3. Differenzieren in $\mathbb{R}^n$

Im folgenden seien $m, n \in \mathbb{N}$, $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}^m$. Die Normen in $\mathbb{R}^n$ sind alle äquivalent. Wir wählen meist die euklidische Norm $|x| := \|x\|_2 := \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ oder die Maximumnorm $|x|_\infty := \|x\|_\infty = \max_{1 \le i \le n} |x_i|$.
Die Ableitung von f kann nicht wie im eindimensionalen Fall definiert werden, da der Differenzenquotient $\frac{f(x+h)-f(x)}{h}$ für $h \in \mathbb{R}^n$ keinen Sinn macht. Aber nach Analysis 1 gilt für differenzierbares $f: \mathbb{R} \to \mathbb{R}$

$$f(x+h) = f(x) + f'(x)h + r(x,h) \quad \text{mit} \quad \lim_{h \to 0} \frac{r(x,h)}{|h|} = 0.$$

Dieser Ansatz geht auch im $\mathbb{R}^n$.
Wir verwenden folgende Schreibweise (Landau-Symbol "klein o"):

$$\varphi(x) = o(|x-x_0|) :\Longleftrightarrow \frac{\varphi(x)}{|x-x_0|} \to 0 \quad (U \ni x \to x_0).$$

Definition 3.1 a) Seien $x \in U$ und $V_x := \{h \in \mathbb{R}^n \mid x+h \in U\}$. f heißt <u>differenzierbar</u> an der Stelle x, falls es eine lineare Abbildung $A = A(x): \mathbb{R}^n \to \mathbb{R}^m$, $h \mapsto A(x)h$ und eine in 0 stetige Abbildung $r(x,\cdot): V_x \to \mathbb{R}^m$ mit $r(x,0) = 0$ gibt mit

$$f(x+h) = f(x) + A(x)h + r(x,h)|h| \qquad (h \in V_x)$$

Kurz: $f(x+h) = f(x) + A(x)h + o(|h|)$.

b) f heißt <u>differenzierbar</u> in U, falls f an jeder Stelle $x \in U$ differenzierbar ist.

c) Falls f in $x \in U$ differenzierbar ist, so heißt

$$Df(x) := f'(x) := A(x)$$

die <u>Ableitung</u> (totales Differential) von f an x.

Bemerkung 3.2. a) Die Abbildung $A(x): \mathbb{R}^n \to \mathbb{R}^m$, $h \mapsto A(x)h$, ist eindeutig bestimmt, d.h., $f'(x)$ ist wohldefiniert. Denn für zwei solche $A_1(x), A_2(x)$ ist

$$(A_1(x) - A_2(x))h + (r_1(x,h) - r_2(x,h))|h| = 0,$$

d.h., für $y \in \mathbb{R}^m \setminus \{0\}$, $t \in \mathbb{R} \setminus \{0\}$, ist

$$|(A_1(x)-A_2(x))\frac{y}{|y|}| = \frac{|(A_1(x)-A_2(x))ty|}{|ty|} = |r_1(x,ty) - r_2(x,ty)| \xrightarrow{t\searrow 0} 0$$

und damit $A_1(x) = A_2(x)$ als Abbildung auf $\mathbb{R}^n$ (x ist fest).

b) Da $A(x): \mathbb{R}^n \to \mathbb{R}^m$ linear ist, ist $A(x)$ stetig und kann als Matrix dargestellt werden.

c) Es gilt

f in x differenzierbar $\iff \forall i \in \{1,\dots,n\}: P_i f$ in x differenzierbar,

wobei $P_i f = f_i$, $f = (f_1,\dots,f_m)^T$. Falls f in x differenzierbar ist, so ist

$$f'(x) = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_m'(x) \end{pmatrix} \in \mathbb{R}^{m\times n} \qquad (f_i'(x) \in \mathbb{R}^{1\times n} \text{ Zeilenvektor}).$$

Denn aus

$$f(x+h) = f(x) + f'(x)h + o(|h|)$$

folgt für jede Komponente

$$f_i(x+h) = f_i(x) + f_i'(x)h + o(|h|).$$

Die Rückrichtung folgt durch Zusammensetzen von $(f_1,\dots,f_m)^T$ aus den Komponenten. $\Diamond$

<u>Beispiel 3.3</u> a) $f: U \to \mathbb{R}^m$, $x \mapsto c \in \mathbb{R}^m$. Dann: $f'(x) = 0$ ($m\times n$-Matrix), da $c = f(x+h) = f(x) + f'(x)h + r(x,h)|h| = c + 0\cdot h + 0$.

b) Seien $M: \mathbb{R}^n \to \mathbb{R}^m$ linear ($M \in \mathbb{R}^{m\times n}$) und $f(x) = Mx$. Dann ist $f'(x) = M$, denn

$$f(x+h) = M(x+h) = