^n, \mathbb{R}^m$ in dieser Definition keine wesentliche Rolle spielen. Eine ganz analoge Definition, bei der man allerdings zusätzlich die Stetigkeit von A fordert, trifft man für beliebige Banachräume.

Wir überlegen uns die folgenden Beziehungen zwischen den eingeführten Differenzierbarkeitsbegriffen:

$$\begin{array}{ccccc} f \text{ stetig partiell} & \text{Satz 10.11} & f \text{ differen-} & \text{Satz 10.9} & f \text{ partiell} \\ \text{differenzierbar} & \implies & \text{zierbar} & \implies & \text{differenzierbar} \\ & & \Downarrow \text{Satz 10.10} & & \\ & & f \text{ stetig} & & \end{array}$$

Satz 10.9 Sei $U \subseteq \mathbb{R}^n$ offen und $f = (f_1, \dots, f_m)^T : U \rightarrow \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist jede Funktion $f_i : U \rightarrow \mathbb{R}$ in x partiell differenzierbar, und die Matrixdarstellung von $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ aus (10.12) bezüglich der Standardbasen von \mathbb{R}^n bzw. \mathbb{R}^m ist

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}. \quad (10.13)$$

Die Abbildung A in (10.12) ist also eindeutig bestimmt. Sie heißt *Ableitung* von f in x und wird mit $(Df)(x)$ oder $f'(x)$ bezeichnet. Die Matrix (10.13) heißt *Jacobi-Matrix* von f in x und wird oft mit $J_f(x)$ bezeichnet.

Beweis Es gelte (10.12) mit $A = (a_{ij})_{i,j=1}^{mn}$ und $r = (r_1, \dots, r_m)^T$, $h = (h_1, \dots, h_n)^T$. Mit diesen Bezeichnungen gilt für jedes $i = 1, \dots, m$

$$f_i(x+h) = f_i(x) + \sum_{j=1}^n a_{ij}h_j + r_i(h) \quad \forall h \in W, \quad (10.14)$$

wobei $\lim_{h \rightarrow 0} \frac{r_i(h)}{\|h\|} = 0$. Letzteres folgt aus $|r_i| \leq \|r\|_\infty$. Wir fixieren nun ein j zwischen 1 und n und wählen $h = (0, \dots, 0, h_j, 0, \dots, 0) = h_j e_j$. Für hinreichend kleine h_j liegen diese Vektoren in W , und (10.14) reduziert sich auf

$$f_i(x + h_j e_j) = f_i(x) + a_{ij}h_j + r_i(h),$$

woraus folgt

$$\frac{f_i(x + h_j e_j) - f_i(x)}{h_j} = a_{ij} + \frac{r_i(h)}{h_j} \quad \text{für } h_j \neq 0.$$

Wegen $|h_j| = \|h\|$ können wir $h_j \rightarrow 0$ streben lassen und bekommen

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{h_j \rightarrow 0} \frac{f_i(x + h_j e_j) - f_i(x)}{h_j} = a_{ij}. \quad \blacksquare$$

Satz 10.10 Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist f in x stetig.

Beweis Da lineare Abbildungen von \mathbb{R}^n nach \mathbb{R}^m stetig sind (Satz 10.4) und 0 in 0 überführen, geht die rechte Seite von (10.12) für $h \rightarrow 0$ gegen $f(x)$. Also existiert $\lim_{h \rightarrow 0} f(x+h)$ und ist gleich $f(x)$. \blacksquare

Satz 10.11 Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ in $x \in U$ stetig partiell differenzierbar. Dann ist f in x differenzierbar.

Beweis Es genügt, diese Aussage für jede Komponente von f zu zeigen, d.h. wir nehmen $m = 1$ an. Für $h = (h_1, \dots, h_n) \in W$ definieren wir

$$x^{(0)} = x \quad \text{und} \quad x^{(i)} := x + \sum_{j=1}^i h_j e_j \quad \text{für} \quad i = 1, \dots, n.$$

Insbesondere ist $x^{(n)} = x + h$. Da sich $x^{(i)}$ und $x^{(i-1)}$ nur in der i . Komponente unterscheiden, gibt es nach dem Mittelwertsatz für Funktionen *einer* Veränderlichen für jedes $i = 1, \dots, n$ ein $t_i \in [0, 1]$ so, dass

$$f(x^{(i)}) - f(x^{(i-1)}) = \frac{\partial f}{\partial x_i}(x^{(i-1)} + t_i h_i e_i) \cdot h_i.$$

Aufsummieren liefert

$$f(x + h) - f(x) = f(x^{(n)}) - f(x^{(0)}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i + r(h)$$

(Teleskopsumme) mit

$$r(h) = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x^{(i-1)} + t_i h_i e_i) - \frac{\partial f}{\partial x_i}(x) \right) h_i.$$

Da alle partiellen Ableitungen in x nach Voraussetzung stetig sind, ist

$$\lim_{h \rightarrow 0} \left(\frac{\partial f}{\partial x_i}(x^{(i-1)} + t_i h_i e_i) - \frac{\partial f}{\partial x_i}(x) \right) = 0$$

für jedes $i = 1, \dots, n$ und folglich $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$. ■

Wir nennen stetig partiell differenzierbare Funktionen daher auch kurz *stetig differenzierbar*. Keine der im Schema vor Satz 10.9 angegebenen Implikationen lässt sich umkehren: In Beispiel 2 aus Abschnitt 10.2 ist eine partiell differenzierbare, aber nicht differenzierbare Funktion angegeben. Ein Beispiel für eine differenzierbare, aber nicht stetig differenzierbare Funktion steht in Heuser, Ana II, Pkt. 164, Aufg. 7.

Die aus Kapitel 7 bekannten Differentiationsregeln übertragen sich ohne Änderung auf den allgemeinen Fall:

Satz 10.12 Seien $U \subseteq \mathbb{R}^n$ offen und $f, g : U \rightarrow \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist für $\alpha, \beta \in \mathbb{R}$ auch die Funktion $\alpha f + \beta g$ in x differenzierbar, und es gilt

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x).$$

Die Produkt- und Quotientenregel vermerken wir nur für skalarwertige Funktionen.

Satz 10.13 Seien $U \subseteq \mathbb{R}^n$ offen und $f, g : U \rightarrow \mathbb{R}$ in $x \in U$ differenzierbar. Dann sind auch die Funktionen $fg : U \rightarrow \mathbb{R}$ und (falls $g(x) \neq 0$) $f/g : U \rightarrow \mathbb{R}$ in x differenzierbar, und es ist

$$\begin{aligned}(fg)'(x) &= g(x)f'(x) + f(x)g'(x), \\ \left(\frac{f}{g}\right)'(x) &= \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)}.\end{aligned}$$

Die Beweise dieser beiden Sätze werden wie im Fall $n = 1$ geführt und sind HA.

Satz 10.14 (Kettenregel) Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offene Mengen, und $g : U \rightarrow \mathbb{R}^m$ und $f : V \rightarrow \mathbb{R}^k$ seien Funktionen mit $g(U) \subseteq V$. Ist g in $x_0 \in U$ und f in $g(x_0) \in V$ differenzierbar, so ist die zusammengesetzte Funktion $f \circ g : U \rightarrow \mathbb{R}^k$ in x_0 differenzierbar, und es gilt

$$(f \circ g)'(x_0) = f'(g(x_0)) \circ g'(x_0).$$

Das \circ auf der rechten Seite steht für die Verkettung der linearen Abbildungen $g'(x_0)$ und $f'(g(x_0))$. In Matrixschreibweise bedeutet das gerade das Matrixprodukt der $k \times m$ -Matrix $f'(g(x_0))$ mit der $m \times n$ -Matrix $g'(x_0)$.

Beweis Differenzierbarkeit von g in x_0 bzw. f in $g(x_0)$ bedeutet

$$g(x) - g(x_0) = g'(x_0)(x - x_0) + r(x - x_0), \quad (10.15)$$

$$f(y) - f(g(x_0)) = f'(g(x_0))(y - g(x_0)) + s(y - g(x_0)), \quad (10.16)$$

wobei

$$\lim_{x \rightarrow x_0} \frac{r(x - x_0)}{\|x - x_0\|} = 0, \quad \lim_{y \rightarrow g(x_0)} \frac{s(y - g(x_0))}{\|y - g(x_0)\|} = 0.$$

Wir setzen in (10.16) $y = g(x)$ und anschließend (10.16) in (10.15) ein:

$$f(g(x)) - f(g(x_0)) = f'(g(x_0))g'(x_0)(x - x_0) + t(x_0, x)$$

mit

$$t(x_0, x) = f'(g(x_0))r(x - x_0) + s(g(x) - g(x_0)).$$

Wir müssen zeigen, dass

$$\frac{t(x_0, x)}{\|x - x_0\|} = f'(g(x_0))\frac{r(x - x_0)}{\|x - x_0\|} + \frac{s(g(x) - g(x_0))}{\|x - x_0\|}$$

für $x \rightarrow x_0$ gegen 0 strebt. Für den ersten Summanden ist dies wegen der Stetigkeit der linearen Abbildung $f'(g(x_0))$ und wegen $\lim_{h \rightarrow 0} r(h)/\|h\| = 0$ klar. Für

den zweiten Summanden beachten wir, dass $s(0) = 0$. Für $g(x) = g(x_0)$ ist also $s(g(x) - g(x_0)) = 0$, und für $g(x) \neq g(x_0)$ haben wir

$$\frac{s(g(x) - g(x_0))}{\|x - x_0\|} = \frac{s(g(x) - g(x_0))}{\|g(x) - g(x_0)\|} \frac{\|g(x) - g(x_0)\|}{\|x - x_0\|}.$$

Wegen der Stetigkeit von g in x_0 und wegen $\lim_{h \rightarrow 0} s(h)/\|h\| = 0$ ist

$$\lim_{x \rightarrow x_0} \frac{s(g(x) - g(x_0))}{\|g(x) - g(x_0)\|} = 0,$$

und wir zeigen noch, dass der Quotient $\|g(x) - g(x_0)\|/\|x - x_0\|$ in einer Umgebung von x_0 beschränkt bleibt. Dies folgt aber aus der Zerlegungsformel für g :

$$\frac{\|g(x) - g(x_0)\|}{\|x - x_0\|} = \frac{\|g'(x_0)(x - x_0) + r(x - x_0)\|}{\|x - x_0\|} \leq \|g'(x_0)\| + \frac{\|r(x - x_0)\|}{\|x - x_0\|}. \quad \blacksquare$$

Wir sehen uns die Kettenregel für einige Spezialfälle an.

Beispiel 1 Die reellwertigen Funktionen f bzw. x_1, \dots, x_n seien auf der offenen Menge $U \subseteq \mathbb{R}^n$ bzw. auf dem offenen Intervall $I \subseteq \mathbb{R}$ definiert, und die verkettete Funktion $F(t) := f(x_1(t), \dots, x_n(t))$ soll auf I erklärt sein. Sind alle Funktionen f und x_i auf I differenzierbar, so ist auch F auf I differenzierbar, und

$$\frac{dF}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt},$$

genauer: für $t_0 \in I$ ist

$$\frac{dF}{dt}(t_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \left(x_1(t_0), \dots, x_n(t_0) \right) \cdot \frac{dx_i}{dt}(t_0).$$

Der Beweis folgt sofort aus der Kettenregel, angewandt auf die äußere Funktion f und die innere Funktion $g(t) := (x_1(t), \dots, x_n(t))^T : I \rightarrow \mathbb{R}^n$. \blacksquare

Beispiel 2 Die reellwertigen Funktionen f bzw. u_1, \dots, u_n seien auf der offenen Menge $U \subseteq \mathbb{R}^n$ bzw. der offenen Menge $V \subseteq \mathbb{R}^m$ definiert. Wir betrachten die verkettete Funktion

$$F(x_1, \dots, x_m) = f\left(u_1(x_1, \dots, x_m), \dots, u_n(x_1, \dots, x_m)\right)$$

auf V . Ist f auf U differenzierbar und jede Funktion u_i auf V partiell differenzierbar, so ist F auf V partiell differenzierbar, und es gilt

$$\frac{\partial F}{\partial x_i} = \frac{\partial f}{\partial u_1} \frac{\partial u_1}{\partial x_i} + \dots + \frac{\partial f}{\partial u_n} \frac{\partial u_n}{\partial x_i} \quad \text{für alle } i.$$

Dies folgt sofort aus Beispiel 1. Ist sogar jede der Funktionen u_i differenzierbar, so ist F auf V differenzierbar, und es ist

$$F' = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_m} \right) \quad \text{auf } V. \quad \blacksquare$$

Beispiel 3 Sei $x : (0, 1) \rightarrow \mathbb{R}^n \setminus \{0\}$ eine differenzierbare Funktion. Dann ist die verkettete Funktion

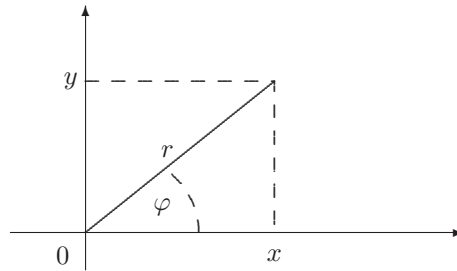
$$f : (0, 1) \rightarrow \mathbb{R}, \quad t \mapsto \|x(t)\|_2$$

nach Beispiel 1 aus Abschnitt 10.2 und nach Satz 10.14 differenzierbar, und die Kettenregel liefert wie in Beispiel 1

$$f'(t) = \sum_{i=1}^n \frac{x_i(t)}{\|x(t)\|_2} \cdot x'_i(t) = \frac{\langle x(t), x'(t) \rangle}{\|x(t)\|_2},$$

wobei $\langle \cdot, \cdot \rangle$ für das übliche Skalarprodukt im \mathbb{R}^n steht. ■

Beispiel 4 Wir sehen uns die Anwendung der Kettenregel bei der Transformation von Differentialausdrücken an. Durch Einführung von Polarkoordinaten $x = r \cos \varphi, y = r \sin \varphi$ auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ wird aus einer Funktion $u = u(x, y)$ eine Funktion $v = v(r, \varphi) = u(r \cos \varphi, r \sin \varphi)$.



Wir zeigen, dass dabei beispielsweise der Ausdruck $x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x}$ in $\frac{\partial v}{\partial r}$ übergeht. Aus $x = r \cos \varphi$ und $y = r \sin \varphi$ folgt $r = \sqrt{x^2 + y^2}$ und $\varphi = \arctan \frac{y}{x}$. Differenzieren liefert

$$\frac{\partial r}{\partial x} = \frac{2x}{2\sqrt{x^2 + y^2}} = \frac{r \cos \varphi}{r} = \cos \varphi$$

und analog $\frac{\partial r}{\partial y} = \sin \varphi$ sowie

$$\frac{\partial \varphi}{\partial x} = \frac{1}{1 + \left(\frac{y}{x}\right)^2} \left(\frac{-y}{x^2}\right) = \frac{-y}{x^2 + y^2} = \frac{-r \sin \varphi}{r^2} = -\frac{\sin \varphi}{r}$$

und analog $\frac{\partial \varphi}{\partial y} = \frac{\cos \varphi}{r}$. Demzufolge ist nach Beispiel 2

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial v}{\partial \varphi} \frac{\partial \varphi}{\partial x} = \frac{\partial v}{\partial r} \cos \varphi - \frac{\partial v}{\partial \varphi} \frac{\sin \varphi}{r}, \\ \frac{\partial u}{\partial y} &= \frac{\partial v}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial v}{\partial \varphi} \frac{\partial \varphi}{\partial y} = \frac{\partial v}{\partial r} \sin \varphi + \frac{\partial v}{\partial \varphi} \frac{\cos \varphi}{r}. \end{aligned}$$

Hieraus folgt schließlich

$$\begin{aligned} x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x} &= r \cos \varphi \left(\frac{\partial v}{\partial r} \sin \varphi + \frac{\partial v}{\partial \varphi} \frac{\cos \varphi}{r} \right) - r \sin \varphi \left(\frac{\partial v}{\partial r} \cos \varphi - \frac{\partial v}{\partial \varphi} \frac{\sin \varphi}{r} \right) \\ &= \frac{\partial v}{\partial r} (r \sin \varphi \cos \varphi - r \sin \varphi \cos \varphi) + \frac{\partial v}{\partial \varphi} (\cos^2 \varphi + \sin^2 \varphi) = \frac{\partial v}{\partial \varphi}. \end{aligned}$$

Es ist also beispielsweise genau dann $x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x} = 0$, wenn $\frac{\partial v}{\partial \varphi} = 0$, d.h. genau dann, wenn v nur von r abhängt, also rotations-symmetrisch ist. ■

10.4 Richtungsableitungen

Definition 10.15 Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$, $x \in U$ und $v \in \mathbb{R}^n$ ein Vektor der Länge $\|v\|_2 = 1$. Man sagt, dass f eine Ableitung in Richtung des Vektors v besitzt, wenn der Grenzwert

$$\lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}$$

existiert. Wir bezeichnen diesen Grenzwert mit $\frac{\partial f}{\partial v}$.

Die partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ sind spezielle Richtungsableitungen in Richtung der Einheitsvektoren $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ mit der 1 an i . Stelle.

Satz 10.16 Seien U, x, f wie in Definition 10.15, und f sei in x differenzierbar. Dann existiert für jeden Einheitsvektor $v = (v_1, \dots, v_n)^T \in \mathbb{R}^n$ die Ableitung von f im Punkt x in Richtung v , und es gilt

$$\frac{\partial f}{\partial v}(x) = f'(x)v = \frac{\partial f}{\partial x_1}(x)v_1 + \dots + \frac{\partial f}{\partial x_n}(x)v_n. \quad (10.17)$$

Beweis Für hinreichend kleines t ist nach Definition von $f'(x)$

$$\frac{f(x + tv) - f(x)}{t} = \frac{f'(x) \cdot tv + r(tv)}{t} = f'(x)v + \frac{r(tv)}{t}.$$

Aus $\lim_{t \rightarrow 0} \frac{\|r(tv)\|}{|t|} = \lim_{t \rightarrow 0} \frac{\|r(tv)\|}{\|tv\|} = 0$ folgt die Behauptung. ■

Mit Hilfe des Gradienten $\text{grad } f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$ können wir (10.17) auch als

$$\frac{\partial f}{\partial v}(x) = \langle (\text{grad } f)(x), v \rangle \quad (10.18)$$

schreiben, wobei $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ das Skalarprodukt auf \mathbb{R}^n ist. Wie im \mathbb{R}^2 führt man den Winkel $\varphi \in [0, \pi]$ zwischen zwei Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ durch

$$\langle x, y \rangle = \|x\| \|y\| \cos \varphi$$

ein. Ist also $(\text{grad } f)(x) \neq 0$, so folgt aus (10.18) für den Winkel φ zwischen v und $(\text{grad } f)(x)$:

$$\frac{\partial f}{\partial v}(x) = \|(\text{grad } f)(x)\| \cos \varphi.$$

Folgerung 10.17 *Seien U, f, x, v wie in Definition 10.15, und sei $(\text{grad } f)(x) \neq 0$. Dann ist die Richtungsableitung $\frac{\partial f}{\partial v}(x)$ genau dann maximal, wenn $\cos \varphi = 1$ bzw. $\varphi = 0$, d.h. wenn $(\text{grad } f)(x)$ und v die gleiche Richtung haben. Der Gradient zeigt also in die Richtung des stärksten Anstieges von f in x . Diese Tatsache wird bei der numerischen Lösung von Extremalaufgaben benutzt.*

10.5 Der Mittelwertsatz

Am Ende von Abschnitt 7.6 haben wir gesehen, dass der Mittelwertsatz in seiner gewohnten Form für vektorwertige Funktionen nicht mehr gilt. Man hat jedoch für reellwertige Funktionen auf \mathbb{R}^n die folgende Version.

Satz 10.18 (Mittelwertsatz) *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ differenzierbar. Weiter sei $x + th \in U$ für alle $t \in [0, 1]$. Dann existiert ein $\tau \in (0, 1)$ mit*

$$f(x + h) - f(x) = f'(x + \tau h)h.$$

Beweis Wir betrachten die differenzierbare Funktion

$$g : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(x + th).$$

Nach dem Mittelwertsatz 7.27 für Funktionen einer Veränderlichen gibt es ein $\tau \in (0, 1)$ mit

$$f(x + h) - f(x) = g(1) - g(0) = \frac{dg}{dt}(\tau).$$

Nach der Kettenregel (Beispiel 1 aus 10.3) ist weiter

$$\frac{dg}{dt}(\tau) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x + \tau h)h_i = f'(x + \tau h)h. \quad \blacksquare$$

Für Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $m \geq 2$ ist die folgende Version des Mittelwertsatzes die nächstbeste und sehr nützlich.

Satz 10.19 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar und $x + th \in U$ für alle $t \in [0, 1]$. Dann ist*

$$f(x + h) - f(x) = \int_0^1 f'(x + \tau h)h \, d\tau.$$

Ist insbesondere $\|f'(x + th)\| \leq M$ für alle $t \in [0, 1]$, so folgt

$$\|f(x + h) - f(x)\| \leq M\|h\|.$$

Das Integral über die vektorwertige Funktion $f'(x + \tau h)h$ ist komponentenweise erklärt.

Beweis Für $g : [0, 1] \rightarrow \mathbb{R}^m$, $g(t) := f(x + th)$ ist $g'(t) = f'(x + th)h$ und daher

$$f(x + h) - f(x) = g(1) - g(0) = \int_0^1 g'(\tau) d\tau = \int_0^1 f'(x + \tau h)h d\tau,$$

womit die erste Aussage gezeigt ist. Die zweite folgt aus

$$\begin{aligned} \left\| \int_0^1 f'(x + \tau h)h d\tau \right\| &\leq \int_0^1 \|f'(x + \tau h)h\| d\tau \\ &\leq \int_0^1 \|f'(x + \tau h)\| \|h\| d\tau \leq \int_0^1 M \|h\| d\tau = M \|h\|. \end{aligned} \quad \blacksquare$$

10.6 Der Satz von Taylor

Wir lernen nun den Satz von Taylor für Funktionen mehrerer Veränderlicher kennen. Dabei beschränken wir uns auf reellwertige Funktionen. Wir führen zunächst einige Bezeichnungen ein, die uns helfen, die Übersicht über die zahlreichen Summanden in der Taylorentwicklung zu behalten.

Ein *Multiindex* ist ein n -Tupel $(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$. Für jeden Multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$ sei $|\alpha| := \alpha_1 + \dots + \alpha_n$ seine *Ordnung* und $\alpha! := \alpha_1! \dots \alpha_n!$ seine *Fakultät*. Für jede $|\alpha|$ -mal partiell differenzierbare Funktion $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ setzen wir

$$D^\alpha f := D_1^{\alpha_1} \dots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}},$$

und für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ sei

$$x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}.$$

Als Vorbereitung für den Satz von Taylor zeigen wir:

Satz 10.20 Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ k -mal stetig partiell differenzierbar, und sei $x + th \in U$ für alle $t \in [0, 1]$. Dann ist die Funktion

$$g : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(x + th)$$

k -mal stetig differenzierbar, und es gilt

$$\frac{d^k g}{dt^k}(\tau) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x + \tau h) h^\alpha \quad (10.19)$$

(die Summation erfolgt über alle Multiindizes der Ordnung k).

Beweis Wir zeigen zuerst mit vollständiger Induktion, dass

$$\frac{d^k g}{dt^k}(\tau) = \sum_{i_1, \dots, i_k=1}^n (D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \cdot \dots \cdot h_{i_k}. \quad (10.20)$$

Für $k = 1$ haben wir dies bereits im Beweis von Satz 10.18 getan. Aus der Kettenregel folgt nämlich

$$\begin{aligned} \frac{dg}{dt}(\tau) &= \frac{df(x_1 + th_1, \dots, x_n + th_n)}{dt}(\tau) = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt} \\ &= \frac{\partial f}{\partial x_1} h_1 + \dots + \frac{\partial f}{\partial x_n} h_n = \sum_{i_1=1}^n \frac{\partial f}{\partial x_{i_1}}(x + \tau h) h_{i_1} \\ &= \sum_{i_1=1}^n (D_{i_1} f)(x + \tau h) h_{i_1}. \end{aligned}$$

Nehmen wir an, dass (10.20) für ein $k - 1 \geq 1$ richtig ist, so folgt analog

$$\begin{aligned} \frac{d^k g}{dt^k}(\tau) &= \frac{d}{dt} \left(\sum_{i_1, \dots, i_{k-1}=1}^n (D_{i_{k-1}} \dots D_{i_1} f)(x + th) h_{i_1} \dots h_{i_{k-1}} \right) (\tau) \\ &= \sum_{i_k=1}^n D_{i_k} \left(\sum_{i_1, \dots, i_{k-1}=1}^n (D_{i_{k-1}} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_{k-1}} \right) h_{i_k} \\ &= \sum_{i_1, \dots, i_k=1}^n (D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_k}. \end{aligned}$$

Damit ist (10.20) gezeigt. Wir überlegen uns nun, dass die rechte Seite von (10.20) gleich der rechten Seite von (10.19) ist. Grundidee ist, dass es nach dem Satz von Schwarz auf die Reihenfolge der partiellen Ableitungen nicht ankommt und wir daher die partiellen Ableitungen umsortieren können. Kommt unter den Indizes i_1, \dots, i_k die Zahl 1 genau α_1 -mal vor, die Zahl 2 genau α_2 -mal, \dots , und die Zahl n genau α_n -mal, so ergibt die Umsortierung und Zusammenfassung gleicher Ableitungen

$$(D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_k} = (D_1^{\alpha_1} \dots D_n^{\alpha_n} f)(x + \tau h) h_1^{\alpha_1} \dots h_n^{\alpha_n}.$$

Da es gerade $\frac{k!}{\alpha_1! \dots \alpha_n!} = \frac{k!}{\alpha!}$ k -Tupel (i_1, \dots, i_k) gibt, in denen die Zahl j genau α_j -mal vorkommt (beachten Sie: $\alpha_1 + \dots + \alpha_n = k$), gilt

$$\begin{aligned} \frac{d^k g}{dt^k}(\tau) &= \sum_{i_1, \dots, i_k}^n (D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_k} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D_1^{\alpha_1} \dots D_n^{\alpha_n} f)(x + \tau h) h_1^{\alpha_1} \dots h_n^{\alpha_n} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x + \tau h) h^\alpha. \quad \blacksquare \end{aligned}$$

Satz 10.21 (Taylor) Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ $(k+1)$ -mal stetig partiell differenzierbar, und sei $x + th \in U$ für alle $t \in [0, 1]$. Dann gibt es ein $\tau \in (0, 1)$ so, dass

$$f(x+h) = \underbrace{\sum_{|\alpha| \leq k} \frac{1}{\alpha!} (D^\alpha f)(x) h^\alpha}_{\text{Taylorpolynom der Ordnung } k} + \underbrace{\sum_{|\alpha|=k+1} \frac{(D^\alpha f)(x + \tau h)}{\alpha!} h^\alpha}_{\text{Restglied}}.$$

Beispiel Wir bestimmen das Taylorpolynom der Ordnung 2 der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x_1, x_2) \mapsto e^{x_1^2 + \cos x_2}$ im Punkt $x = (0, 0)$. Dieses ist gleich

$$\begin{aligned} f(0) &+ (D_1 f)(0)h_1 + (D_2 f)(0)h_2 + \frac{1}{2} (D_1^2 f)(0)h_1^2 \\ &+ \frac{1}{2} (D_2^2 f)(0)h_2^2 + (D_1 D_2 f)(0)h_1 h_2, \end{aligned}$$

und wir bestimmen die partiellen Ableitungen von f bis zur 2. Ordnung:

$$\begin{aligned} (D_1 f)(x_1, x_2) &= 2x_1 e^{x_1^2 + \cos x_2} &\implies (D_1 f)(0) &= 0. \\ (D_2 f)(x_1, x_2) &= -\sin x_2 e^{x_1^2 + \cos x_2} &\implies (D_2 f)(0) &= 0. \\ (D_1^2 f)(x_1, x_2) &= (2 + 4x_1^2) e^{x_1^2 + \cos x_2} &\implies (D_1^2 f)(0) &= 2e. \\ (D_2^2 f)(x_1, x_2) &= (-\cos x_2 + \sin^2 x_2) e^{x_1^2 + \cos x_2} &\implies (D_2^2 f)(0) &= -e. \\ (D_1 D_2 f)(x_1, x_2) &= -2x_1 \sin x_2 e^{x_1^2 + \cos x_2} &\implies (D_1 D_2 f)(0) &= 0. \end{aligned}$$

Das gesuchte Taylorpolynom ist also

$$(h_1, h_2) \mapsto e + eh_1^2 - \frac{e}{2} h_2^2. \quad \blacksquare$$

Beweis von Satz 10.21 Die Funktion $g : [0, 1] \rightarrow \mathbb{R}$, $t \mapsto f(x + th)$ ist nach Satz 10.20 $(k+1)$ -mal stetig differenzierbar. Der Satz von Taylor für Funktionen einer Veränderlichen behauptet die Existenz eines $\tau \in (0, 1)$ so, dass

$$f(x+h) = g(1) = \sum_{m=0}^k \frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\tau)}{(k+1)!}.$$

Wieder nach Satz 10.20 ist

$$\frac{g^{(m)}(0)}{m!} = \sum_{|\alpha|=m} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha \quad \text{sowie} \quad \frac{g^{(k+1)}(\tau)}{(k+1)!} = \sum_{|\alpha|=k+1} \frac{(D^\alpha f)(x + \tau h)}{\alpha!} h^\alpha,$$

woraus die Behauptung folgt. \blacksquare

Anmerkung 1 Seien α, β Multiindizes gleicher Länge. Man rechnet leicht nach, dass für $f(x) := x^\beta$ gilt

$$(D^\alpha f)(x) = \begin{cases} \frac{\beta!}{(\beta - \alpha)!} x^{\beta - \alpha} & \text{falls } \beta \geq \alpha \\ 0 & \text{sonst} \end{cases}$$

($\beta - \alpha$ und $\beta \geq \alpha$ sind komponentenweise zu verstehen). Insbesondere ist

$$(D^\alpha f)(0) = \begin{cases} \beta! & \text{falls } \alpha = \beta \\ 0 & \text{sonst.} \end{cases}$$

Hieraus folgt, dass wie bei Funktionen einer Veränderlichen alle partiellen Ableitungen in x bis zur k . Ordnung der Funktion f mit denen ihres Taylorpolynoms der Ordnung k übereinstimmen. ■

Anmerkung 2 Man kann den Satz von Taylor auch für vektorwertige Funktionen $f = (f_1, \dots, f_m)^T : U \rightarrow \mathbb{R}^m$ formulieren und beweisen. Definieren wir für solche Funktionen $D_i f := (D_i f_1, \dots, D_i f_m)^T : U \rightarrow \mathbb{R}^m$, so sieht das entsprechende Taylorpolynom (welches nun ein Vektor ist) formal genauso aus wie in Satz 10.21. Für $m > 1$ muss jedoch das Restglied modifiziert werden (vgl. Abschnitt 10.5 für den Mittelwertsatz). ■

Wir sehen uns die Polynome $P_m(h) := \sum_{|\alpha|=m} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha$ für $m = 0, 1, 2$ genauer an.

$m = 0$: Notwendigerweise ist $\alpha = (0, \dots, 0)$ und daher $P_0(h) = f(x)$.

$m = 1$: Die einzigen n -Tupel $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 1$ sind die "Einheitsvektoren" $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ mit der 1 an der j . Stelle. Wegen $D^{e_j} f = D_j f$, $e_j! = 1$ und $h^{e_j} = h_j$ ist

$$P_1(h) = \sum_{j=1}^n (D_j f)(x) h_j = \langle (\text{grad } f)(x), h \rangle.$$

$m = 2$: Wir haben im Beweis von Satz 10.20 gesehen, dass

$$P_2(h) = \sum_{|\alpha|=2} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha = \frac{1}{2} \sum_{i,j=1}^n (D_i D_j f)(x) h_i h_j.$$

Um dies kompakter zu schreiben, bezeichnen wir für jede zweimal stetig partiell differenzierbare Funktion $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ die $n \times n$ - Matrix

$$\left((D_i D_j f)(x) \right)_{i,j=1}^n = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} (x) \right)_{i,j=1}^n$$

mit $(\text{Hess } f)(x)$ und nennen sie die *Hesse-Matrix* oder den *Hessian* von f in x . Nach dem Satz von Schwarz ist $(\text{Hess } f)(x)$ eine symmetrische Matrix. Mit dieser können wir $P_2(h)$ schreiben als

$$P_2(h) = \frac{1}{2} \langle (\text{Hess } f)(x)h, h \rangle.$$

Folgerung 10.22 *Ist $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ dreimal stetig partiell differenzierbar und $x + th \in U$ für alle $t \in [0, 1]$, so ist*

$$f(x + h) = c + \langle a, h \rangle + \frac{1}{2} \langle Ah, h \rangle + R_2(x, h)$$

mit $c = f(x)$, $a = (\text{grad } f)(x)$, $A = (\text{Hess } f)(x)$ und einem Restglied R_2 wie im Satz 10.21.

10.7 Lokale Extrema

Wir benutzen nun Folgerung 10.22 zur Untersuchung des lokalen Verhaltens von Funktionen $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$. Für offenes $U \subseteq \mathbb{R}^n$ sei $C^k(U)$ die Menge aller k -mal stetig partiell differenzierbaren Funktionen $f : U \rightarrow \mathbb{R}$. Eine Funktion $f : U \rightarrow \mathbb{R}$ besitzt in $x_0 \in U$ ein *lokales Minimum* (bzw. *Maximum*), wenn für alle x aus einer Umgebung $V \subseteq U$ von x_0 gilt

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

Tritt die Gleichheit nur für $x = x_0$ ein, nennen wir x_0 ein *isoliertes* lokales Minimum (bzw. Maximum).

Satz 10.23 (Notwendige Bedingung) *Sei U offen und $f : U \rightarrow \mathbb{R}$ partiell differenzierbar. Besitzt f in $x_0 \in U$ ein lokales Extremum (Minimum oder Maximum), so ist $(\text{grad } f)(x_0) = 0$.*

Beweis Für $i = 1, \dots, n$ betrachten wir die Funktionen $g_i(t) := f(x_0 + te_i)$. Diese sind auf einem Intervall $(-\varepsilon, \varepsilon)$ mit $\varepsilon > 0$ definiert und differenzierbar, und sie besitzen in $t = 0$ ein lokales Extremum. Nach Lemma 7.9 ist $g_i'(0) = 0$. Nun ist aber $g_i'(0) = \frac{\partial f}{\partial x_i}(x_0)$. Also ist $(\text{grad } f)(x_0) = 0$. ■

Das Verschwinden aller partiellen Ableitungen in x_0 ist also eine *notwendige* Bedingung für das Vorliegen eines lokalen Extremums. Wie bei Funktionen einer Veränderlichen erhält man hinreichende Bedingungen durch Betrachten der zweiten Ableitungen. Wir treffen dazu einige Vorbereitungen.

