2 Partielle Ableitungen

Wir kommen nun zur Differentiation von Funktionen im $\mathbb{R}^n$. Um für diese Ableitungen zu definieren, ist die einfachste und vielfach beste Idee, alle Variablen bis auf x_j als konstant aufzufassen und die resultierende Funktion der einen Variablen x_j wie üblich zu differenzieren. Auf diese Weise ergeben sich für $j = 1, \ldots, n$ die n partiellen Ableitungen der Funktion. Im Folgenden bezeichnet $e_1, \ldots, e_n$ die Standardbasis des $\mathbb{R}^n$, also $e_j = (0, \ldots, 1, \ldots, 0)$ mit der 1 an der j-ten Stelle.

Definition 2.1 *Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Die partielle Ableitung von f nach x_j an der Stelle $x \in \Omega$ ist der Grenzwert (falls existent)*

$$\partial_j f(x) = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t}.$$

Andere Bezeichnungen sind $\dfrac{\partial f}{\partial x_j}(x)$ und $D_j f(x)$.

Die partielle Ableitung $\partial_j f(x)$ ist einfach die Ableitung der reellen Funktion $(-\delta, \delta) \to \mathbb{R}^m$, $t \mapsto f(x + te_j)$, an der Stelle $t = 0$, oder alternativ die Ableitung der reellen Funktion

$$(x_j - \delta, x_j + \delta) \to \mathbb{R}^m, \quad y \mapsto f(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$$

an der Stelle $y = x_j$. Der folgende Satz formuliert in diesem Zusammenhang wohlbekannte Ergebnisse der eindimensionalen Analysis.

Satz 2.1 (Ableitungsregeln) *Sei $\Omega \subset \mathbb{R}^n$ offen und $x \in \Omega$. Die Existenz der partiellen Ableitungen $\partial_j f(x)$ und $\partial_j g(x)$ sei vorausgesetzt. Dann gelten folgende Aussagen:*

(a) Linearität: *für $f, g : \Omega \to \mathbb{R}^m$ und $\alpha, \beta \in \mathbb{R}$ gilt*

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

(b) Komponentenweise Differentiation: *für $f : \Omega \to \mathbb{R}^m$ gilt, wenn eine der Seiten existiert,*

$$\partial_j f(x) = \sum_{i=1}^{m} \partial_j f_i(x)\, e_i.$$

(c) Produktregel: *für $f : \Omega \to \mathbb{R}^m$ und $g : \Omega \to \mathbb{R}$ gilt*

$$\partial_j(fg)(x) = (\partial_j f)(x) g(x) + f(x)(\partial_j g)(x).$$

(d) Quotientenregel: *für $f : \Omega \to \mathbb{R}^m$ und $g : \Omega \to \mathbb{R}$ mit $g(x) \neq 0$ gilt*

$$\partial_j \left(\frac{f}{g}\right)(x) = \frac{(\partial_j f)(x) g(x) - f(x)(\partial_j g)(x)}{g(x)^2}.$$

(e) Kettenregel: *ist $f : \Omega \to I \subset \mathbb{R}$ und $\varphi : I \to \mathbb{R}$ differenzierbar in $f(x)$, so folgt*

$$\partial_j(\varphi \circ f)(x) = \varphi'(f(x)) \partial_j f(x).$$

Beispiel 2.1 Wir betrachten die Euklidische Abstandsfunktionen vom Nullpunkt

$$r : \mathbb{R}^n \to \mathbb{R},\ r(x) = |x| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

In $x \neq 0$ existieren die partiellen Ableitungen, und zwar gilt mit der Kettenregel

$$\partial_j r(x) = \frac{2x_j}{2\sqrt{x_1^2 + \ldots + x_n^2}} = \frac{x_j}{r} \qquad \text{für } r = r(x) = |x|.$$

