2 Ableitungsregeln

Wir beginnen diesen Abschnitt mit der Kettenregel für differenzierbare Funktionen.

2.1 Satz. (Kettenregel). *Es seien $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ offene Mengen, $f : U \to \mathbb{R}^m$ und $g : V \to \mathbb{R}^l$ Abbildungen mit $f(U) \subset V$. Ferner sei f in $x_0 \in U$ und g in $y_0 := f(x_0)$ differenzierbar. Dann ist die Abbildung $g \circ f : U \to \mathbb{R}^l$ in x_0 differenzierbar und es gilt*

$$D(g \circ f)(x_0) = Dg(f(x_0)) \cdot Df(x_0),$$

bzw. in Matrixschreibweise

$$J_{g \circ f}(x_0) = J_g(f(x_0)) \cdot J_g(x_0).$$

Im obigen Satz bedeutet „·“ zum einen die Hintereinanderausführung von linearen Abbildungen und zum anderen die Multiplikation von Matrizen.

Beweis. Wir setzen $A := Df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ und $B := Dg(f(x_0)) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^l)$. Nach Voraussetzung gilt für $x = x_0 + h \in U$ und $y = y_0 + k \in V$

$$\begin{aligned} f(x_0 + h) &= f(x_0) + Ah + r_f(x) \\ g(y_0 + k) &= g(y_0) + Bk + r_g(y), \end{aligned}$$

mit

$$\lim_{h \to 0} \frac{r_f(x)}{\|h\|} = 0 = \lim_{k \to 0} \frac{r_g(y)}{\|k\|}.$$

Setzt man $k := f(x_0 + h) - f(x_0) = Ah + r_f(x)$, so erhalten wir

$$\begin{aligned} (g \circ f)(x_0 + h) &= g(\underbrace{f(x_0)}_{=y_0} + \underbrace{Ah + r_f(x)}_{=:k}) \\ &= g(f(x_0)) + \underbrace{BAh + Br_f(x)}_{=Bk} + r_g(y), \end{aligned}$$

und es bleibt zu zeigen, dass

$$\lim_{h \to 0} \frac{Br_f(x)}{\|h\|} = 0 \quad \text{und} \quad \lim_{h \to 0} \frac{r_g(y)}{\|h\|} = 0$$

gilt. Da $B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^l)$ nach Satz VI.2.6 stetig ist, folgt

$$\lim_{h \to 0} \frac{Br_f(x)}{\|h\|} = B \lim_{h \to 0} \frac{r_f(x)}{\|h\|} = B0 = 0,$$

und somit die erste Aussage.
Für den Beweis der zweiten Aussage notieren wir, dass $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ nach Satz

VI.2.6 wiederum stetig ist. Nach demselben Satz existiert eine Konstante $M > 0$ mit $\|Ah\|_{\mathbb{R}^m} \leq L\|h\|_{\mathbb{R}^n}$ für alle $h \in \mathbb{R}^n$ und daher gilt

$$\|k\| = \|Ah + r_f(x)\| \leq (M + \frac{\|r_f(x)\|}{\|h\|})\|h\|.$$

Für $h \to 0$ gilt auch $k = f(x_0 + h) - f(x_0) \to 0$ und somit ist

$$\frac{\|r_g(y)\|}{\|h\|} = \frac{\|r_g(y)\|}{\|k\|} \cdot \frac{\|k\|}{\|h\|} \leq \frac{\|r_g(y)\|}{\|k\|} \cdot (M + \frac{\|r_f(x)\|}{\|h\|}).$$

Also gilt

$$\lim_{h \to 0} \frac{r_g(y)}{\|h\|} = 0.$$

□

2.2 Beispiel. Wir betrachten die Funktionen $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}^2$ gegeben durch

$$\begin{aligned} f(x,y) &:= (x^2, xy, xy^2) \\ g(u,v,w) &:= (\sin v, \cos(uvw)) \end{aligned}$$

Die Funktion $h := g \circ f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $h(x,y) = (\sin x^2, \cos(x^4y^3))$ ist differenzierbar und es gilt

$$Dh(x,y) = \begin{pmatrix} 2x\cos x^2 & 0 \\ -4x^3y^3\sin(x^4y^3) & -3x^4y^2\sin(x^4y^3) \end{pmatrix}. \tag{2.1}$$

Für die Ableitungen von f bzw. g gilt

$$Dg(u,v,w) = \begin{pmatrix} \cos u & 0 & 0 \\ -vw\sin(uvw) & -uw\sin(uvw) & -uv\sin(uvw) \end{pmatrix}, \; Df(x,y) = \begin{pmatrix} 2x & 0 \\ y & x \\ y^2 & 2xy \end{pmatrix}$$

und wir verifizieren, dass das Produkt dieser Matrizen an der Stelle $(u,v,w) = f(x,y)$ mit (2.1) übereinstimmt.

Aus der obigen Kettenregel können wir nun relativ einfach Ableitungsregeln für Summen und Produkte differenzierbarer Funktionen ableiten.

2.3 Korollar. *Es seien $U \subset \mathbb{R}^n$ offen und $f, g : U \to \mathbb{R}^m$ in $x_0 \in U$ differenzierbare Funktionen. Dann ist $\alpha f + \beta g$ für beliebige $\alpha, \beta \in \mathbb{R}$ in x_0 differenzierbar und es gilt*

$$D(\alpha f + \beta g)(x_0) = \alpha Df(x_0) + \beta Dg(x_0).$$

Beweis. Setzt man $F := \binom{f}{g} : U \to \mathbb{R}^m \times \mathbb{R}^m$ und $G : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m, G(u,v) := \alpha u + \beta v$, so ist G linear und somit differenzierbar mit Ableitung $DG(u,v) = G$ für alle $u, v \in \mathbb{R}^m$ und F ist ebenfalls in x_0 differenzierbar mit Ableitung $DF(x_0) := \binom{Df(x_0)}{Dg(x_0)}$. Die Kettenregel impliziert daher, dass $G \circ F$ gegeben durch $G \circ F(x) = \alpha f(x) + \beta g(x)$ in x_0 differenzierbar ist mit der Ableitung

$$D(G \circ F)(x_0) = DG(F(x_0)) \cdot DF(x_0) = G\begin{pmatrix} Df(x_0) \\ Dg(x_0) \end{pmatrix} = \alpha \cdot Df(x_0) + \beta \cdot Dg(x_0).$$

□

.

2.4 Korollar. (Produktregel). *Es seien $U \subset \mathbb{R}^n$ offen und $f, g : U \to \mathbb{R}$ in $x_0 \in U$ differenzierbare Funktionen. Dann ist $f \cdot g$ in x_0 differenzierbar und es gilt*

$$D(f \cdot g)(x_0) = f(x_0)Dg(x_0) + g(x_0)Df(x_0).$$

Beweis. Setzt man $F := \binom{f}{g} : U \to \mathbb{R} \times \mathbb{R}$ und $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ G(\alpha, \beta) := \alpha \cdot \beta$, so ist $(f \cdot g)(x) = (G \circ F)(x)$ und die Behauptung folgt aus der obigen Kettenregel.

□

2.5 Korollar. *Es seien $J \subset \mathbb{R}$ ein Intervall, $V \subset \mathbb{R}^m$ eine offene Menge und $\gamma = (\gamma_1, \dots, \gamma_m) : J \to V$ in $t_0 \in J$ und $f : V \to \mathbb{R}$ in $s_0 = \gamma(t_0)$ differenzierbare Funktionen. Dann ist $f \circ \gamma : J \to \mathbb{R}$ in t_0 differenzierbar und es gilt*

$$D(f \circ \gamma)(t_0) = (\mathrm{grad} f(\gamma(t_0)) | \gamma'(t_0)) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\gamma(t_0))\gamma_j'(t_0).$$

Für den Beweis verweisen wir auf die Übungen.

2.6 Bemerkung. Stetige Abbildungen $\gamma : J \to \mathbb{R}^m$ werden auch *Kurven* genannt. Letztere werden wir später noch genauer untersuchen. Es sei an dieser Stelle jedoch erwähnt, dass, fasst man $t \in J$ als Zeit und $\gamma(t) \in \mathbb{R}^m$ als Ort auf, γ die zeitliche Bewegung eines Punktes in $\mathbb{R}^m$ beschreibt. Jede Kurve $\gamma : J \to \mathbb{R}^m$ kann durch ein m-Tupel $\gamma = (\gamma_1, \dots, \gamma_m)$ beschrieben werden und für eine differenzierbare Kurve γ gilt

$$\gamma'(t_0) = (\gamma_1'(t_0), \dots, \gamma_m'(t_0))^T \in \mathcal{L}(\mathbb{R}, \mathbb{R}^m) = \mathbb{R}^m.$$

Der Vektor $\gamma'(t_0)$ heißt der *Tangentialvektor* der Kurve γ in t_0.

Mithilfe der obigen Formulierung der Kettenregel lassen sich die Begriffe Gradient und Niveaumenge geometrisch wie folgt interpretieren: es sei $f : U \to \mathbb{R}$ eine differenzierbare Funktion, $U \subset \mathbb{R}^n$ offen und $\gamma : J \to U$ eine differenzierbar Kurve definiert auf einem Intervall $J \subset \mathbb{R}$. Verläuft γ auf einer Niveaumenge von f, d.h. gilt $f(\gamma(t)) = c$ für alle $t \in J$ und eine Konstante $c \in \mathbb{R}$, so steht der Gradient von f im Punkte $\gamma(t)$ senkrecht auf dem Tangentialvektor $\gamma'(t)$, d.h. es gilt

$$\mathrm{grad} g(\gamma(t)) \perp \gamma'(t), \quad t \in J.$$

Gilt $U \subset \mathbb{R}^2$, so lässt sich der Graph von f als „Landschaft" über U mit $f(x)$ als „Höhe" über x interpretieren. Die Niveaumengen von f entsprechen dann den Höhenlinien des Graphen. Die obige Aussage besagt in diesem Bild also, dass der Gradient von f in x senkrecht auf der Höhenlinie durch x steht. Ferner zeigt $\mathrm{grad} f(x)$ in die Richtung des stärkesten Anstiegs von f und $-\mathrm{grad} f(x)$ in die Richtung des steilsten Abfalls.

2.7 Beispiel. Eine differenzierbare Funktion $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ heißt *positiv homogen vom Grade* $\alpha \in \mathbb{R}$, falls

$$f(tx) = t^\alpha f(x), \qquad \text{für alle } x \in \mathbb{R}^n \setminus \{0\},\ t > 0$$

gilt. Das obige Korollar 2.5 impliziert, dass

$$(\text{grad } g(x)|x) = \alpha g(x), \quad x \in \mathbb{R}^n \setminus \{0\}$$

gilt. Diese Beziehung wird auch *Eulersche Relation* genannt. Für den Beweis verweisen wir auf die Übungen.

Eine weitere Folgerung aus der Kettenregel ist der folgende Mittelwertsatz. Wie im Fall einer Variablen kann man hiermit die Differenz von Funktionswerten durch die Ableitung ausdrücken.

2.8 Satz. (Mittelwertsatz). *Es seien* $U \subset \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}$ *eine differenzierbare Funktion. Ferner seien* $a, b \in U$ *so, dass die Verbindungssgtrecke* $\overline{ab} := \{a + t(b-a), t \in [0,1]\}$ *ganz in* U *liegt. Dann existiert ein* $\xi \in \overline{ab}$ *mit*

$$f(b) - f(a) = f'(\xi)(b-a).$$

Beweis. Definiert man $g : [0,1] \to U$ durch $g(t) = a + t(b-a)$, so ist g differenzierbar mit $g'(t) = b - a$ für alle $t \in (0,1)$. Nach Korollar 2.5 ist $F = f \circ g : [0,1] \to \mathbb{R}$ ebenfalls differenzierbar mit $F'(t) = f'(g(t))(b-a)$ für alle $t \in (0,1)$. Nach dem klassischen Mittelsatz, Theorem IV.2.4 aus Analysis I, existiert ein $\tau \in (0,1)$ mit

$$f(b) - f(a) = F(1) - F(0) = F'(\tau) = f'(\xi)(b-a)$$

für $\xi := g(\tau) \in \overline{ab}$.

$\square$

.

In diesem Zusammenhang tritt in natürlicher Weise der Begriff der konvexen Menge auf. Eine Menge $U \subset \mathbb{R}^n$ heißt *konvex*, falls $\overline{ab} = \{a + t(b-a), t \in [0,1]\} \subset U$ gilt für alle $a, b \in U$.

Für konvexe Definitionsbereiche gilt die folgende Variante des Mittelwertsatzes.

2.9 Satz. (Schrankensatz). *Es sei $U \subset \mathbb{R}^n$ eine offene und konvexe Menge. Ist $f : U \to \mathbb{R}$ differenzierbar und existiert ein $L \geq 0$ mit $\|\mathrm{grad} f(x)\| \leq L$ für alle $x \in U$, so gilt*

$$|f(x) - f(y)| \leq L\|x-y\|, \quad x, y \in U,$$

d.h. f ist Lipschitz-stetig mit Lipschitzkonstante L.

Beweis. Nach dem Mittelwertsatz und der Cauchy-Schwarzschen Ungleichung gilt

$$|f(x) - f(y)| = |(\mathrm{grad} f(\xi)|(x-y))| \leq \|\mathrm{grad} f(\xi)\| \, \|x-y\| \leq L\|x-y\|$$

für ein geeignetes $\xi \in \overline{ab}$.

□

2.10 Korollar. *Es sei $U \subset \mathbb{R}^n$ eine offene Menge, so dass für je zwei Punkte $x, y \in U$ ein Streckenzug $x = z_0, z_1, \dots, z_l = y$ existiert mit $\overline{z_{k-1} z_k} \in U$ für alle $k = 1, \dots, l$. Dann ist eine differenzierbare Funktion $f : U \to \mathbb{R}$ genau dann konstant auf U, wenn* $\mathrm{grad} f(x) = 0$ *für alle $x \in U$ gilt.*

Für den Beweis verweisen wir auf die Übungen.

Wir beenden diesen Abschnitt mit einer nützlichen Variante des Mittelwertsatzes, die jedoch auf der stärkeren Voraussetzung der stetigen Differenzierbarkeit beruht.

2.11 Satz. (Mittelwertsatz in Integralform). *Es sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine stetig differenzierbare Funktion. Dann gilt*

$$f(y) - f(x) = \int_0^1 Df(x + t(y-x))(y-x)dt$$

für alle $x, y \in U$ *mit* $\overline{xy} \subset U$.

Beweis. Für $t \in [0,1]$ definieren wir $\varphi(t) := f(x + t(y-x))$. Dann ist φ stetig differenzierbar und der Hauptsatz der Differential- und Integralrechnung impliziert

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt = \int_0^1 Df(x + t(y-x))(y-x)dt.$$

$\square$

3 Höhere Ableitungen

Betrachtet man eine Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}$ mit partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$, so können diese wiederum partiell differenzierbar sein. Die Funktionen $\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})$ heißen *partielle Ableitungen 2. Ordnung* von f und werden oft auch als

$$\frac{\partial^2 f}{\partial x_j \partial x_i}$$

geschrieben.

Allgemeiner heißt eine Funktion $f : U \to \mathbb{R}$ $(\ell+1)$*-mal (stetig) partiell differenzierbar*, falls f ℓ-mal partiell differenzierbar ist und alle Ableitungen $\ell-$ter Ordnung (stetig) partiell differenzierbar sind. In den folgenden Untersuchungen wird der Vektorraum

$$C^k(U) := \{f : U \to \mathbb{R}, f \text{ ist } k - \text{mal stetig partiell differenzierbar}\}$$

eine wichtige Rolle spielen. Ist f k-mal stetig partiell differenzierbar, so schreiben wir auch

$$\frac{\partial}{\partial x_{i_k}} \dots \frac{\partial}{\partial x_{i_1}} f =: \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} =: f_{x_{i_n} \dots x_{i_1}} \text{ und } \partial_i \dots \partial_i f =: \frac{\partial^k}{\partial x_i^k}.$$

Im folgenden Beispiel berechnen wir die zweiten partiellen Ableitungen der Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) := x^2 \sin y$. Es gilt dann

$$\begin{array}{lclllcl} f_x(x,y) & = & 2x \sin y & & f_y(x,y) & = & x^2 \cos y \\ f_{xx}(x,y) & = & 2 \sin y & & f_{xy}(x,y) & = & 2x \cos y \\ f_{yx}(x,y) & = & 2x \cos y & & f_{yy}(x,y) & = & -x^2 \sin y \end{array}$$

und insbesondere ist $f_{xy} = f_{yx}$ in diesem Beispiel. Im Allgemeinen ist dies jedoch nicht der Fall wie das folgende Beispiel aufzeigt.

3.1 Beispiel. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) := \begin{cases} \frac{x^3 y - x y^3}{x^2 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Dann ist $f \in C^1(\mathbb{R}^2)$, die partiellen Ableitungen f_{xy} und f_{yx} existieren und sind stetig auf $\mathbb{R}^2 \backslash \{(0,0)\}$, aber es gilt

$$f_{xy}(0,0) = 1 \text{ und } f_{yx}(0,0) = -1.$$

Diese Tatsache verifizieren wir in den Übungen.

Es kann sogar die Situation eintreten, dass nur eine der beiden partiellen Ableitungen $\partial_{ij} f$ oder $\partial_{ji} f$ existieren. Der folgende Satz von H.A. SCHWARZ (1843-1921) besagt, dass derartiges nicht eintritt, wenn eine der beiden partiellen Ableitungen $\partial_{ij} f$ oder $\partial_{ji} f$ stetig ist.

3.2 Satz. (Satz von Schwarz). *Es sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ besitze für eine Wahl von $i, j \in \{1, \ldots, n\}$ in einer Umgebung von $x_0 \in U$ partielle Ableitungen $\partial_i f$, $\partial_j f$, $\partial_{ij} f$. Ist $\partial_{ij} f$ stetig in x_0, so existiert $\partial_{ji} f$ und es gilt*

$$\partial_{ij} f(x_0) = \partial_{ji} f(x_0).$$

Beweis. Wir wählen $\delta_i, \delta_j > 0$ so klein, dass $x_0 + se_i + te_j \in U$ für alle $(s,t) \in (-\delta_i, \delta_i) \times (-\delta_j, \delta_j) =: Q \subset \mathbb{R}^2$ gilt. Dann ist die Funktion

$$\varphi : Q \to \mathbb{R},\ \varphi(s,t) := f(x_0 + se_i + te_j)$$

wohl definiert und partiell differenzierbar. Ferner existieren die partiellen Ableitungen $\partial_1 \partial_2 \varphi$ und diese sind auch stetig in $(0,0)$.

Es ist zu zeigen, dass $\partial_2 \partial_1 \varphi(0,0)$ existiert und mit $\partial_1 \partial_2 \varphi(0,0)$ übereinstimmt. Nach Definition gilt

$$\begin{aligned} \partial_2 \partial_1 \varphi(0,0) &= [\frac{d}{dt}(\lim_{s \to 0} \frac{\varphi(s,t) - \varphi(0,t)}{s})](0) \\ &= \lim_{t \to 0} \lim_{s \to 0} \frac{1}{s} \frac{[\varphi(s,t) - \varphi(0,t)] - [\varphi(s,0) - \varphi(0,0)]}{t}. \end{aligned}$$

Wenden wir den Mittelwertsatz auf den Differenzenquotienten bzgl. der zweiten Variablen an, so erhalten wir

$$\begin{aligned} \frac{1}{s} \frac{[\varphi(s,t) - \varphi(0,t)] - [\varphi(s,0) - \varphi(0,0)]}{t} &= \frac{1}{s} \partial_2 [\varphi(s, \xi t) - \varphi(0, \xi t)] \\ &= \frac{1}{s} [\partial_2 \varphi(s, \xi t) - \partial_2 \varphi(0, \xi t)] \end{aligned}$$

für ein $\xi \in (0,1)$. Dies ist jedoch ein Differenzenquotient von $\partial_2 \varphi$ bezüglich der ersten Variable s. Nach Voraussetzung ist $\partial_2 \varphi$ bzgl. der ersten Variablen differenzierbar und nach dem Mittelwertsatz existiert ein $\eta \in (0,1)$ mit $\frac{1}{s}[\partial_2 \varphi(s, \xi t) - \partial_2 \varphi(0, \xi t)] = \partial_1 \partial_2 \varphi(\eta s, \xi t)$. Ferner ist $\partial_1 \partial_2 \varphi$ nach Voraussetzung stetig in $(0,0)$ und daher gilt

$$\partial_2 \partial_1 \varphi(0,0) = \lim_{t \to 0} \lim_{s \to 0} \partial_1 \partial_2 \varphi(\eta s, \xi t) = \partial_1 \partial_2 \varphi(0,0).$$

□

3.3 Korollar. *Ist $U \subset \mathbb{R}^n$ eine offene Menge und $f \in C^k(U)$ für ein $k \in \mathbb{N}$, so gilt*

$$\frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial f}{\partial x_{i_1}} = \frac{\partial}{\partial x_{\pi(k)}} \cdots \frac{\partial f}{\partial x_{\pi(1)}}$$

für jede Permutation $\pi : \{1, \ldots, k\} \to \{1, \ldots, k\}$.

Den Beweis via Induktion nach k überlassen wir dem Leser.

Im Folgenden wollen wir die Ableitungen k-mal stetig differenzierbarer Funktionen f in Verallgemeinerung des Differentials als symmetrische, k-fach lineare Abbildung

$$D^k f(x_0) : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$$

auffassen.

Wir beginnen mit dem Fall $k = 2$ und betrachten 2-mal stetig differenzierbare Funktionen, d.h. differenzierbare Funktionen f für welche Df stetig differenzierbar ist und definieren eine Bilinearform $a(u, v)$ für $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n$ durch

$$D^2 f(x_0)(u, v) := D_u(D_v f)(x_0), \quad x_0 \in \mathbb{R}^n.$$

Aufgrund von Satz 1.6 ist die Richtungsableitung $D_v f(x_0)$ von f in Richtung v gegeben durch $D_v f(x_0) = Df(x_0)v$. Ferner ist die Funktion $D_v f$ wiederum in x_0 differenzierbar, da dies für Df gilt. Nach Satz 1.6 besitzt daher $D_v f$ in x_0 Richtungsableitungen in Richtung u und es gilt

$$D_u(D_v f)(x_0) = D(D_v f(x_0)) \cdot u = \sum_{i,j=1}^{n} \partial_{ij} f(x_0) v_i u_j.$$