Definition 10.24 *Sei A eine symmetrische ($A = A^T$) reelle $n \times n$ Matrix. A heißt*

- positiv definit, wenn $\langle Ax, x \rangle > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.

- positiv semidefinit, wenn $\langle Ax, x \rangle \geq 0$ für alle $x \in \mathbb{R}^n$.
- negativ definit (semidefinit), wenn $-A$ positiv definit (semidefinit) ist.
- indefinit, wenn es $x, y \in \mathbb{R}^n$ gibt mit $\langle Ax, x \rangle > 0$ und $\langle Ay, y \rangle < 0$.

Aus der linearen Algebra kennen wir Kriterien für die Definitheit. So gilt:

Satz 10.25 Sei A symmetrische reelle $n \times n$ - Matrix. Dann ist A

- positiv definit \iff alle Eigenwerte von A sind positiv.
- positiv semidefinit \iff alle Eigenwerte von A sind nichtnegativ.
- indefinit \iff A hat sowohl positive als auch negative Eigenwerte.

Satz 10.26 (Hurwitz-Kriterium) Eine symmetrische reelle $n \times n$ - Matrix $A = (a_{ij})_{i,j=1}^n$ ist genau dann positiv definit, wenn

$$\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0 \quad \text{für } k = 1, \dots, n.$$

Insbesondere für $n = 2$ ist die Anwendung dieses Kriteriums sehr einfach. Wir formulieren nun hinreichende Kriterien für das Vorliegen von Extremwerten.

Satz 10.27 Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^2(U)$, $x_0 \in U$ und $(\text{grad } f)(x_0) = 0$. Dann gilt:

- ist $(\text{Hess } f)(x_0)$ positiv definit, so hat f in x_0 ein isoliertes Minimum.
- ist $(\text{Hess } f)(x_0)$ negativ definit, so hat f in x_0 ein isoliertes Maximum.
- ist $(\text{Hess } f)(x_0)$ indefinit, so besitzt f in x_0 kein lokales Extremum.

Ist $(\text{Hess } f)(x_0)$ nur semidefinit, so ist keine Entscheidung möglich.

Beweis Wir zeigen nur die Aussage (a) und nehmen der Einfachheit halber $f \in C^3(U)$ an. Nach Folgerung 10.22 gilt für alle $x_0 + x$ aus einer Umgebung von x_0

$$f(x_0 + x) = f(x_0) + \frac{1}{2} \langle Ax, x \rangle + R(x) \quad (10.21)$$

mit $A := (\text{Hess } f)(x_0)$. Da $S := \{y \in \mathbb{R}^n : \|y\| = 1\}$ kompakt ist, nimmt die stetige Funktion

$$S \rightarrow \mathbb{R}, \quad y \mapsto \langle Ay, y \rangle$$

auf S ihr Minimum α an. Da $\langle Ay, y \rangle > 0$ für alle $y \in S$, ist insbesondere $\alpha > 0$. Ist nun $x \in \mathbb{R}^n \setminus \{0\}$, so ist $x/\|x\| \in S$ und folglich

$$\left\langle A \frac{x}{\|x\|}, \frac{x}{\|x\|} \right\rangle \geq \alpha > 0 \quad \text{bzw.} \quad \langle Ax, x \rangle \geq \alpha \|x\|^2. \quad (10.22)$$

Die letztere Abschätzung gilt offenbar auch für $x = 0$ und damit für alle $x \in \mathbb{R}$. (Umgekehrt folgt aus dieser Abschätzung natürlich die positive Definitheit von A ; beide Aussagen sind also äquivalent.)

Weiter: für das Restglied $R(x)$ in (10.21) haben wir wegen $|x_i| \leq \|x\|$ offenbar die Abschätzung

$$|R(x)| = \left| \sum_{|\alpha|=3} \frac{(D^\alpha f)(x_0 + \tau x)}{\alpha!} x^\alpha \right| \leq C \|x\|^3$$

für alle x aus einer hinreichend kleinen Umgebung von 0. Wählen wir diese Umgebung so klein, dass auch $C\|x\| \leq \alpha/4$, so wird

$$|R(x)| \leq C \|x\|^3 = C \|x\| \|x\|^2 \leq \frac{\alpha}{4} \|x\|^2.$$

Hieraus und aus (10.21) und (10.22) folgt schließlich

$$f(x_0 + x) \geq f(x_0) + \frac{\alpha}{2} \|x\|^2 - \frac{\alpha}{4} \|x\|^2 = f(x_0) + \frac{\alpha}{4} \|x\|^2$$

für alle x aus einer (hinreichend kleinen) Umgebung von 0. Also besitzt f in x_0 ein isoliertes lokales Minimum. ■

Beispiel Für $f(x, y) = x^2 + \frac{1}{2}y^2 + xy + 10$ ist $(\text{grad } f)(0, 0) = 0$, und die Matrix

$$(\text{Hess } f)(0, 0) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

ist positiv definit (Hurwitz-Kriterium). Also besitzt f in $(0, 0)$ ein lokales isoliertes Minimum. ■

10.8 Parameterabhängige Integrale

In diesem Abschnitt betrachten wir folgendes Problem: Wird eine Funktion f zweier Veränderlicher bzgl. einer der Veränderlichen integriert, so hängt das Ergebnis von der zweiten Veränderlichen ab. Beispielsweise ist für $f(x, y) = x^y$ für $x, y \geq 0$

$$\int_1^2 f(x, y) dx = \int_1^2 x^y dx = \frac{x^{y+1}}{y+1} \Big|_1^2 = \frac{2^{y+1} - 1}{y+1}.$$

Die Frage ist, unter welchen Voraussetzungen an f diese Abhängigkeit stetig oder sogar differenzierbar ist.

Satz 10.28 Sei $U \subseteq \mathbb{R}^n$ offen und $D = [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall. Die Funktion $f : D \times U \rightarrow \mathbb{R}$ sei stetig. Dann ist auch die Funktion

$$F : U \rightarrow \mathbb{R}, \quad x \mapsto \int_a^b f(t, x) dt$$

stetig. Hat f zusätzlich stetige partielle Ableitungen $\frac{\partial f}{\partial x_i} : D \times U \rightarrow \mathbb{R}$, $i = 1, \dots, n$, so ist auch F stetig partiell differenzierbar, und es gilt

$$\frac{\partial}{\partial x_i} \int_a^b f(t, x) dt = \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt.$$

Unter den getroffenen Voraussetzungen dürfen Integration und Differentiation also vertauscht werden.

Beweis Wir überlegen uns zuerst die Stetigkeit von F . Sei $x \in U$. Da U offen ist, gibt es ein $r > 0$ so, dass $U_{2r}(x) \subseteq U$. Dann liegt aber erst recht die abgeschlossene Kugel $\overline{U_r(x)} = \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$ in U . Weiter: die Menge

$$D \times \overline{U_r(x)} = [a, b] \times \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$$

ist abgeschlossen und beschränkt in $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$, also kompakt. Also ist f auf dieser Menge sogar gleichmäßig stetig (Satz 6.41). Insbesondere gibt es zu jedem $\varepsilon > 0$ ein $\delta \in (0, r)$ so, dass

$$|f(t, x + h) - f(t, x)| < \frac{\varepsilon}{b - a}$$

für alle $t \in [a, b]$ und alle $h \in \mathbb{R}^n$ mit $\|h\| < \delta$. Integration liefert

$$|F(x + h) - F(x)| \leq \int_a^b |f(t, x + h) - f(t, x)| dt \leq \frac{\varepsilon}{b - a} (b - a) = \varepsilon$$

für alle h mit $\|h\| < \delta$. Also ist F in x stetig.

Für den Beweis der zweiten Aussage sei wieder $x \in U$ beliebig. Wir wählen ein $r \in \mathbb{R}$ so, dass $x + he_i \in U$ für alle $h \in (-r, r)$. Wir zeigen, dass die Funktion

$$g(t, x, h) := \begin{cases} \frac{f(t, x + he_i) - f(t, x)}{h} & \text{falls } h \neq 0 \\ \frac{\partial f}{\partial x_i}(t, x) & \text{falls } h = 0 \end{cases}$$

auf der Menge

$$\tilde{U} := \{(t, x, h) \in [a, b] \times U \times (-r, r)\}$$

stetig ist. In allen Punkten $(t, x, h) \in \tilde{U}$ mit $h \neq 0$ ist dies klar, so dass wir noch die Stetigkeit in allen Punkten $(t, x, 0) \in \tilde{U}$ zeigen müssen. Seien (t_n, x_n, h_n)

Punkte aus \tilde{U} mit $\lim_{n \rightarrow \infty} (t_n, x_n, h_n) = (t, x, 0)$ und $h_n \neq 0$ (für $h_n = 0$ ist die Aussage wieder offensichtlich). Nach dem Mittelwertsatz der Differentialrechnung für Funktionen einer Veränderlichen, angewandt auf die Funktion

$$a : s \mapsto f(t_n, x_n + sh_n e_i),$$

gibt es ein $\xi_n \in [0, 1]$ mit

$$\frac{a(1) - a(0)}{h_n} = \frac{f(t_n, x_n + h_n e_i) - f(t_n, x_n)}{h_n} = \frac{\partial f}{\partial x_i}(t_n, x_n + \xi_n h_n e_i).$$

Für $n \rightarrow \infty$ ist $h_n \rightarrow 0$ und folglich auch $\xi_n h_n \rightarrow 0$. Wegen der Stetigkeit von $\frac{\partial f}{\partial x_i}$ ist daher

$$\begin{aligned} g(t_n, x_n, h_n) &= \frac{f(t_n, x_n + h_n e_i) - f(t_n, x_n)}{h_n} \\ &= \frac{\partial f}{\partial x_i}(t_n, x_n + \xi_n h_n e_i) \rightarrow \frac{\partial f}{\partial x_i}(t, x) = g(t, x, 0). \end{aligned}$$

Also ist g stetig. Wenden wir die Aussage des ersten Teils des Satzes an, erhalten wir

$$\begin{aligned} \frac{\partial F}{\partial x_i}(x) &= \frac{\partial}{\partial x_i} \int_a^b f(t, x) dt = \lim_{h \rightarrow 0} \int_a^b g(t, x, h) dt \\ &= \int_a^b g(t, x, 0) dt = \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt. \end{aligned}$$

Schließlich hängt diese Funktion – wieder nach dem ersten Teil des Satzes – stetig von x ab. ■

Beispiel Für $|t| < 1$ berechnen wir das Integral

$$F(t) := \int_0^\pi \ln(1 - 2t \cos x + t^2) dx.$$

Um Satz 10.28 benutzen zu können, wählen wir ein $a \in (|t|, 1)$ und betrachten F auf dem kompakten Intervall $[-a, a]$. (Man beachte, dass $1 - 2t \cos x + t^2 > 0$ für $|t| < 1$.) Die Funktion

$$f(x, t) := \ln(1 - 2t \cos x + t^2)$$

ist nach t stetig partiell differenzierbar und hat die Ableitung

$$\frac{\partial f}{\partial t} = \frac{2t - 2 \cos x}{1 - 2t \cos x + t^2} \quad \text{auf} \quad [0, \pi] \times [-a, a].$$

Aus der zweiten Aussage von Satz 10.28 folgt

$$F'(t) = \int_0^\pi \frac{2t - 2 \cos x}{1 - 2t \cos x + t^2} dx.$$

Offenbar ist $F'(0) = 0$. Für $t \neq 0$ substituieren wir

$$s := \tan \frac{x}{2} \quad \text{bzw.} \quad x = 2 \arctan s.$$

Dann ist $\frac{dx}{ds} = \frac{2}{1+s^2}$ bzw. $dx = \frac{2ds}{1+s^2}$ sowie

$$\cos x = \frac{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}} = \frac{1 - \tan^2 \frac{x}{2}}{1 + \tan^2 \frac{x}{2}} = \frac{1 - s^2}{1 + s^2}.$$

Das gesuchte Integral geht damit über in

$$\int_0^\infty \frac{2t - 2\frac{1-s^2}{1+s^2}}{1 - 2t\frac{1-s^2}{1+s^2} + t^2} \cdot \frac{2}{1+s^2} ds = 4 \int_0^\infty \frac{s^2(1+t) - (1-t)}{(1+t)^2 s^4 + 2(1+t^2)s^2 + (1-t)^2} ds.$$

Partialbruchzerlegung liefert

$$4 \frac{s^2(1+t) - (1-t)}{(1+t)^2 s^4 + 2(1+t^2)s^2 + (1-t)^2} = \frac{2}{t} \left(\frac{1}{s^2+1} + \frac{t^2-1}{(1+t)^2 s^2 + (1-t)^2} \right).$$

Mit dem Grundintegral

$$\int \frac{1}{ax^2+c} dx = \frac{1}{\sqrt{ac}} \arctan \sqrt{\frac{a}{c}} x \quad \text{für} \quad ac > 0$$

erhalten wir

$$F'(t) = \frac{2}{t} \left(\arctan s - \arctan \left(\frac{1+t}{1-t} s \right) \right) \Big|_0^\infty = 0$$

(beachten Sie: $\arctan 0 = 0$, $\lim_{s \rightarrow \infty} \arctan s = \frac{\pi}{2}$). Also ist $F'(t) = 0$ für alle $t \in (-1, 1)$, und F eine konstante Funktion. Wir bestimmen ihren (einzigsten) Wert, indem wir im Ausgangsintegral $t = 0$ setzen:

$$F(t) = F(0) = \int_0^\pi \ln 1 dx = 0. \quad \blacksquare$$

Als weitere Anwendung von Satz 10.28 betrachten wir iterierte Integrale und ihre Berechnung. Sei $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ eine stetige Funktion. Nach Satz 10.28 ist die Funktion

$$F(y) := \int_a^b f(x, y) dx$$

auf $[c, d]$ stetig und kann folglich integriert werden. Man bezeichnet

$$\int_c^d F(y) dy = \int_c^d \int_a^b f(x, y) dx dy$$

als *iteriertes Integral*. Der folgende Satz sagt, dass es auf die Reihenfolge der Integrationen nicht ankommt.

Satz 10.29 Sei $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ stetig. Dann ist

$$\int_c^d \int_a^b f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx.$$

Beweis Wir erklären $\varphi : [c, d] \rightarrow \mathbb{R}$ durch

$$\varphi(y) := \int_a^b \int_c^y f(x, t) dt dx.$$

Dann ist $\varphi(c) = 0$, φ ist nach Satz 10.28 differenzierbar, und es gilt

$$\varphi'(y) = \int_a^b \frac{\partial}{\partial y} \int_c^y f(x, t) dt dx = \int_a^b f(x, y) dx.$$

Hieraus folgt

$$\int_c^d \int_a^b f(x, y) dx dy = \int_c^d \varphi'(y) dy = \varphi(d) - \varphi(c) = \varphi(d) = \int_a^b \int_c^d f(x, y) dy dx. \quad \blacksquare$$

Ein analoger Satz gilt für n -fache Integrale einer stetigen Funktion auf einem Quader im \mathbb{R}^n .

Abschließend betrachten wir noch uneigentliche Parameterintegrale der Gestalt

$$\int_a^\infty f(x, y) dx, \quad y \in [c, d].$$

Da das Integrationsintervall nicht mehr kompakt ist, lässt sich der Beweis von Satz 10.28 nicht unmittelbar übertragen, und wir benötigen stärkere Voraussetzungen.

Definition 10.30 Das Integral $\int_a^\infty f(x, y) dx$ heißt auf $[c, d]$ gleichmäßig konvergent, wenn es für jedes $y \in [c, d]$ konvergiert und wenn für jedes $\varepsilon > 0$ ein $b_0 \geq a$ existiert, so dass

$$\left| \int_b^\infty f(x, y) dx \right| < \varepsilon \text{ für alle } b \geq b_0 \text{ und alle } y \in [c, d].$$

Man beachte, dass b_0 unabhängig von y ist.

Satz 10.31 (a) Sei $f : [a, \infty) \times [c, d] \rightarrow \mathbb{R}$ stetig und beschränkt, und das Integral $\int_a^\infty f(x, y) dx$ konvergiere gleichmäßig auf $[c, d]$. Dann ist die Funktion

$$\varphi : [c, d] \rightarrow \mathbb{R}, \quad y \mapsto \int_a^\infty f(x, y) dx \quad (10.23)$$

stetig. Außerdem konvergiert das Integral $\int_a^\infty \int_c^d f(x, y) dy dx$, und es gilt

$$\int_c^d \int_a^\infty f(x, y) dx dy = \int_a^\infty \int_c^d f(x, y) dy dx.$$

(b) Sei $f : [a, \infty) \times [c, d] \rightarrow \mathbb{R}$ stetig, beschränkt und nach y partiell differenzierbar, die Ableitung $\frac{\partial f}{\partial y}$ sei auf $[a, \infty) \times [c, d]$ stetig und beschränkt, das Integral $\int_a^\infty \frac{\partial f}{\partial y}(x, y) dx$ sei gleichmäßig konvergent, und das Integral $\int_a^\infty f(x, c) dx$ konvergiere. Dann konvergiert das Integral $\int_a^\infty f(x, y) dx$ für jedes $y \in [c, d]$, die Funktion (10.23) ist differenzierbar, und

$$\varphi'(y) = \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx.$$

Beweis (a) Für alle $y, y + h \in [c, d]$ und $b \geq a$ ist

$$\begin{aligned} \varphi(y + h) - \varphi(y) &= \int_a^\infty (f(x, y + h) - f(x, y)) dx & (10.24) \\ &= \int_a^b (f(x, y + h) - f(x, y)) dx + \int_b^\infty f(x, y + h) dx - \int_b^\infty f(x, y) dx. \end{aligned}$$

Sei $\varepsilon > 0$. Wir wählen b so groß, dass der Betrag der letzten beiden Integrale in (10.24) jeweils kleiner als $\varepsilon/3$ wird (gleichmäßige Konvergenz!). Weiter wissen wir aus Satz 10.28, dass $y \mapsto \int_a^b f(x, y) dx$ eine stetige Funktion ist, und daher wird auch das erste Integral in (10.24) kleiner als $\varepsilon/3$, wenn nur h hinreichend klein ist, etwa für $h \leq h_0$. Für alle $h \leq h_0$ ist also

$$|\varphi(y + h) - \varphi(y)| < \varepsilon.$$

Für die zweite Aussage von (a) sei wieder $\varepsilon > 0$. Wir wählen $b_0 \geq a$ so, dass

$$\left| \int_a^\infty f(x, y) dx - \int_a^{b_0} f(x, y) dx \right| < \varepsilon \quad \text{für alle } y \in [c, d] \text{ und } b \geq b_0$$

(gleichmäßige Konvergenz!). Integrieren liefert

$$\left| \int_c^d \int_a^\infty f(x, y) dx dy - \int_c^d \int_a^{b_0} f(x, y) dx dy \right| < \varepsilon(d - c).$$

Mit Satz 10.29 vertauschen wir die Integrationsreihenfolge im 2. Integral und erhalten

$$\left| \int_c^d \int_a^\infty f(x, y) dx dy - \int_a^{b_0} \int_c^d f(x, y) dy dx \right| < \varepsilon(d - c).$$

Da dies für jedes $b \geq b_0$ gilt, folgt die Behauptung.

(b) Nach Teil (a) konvergiert für jedes $t \in [c, d]$ das uneigentliche Integral

$$\int_a^\infty \int_c^t \frac{\partial f}{\partial y}(x, y) dy dx,$$

und es gilt

$$\int_c^t \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx dy = \int_a^\infty \int_c^t \frac{\partial f}{\partial y}(x, y) dy dx. \quad (10.25)$$

Da $\frac{\partial f}{\partial y}$ stetig ist, folgt mit dem Hauptsatz der Differential- und Integralrechnung

$$\int_c^t \frac{\partial f}{\partial y}(x, y) dy = f(x, t) - f(x, c).$$

Integration bzgl. x über $[a, \infty)$ liefert wegen (10.25) (und wegen der Konvergenz von $\int_a^\infty f(x, c) dx$ nach Voraussetzung) für $t \in [c, d]$

$$\varphi(t) = \int_a^\infty f(x, t) dx = \int_a^\infty f(x, c) dx + \int_c^t \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx dy.$$

Nun ist $y \mapsto \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx$ stetig nach Teil (a). Wieder nach dem Hauptsatz der Differential- und Integralrechnung ist φ differenzierbar, und

$$\varphi'(t) = \int_a^\infty \frac{\partial f}{\partial y}(x, t) dx. \quad \blacksquare$$

Aufgabe: Benutzen Sie diesen Satz um zu zeigen, dass die Eulersche Gammafunktion

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt, \quad x \geq 1$$

stetig und sogar unendlich oft differenzierbar ist. (Für $x \in (0, 1)$ ist der Integrand unbeschränkt und Satz 10.31 nicht mehr unmittelbar anwendbar.)