Im Nullpunkt sind die partiellen Ableitungen dagegen nicht definiert, denn $r(0+te_i) = |t|$ ist in $t = 0$ nicht differenzierbar. Die Funktion $\partial_j r$ ist in $x \neq 0$ ihrerseits partiell differenzierbar, und wir erhalten die zweiten partiellen Ableitungen

$$\partial_i(\partial_j r)(x) = \frac{(\partial_i x_j)r - x_j \partial_i r}{r^2} = \frac{1}{r}\left(\delta_{ij} - \frac{x_i x_j}{r^2}\right).$$

Ist $\varphi : (0,\infty) \to \mathbb{R}$ zweimal differenzierbar, so berechnen wir weiter für $f = \varphi \circ r : \mathbb{R}^n \to \mathbb{R}$

$$\begin{aligned}
\partial_j f(x) &= \varphi'(r)\partial_j r = \varphi'(r)\frac{x_j}{r},\\
\partial_i(\partial_j f)(x) &= \varphi''(r)\partial_i r \partial_j r + \varphi'(r)\partial_i(\partial_j r) = \varphi''(r)\frac{x_i x_j}{r^2} + \frac{\varphi'(r)}{r}\left(\delta_{ij} - \frac{x_i x_j}{r^2}\right).
\end{aligned}$$

Der Operator $\Delta = \sum_{i=1}^n \partial_i^2$ heißt *Laplaceoperator*, die Lösungen der Gleichung $\Delta f = 0$ heißen *harmonische Funktionen*. Wir können jetzt die rotationssymmetrischen harmonischen Funktionen ausrechnen, und zwar erhalten wir

$$0 \stackrel{!}{=} \Delta f(x) = \varphi''(r) + \frac{n-1}{r}\varphi'(r) = r^{1-n}\big(r^{n-1}\varphi'(r)\big)'.$$

Diese Gleichung hat die Lösungen, mit Integrationskonstanten $a, b \in \mathbb{R}$,

$$\varphi(r) = \begin{cases} a\,\dfrac{r^{2-n}}{2-n} + b & \text{für } n \geq 3\\ a\,\log r + b & \text{für } n = 2. \end{cases}$$

Für $n = 3$ ist f das Newtonsche Gravitationspotential.

Im vorangegangenen Beispiel traten zweite partielle Ableitungen auf. Ist für $f : \Omega \to \mathbb{R}^m$ die Ableitungsfunktion $\partial_j f : \Omega \to \mathbb{R}^m$ definiert und ihrerseits in $x \in \Omega$ nach x_i partiell differenzierbar, so setzen wir

$$\partial_i \partial_j f(x) := \partial_i(\partial_j f)(x) \qquad \left(\text{alternative Notation } D_{ij}^2 f(x) \text{ oder } \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right). \tag{2.1}$$

Entsprechend für Ableitungen beliebiger Ordnung: ist für $i_1, \ldots, i_k \in \{1, \ldots, n\}$ die Ableitungsfunktion $\partial_{i_1} \ldots \partial_{i_k} f : \Omega \to \mathbb{R}^n$ definiert und in $x \in \Omega$ nach x_i partiell differenzierbar, so setzen wir induktiv

$$\partial_i \partial_{i_1} \ldots \partial_{i_k} f(x) = \partial_i(\partial_{i_1} \ldots \partial_{i_k} f)(x). \tag{2.2}$$

Die folgenden Klassen von Funktionen spielen in der Analysis eine wichtige Rolle. Wir werden sehen, dass die Eigenschaft der Stetigkeit der partiellen Ableitungen in vielen Anwendungen wesentlich ist.

Definition 2.2 (C^k-Räume) *Sei $\Omega \subset \mathbb{R}^n$ offen und $k \in \mathbb{N}_0 \cup \{\infty\}$. Wir bezeichnen mit $C^k(\Omega, \mathbb{R}^m)$ die Menge aller k-mal stetig differenzierbaren Funktionen auf Ω mit Werten im $\mathbb{R}^m$, das heißt alle partiellen Ableitungen $\partial_{i_1} \ldots \partial_{i_j} f$ der Ordnung $j \leq k$ (bzw. $j < \infty$ im Fall $k = \infty$) sind definiert und stetig auf Ω. Im reellwertigen Fall, also $m = 1$, setzen wir $C^k(\Omega, \mathbb{R}) = C^k(\Omega)$.*

Wir wollen nun zeigen, dass die Operatoren ∂_i und ∂_j auf C^2-Funktionen vertauschen. Aus der Existenz der partiellen Ableitungen $\partial_i \partial_j f$ und $\partial_j \partial_i f$ allein folgt das nicht, wie das Beispiel

$$f(x,y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{für } (x,y) \neq (0,0) \\ 0 & \text{für } (x,y) = (0,0) \end{cases} \tag{2.3}$$

zeigt. Man kann berechnen

$$\partial_1 f(x,y) = \begin{cases} y\frac{x^2-y^2}{x^2+y^2} + xy\frac{4xy^2}{(x^2+y^2)^2} & \text{für } (x,y) \neq (0,0) \\ 0 & \text{für } (x,y) = (0,0) \end{cases}$$

$$\partial_2 f(x,y) = \begin{cases} x\frac{x^2-y^2}{x^2+y^2} + xy\frac{4x^2y}{(x^2+y^2)^2} & \text{für } (x,y) \neq (0,0) \\ 0 & \text{für } (x,y) = (0,0) \end{cases}$$

und $\partial_1 \partial_2 f(0,0) = 1$, aber $\partial_2 \partial_1 f(0,0) = -1$.

Satz 2.2 (Symmetrie der 2. Ableitung, H. A. Schwarz) *Sei $\Omega \subset \mathbb{R}^n$ offen. Ist $f \in C^2(\Omega)$, so vertauschen für $1 \leq i,j \leq n$ die Ableitungen nach x_i und x_j:*

$$\partial_i \partial_j f = \partial_j \partial_i f \quad \text{auf } \Omega.$$

Beweis: Nach Definition ist $\partial_j f(x)$ Grenzwert der Differenzenquotienten

$$\Delta_j^t f(x) = \frac{f(x + te_j) - f(x)}{t} \quad \text{für } t \to 0.$$

Für $\Delta_i^s(\Delta_j^t f)(x) = \frac{1}{s}\Big(\Delta_j^t f(x + se_i) - \Delta_j^t f(x)\Big)$ folgt

$$\lim_{s \to 0} \Big(\lim_{t \to 0} \Delta_i^s (\Delta_j^t f)(x) \Big) = \partial_i (\partial_j f)(x).$$

Das Problem besteht darin, die beiden Grenzwerte zu vertauschen. Der Differenzenquotient vertauscht immerhin mit der partiellen Ableitung:

$$\partial_i(\Delta_j^t f)(x) = \frac{1}{t}\big(\partial_i f(x + te_j) - \partial_i f(x)\big) = \Delta_j^t(\partial_i f)(x).$$

Wir verwenden nun den Mittelwertsatz der Differentialrechnung. Ist g nach x_i partiell differenzierbar, so gilt für ein $\alpha \in [0,1]$:

$$\Delta_i^s g(x) = \frac{g(x + se_i) - g(x)}{s} = \partial_i g(x + \alpha s e_i).$$

Wir wenden das an auf $g = \Delta_j^t f$ und auf $g = \partial_i f$, wobei $\alpha, \beta \in [0,1]$ von s,t abhängen:

$$\Delta_i^s(\Delta_j^t f)(x) = \partial_i(\Delta_j^t f)(x + \alpha s e_i) = \Delta_j^t(\partial_i f)(x + \alpha s e_i) = \partial_j(\partial_i f)(x + \alpha s e_i + \beta t e_j).$$

Da nach Voraussetzung $\partial_j \partial_i f$ stetig in x ist, folgt die Behauptung, indem wir erst $t \to 0$ und dann $s \to 0$ gehen lassen. $\square$

Bemerkung. Tatsächlich haben wir gezeigt: existieren $\partial_i f$, $\partial_j f$, $\partial_j \partial_i f$ und ist $\partial_j \partial_i f$ stetig in x, so existiert auch $\partial_i \partial_j f(x)$ und es gilt $\partial_i \partial_j f(x) = \partial_j \partial_i f(x)$.

Folgerung 2.1 *Für eine Funktion $f \in C^k(\Omega, \mathbb{R}^m)$ vertauschen die partiellen Ableitungen bis zur Ordnung k, das heißt für jede Permutation $\sigma \in S_k$ gilt*

$$\partial_{i_{\sigma(1)}} \ldots \partial_{i_{\sigma(k)}} f = \partial_{i_1} \ldots \partial_{i_k} f.$$

BEWEIS: Nach Satz 2.2 können benachbarte Operatoren ∂_i, ∂_j vertauscht werden. Die symmetrische Gruppe wird durch Vertauschungen erzeugt (siehe Lineare Algebra). $\square$

Der Begriff der partiellen Ableitung allein ist nicht geeignet, um die mehrdimensionale Differentialrechnung zu entwickeln. Ein entscheidender Mangel ist, dass aus der Existenz der partiellen Ableitungen $\partial_1 f, \ldots, \partial_n f$ in $x \in \Omega$ nicht die Stetigkeit von f im Punkt x folgt.

Beispiel 2.2 Sei $\Omega = \mathbb{R}^2$ und

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq 0 \\ 0 & (x,y) = (0,0). \end{cases}$$

Dann gilt $f(x,0) = 0 = f(0,y)$, insbesondere $\partial_1 f(0,0) = 0 = \partial_2 f(0,0)$. Aber für $c(t) = (t,t)$ gilt $f(c(t)) = 1/2$ für alle $t \neq 0$, das heißt f ist nicht stetig im Nullpunkt.

Insbesondere können wir im allgemeinen keine Kettenregel für $f \circ c$ formulieren, da die Funktion $f \circ c$ nicht einmal stetig sein muss. Die Definition der partiellen Ableitungen macht explizit von den Koordinaten auf $\mathbb{R}^n$ Gebrauch. Es wäre denkbar, dass sich ein besserer Ableitungsbegriff ergibt, wenn alle Richtungen gleichberechtigt betrachtet werden. Dies führt auf den Begriff der Richtungsableitung.

Definition 2.3 (Richtungsableitung) *Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Für $x \in \Omega$ und $v \in \mathbb{R}^n$ heißt der Grenzwert (falls existent)*

$$\partial_v f(x) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

Richtungsableitung von f an der Stelle x in Richtung v. Dies ist die gewöhnliche Ableitung der Funktion $t \mapsto f(x+tv)$ an der Stelle $t = 0$.

Beispiel 2.3 Die Richtungsableitung von $r(x) = |x|$ in $x \in \mathbb{R}^n \backslash \{0\}$ in Richtung $v \in \mathbb{R}^n$ ist

$$\partial_v r(x) = \frac{d}{dt} \sqrt{|x|^2 + 2t\langle x, v\rangle + |v|^2}\ |_{t=0} = \left\langle \frac{x}{|x|}, v \right\rangle.$$

Leider reicht aber selbst die Existenz aller Richtungsableitungen von f in $x \in \Omega$ nicht aus, damit f auch stetig im Punkt x ist.

Beispiel 2.4 Betrachte jetzt auf $\Omega = \mathbb{R}^2$ die Funktion

$$f(x,y) = \begin{cases} \dfrac{2xy^2}{x^2+y^4} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Dann existieren im Punkt (0,0) alle Richtungsableitungen, denn für $v = (a,b) \neq (0,0)$ ist

$$\partial_v f(0,0) = \lim_{t\to 0} \frac{2ab^2}{a^2+t^2b^4} = \begin{cases} 2b^2/a & \text{für } a \neq 0 \\ 0 & \text{für } a = 0. \end{cases}$$

Dennoch ist f im Nullpunkt unstetig, denn für $c(t) = (t^2, t)$ gilt $f(c(t)) = 1$ für alle $t \neq 0$.