Die Abbildung

$$(u, v) \mapsto D_u D_v f(x_0)$$

ist linear in u und v und ferner nach dem obigen Satz von Schwarz auch symmetrisch. Sie wird als *Differential zweiter Ordnung* von f in x_0 bezeichnet. Bezüglich der kanonischen Basis des $\mathbb{R}^n$ ist ihre Matrixdarstellung gegeben durch

$$H_f(x_0) = \begin{pmatrix} \partial_{11} f(x_0) & \ldots & \partial_{1n} f(x_0) \\ \vdots & & \vdots \\ \partial_{n1} f(x_0) & \ldots & \partial_{nn} f(x_0) \end{pmatrix}.$$

Diese Matrix heißt *Hesse-Matrix*. Nach dem Satz von Schwarz ist sie eine symmetrische Matrix und es gilt

$$D^2 f(x_0)(u, v) = u^T \cdot H_f(x_0) \cdot v. \tag{3.1}$$

Wir werden im Anschluss an die Taylorapproximation noch genauer auf die geometrische Bedeutung der zweiten Ableitung eingehen.

Für beliebiges $k \in \mathbb{N}$ definieren wir $D^k f(x_0)$ analog zum Fall $k = 2$ als

$$D^k f(x_0)(v^1, \ldots, v^k) := D_{v^1} \ldots D_{v^k} f(x_0), \quad v^1, \ldots, v^k \in \mathbb{R}^n.$$

Diese Abbildung ist wiederum linear in den Variablen $v^1, \ldots, v^k$.

In diesem Zusammenhang erweist es sich als nützlich, den Begriff des *Multiindex* einzuführen. Darunter versteht man ein n-Tupel $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. Die natürliche Zahl

$$|\alpha| := \alpha_1 + \ldots + \alpha_n$$

heißt *Ordnung* von α. Ferner definiert man

$$\alpha! := \alpha_1!\alpha_2! \ldots \alpha_n!$$

und für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ setzt man

$$\begin{aligned} x^\alpha &:= x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} \text{ und} \\ D^\alpha f &:= D_1^{\alpha_1} D_2^{\alpha_2} \ldots D_n^{\alpha_n} f := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} f, \\ D^0 f &:= f. \end{aligned}$$

Ersetzt man in einem Polynom p von n Variablen $\xi_1, \ldots, \xi_n$ vom Grade $m \in \mathbb{N}$, d.h.

$$p(\xi) = \sum_{|\alpha| \le m} a_\alpha \xi^\alpha,$$

die Variablen ξ_i durch Ableitungsoperatoren ∂_i, so entsteht ein sogenannter *linearer Differentialoperatoren $P(D)$ der Form*

$$P(D) : C^m(\mathbb{R}^n) \to C(\mathbb{R}^n), \quad P(D) := \sum_{|\alpha| \le m} a_\alpha D^\alpha$$

mit Koeffizienten a_α.

3.4 Beispiele. a) Ein sehr wichtiges Beispiel eines Differentialoperators ist der *Laplace-Operator* definiert durch

$$\Delta := \partial_1^2 + \ldots + \partial_n^2.$$

Das zugehörige Polynom p ist in diesem Fall gegeben durch $p(\xi) = \xi_1^2 + \ldots + \xi_n^2$.
b) Für $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ betrachten wir das Polynom $p(\xi) = h_1\xi_1 + \ldots + h_n\xi_n$ und setzen

$$\nabla h := p(D) = h_1\partial_1 + \ldots + h_n\partial_n.$$

c) Für $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ und $\ell \in \mathbb{N}$ gilt

$$(a_1 + \ldots + a_n)^\ell = \sum_{|\alpha| = \ell} \frac{\ell! a^\alpha}{\alpha!}.$$

Den Induktionsbeweis überlassen wir dem Leser.
d) Für $h = (h_1, \ldots, h_n)$ und $\ell \in \mathbb{N}$ gilt

$$(\nabla h)^\ell = (h_1\partial_1 + \ldots + h_n\partial_n)^\ell = \ell! \sum_{|\alpha| = \ell} \frac{h^\alpha \partial^\alpha}{\alpha!}.$$

Den Laplace-Operator kann man auch als die Spur der Hesse-Matrix $H_f(x)$ einer zweimal stetig differenzierbaren Funktion identifizieren, d.h. es gilt

$$\text{spur } H_f(x) = \sum_{i=1}^{n} \partial_i^2 f(x) = \Delta f(x). \tag{3.2}$$

Aufgrund dieser Beziehung, können wir nun die Drehinvarianz der Spur einer Matrix auf die Drehinvarianz des Laplace-Operators übertragen. Genauer gesagt, gilt der folgende Satz.

3.5 Satz. (Drehinvarianz des Laplace-Operators). *Für jede Orthonormalbasis* $v_1, \ldots, v_n$ *des* $\mathbb{R}^n$ *gilt*

$$\Delta = \partial_{v_1}^2 + \ldots + \partial_{v_n}^2.$$

Beweis. Nach Gleichung (3.1) gilt

$$\partial_{v_i}\partial_{v_i} f(x) = v_i^T H_f(x) v_i = e_i^T \tilde{H} e_i,$$

mit $\tilde{H} = V^T H_f V$, $V = (v_1, \ldots, v_n)$ und den kanonischen Basisvektoren $e_1, \ldots, e_n$. Somit gilt

$$\sum_{i=1}^{n} \partial_{v_i}^2 f = \text{spur } \tilde{H}.$$

Da V nach Voraussetung orthogonal ist, besitzen die Matrizen H_f und $\tilde{H}$ diesselbe Spur und die Behauptung folgt aus Gleichung (3.2).

□

Der Laplace Operator tritt in vielen Differentialgleichungen der Analysis und der Physik auf. Beispielhaft erwähnen wir an dieser Stelle die folgenden Gleichungen:

1. Die Potentialgleichung

$$\Delta u = 0.$$

Sie beschreibt Diffusionsprozesse und tritt auch in der Wahrscheinlichkeitstheorie auf. Ihre Lösungen heißen *harmonische* Funktionen. In der Dimension 2 bilden diese den Ausgangspunkt der Funktionentheorie.

2. Die Wellengleichung

$$u_{tt} = c\Delta u$$

beschreibt die Auslenkung eines elastischen Körpers.

3. Die Wärmeleitungsgleichung

$$u_t = c\Delta u$$

beschreibt die Wärmeleitung in homogenen Medien.

4. Die Schrödingergleichung

$$u_t = i\Delta u$$

ist die zentrale Gleichung der Quantenmechanik.

Zum Abschluss dieses Abschnitts berechnen wir noch Δf für eine rotationssymmetrische Funktion. Es sei $F \in C^2(J)$, wobei $J \subset (0,\infty)$ ein Intervall ist. Setzt man $f(x) := F(\|x\|_2)$ und $r := \|x\|_2$, so gilt

$$\partial_i f(x) = F'(r)\frac{x_i}{r}$$

und

$$\partial_i^2 f(x) = F''(r)\frac{x_i^2}{r^2} + F'(r)(\frac{1}{r} - \frac{x_i^2}{r^3}).$$

Somit ergibt sich

$$\Delta f(x) = F''(r) + \frac{n-1}{r}F'(r)$$

und es gilt $\Delta f = 0$ genau dann wenn die Gleichung $F''(r)+\frac{n-1}{r}F'(r) = 0$ erfüllt ist. Wir können leicht nachprüfen, dass für $n > 2$ die Funktion F gegeben durch $F(r) = r^{2-n}$ eine Lösung dieser Gleichung ist. Somit ist die durch

$$N(x) := \frac{1}{\|x\|_2^{n-2}}$$

auf $\mathbb{R}^n\backslash\{0\}$ definierte Funktion eine Lösung der Potentialgleichung $\Delta f = 0$. Die Funktion N stimmt bis auf einen Skalierungsfaktor mit dem sogenannten *Newton-Potential* auf $\mathbb{R}^n\backslash\{0\}$ überein.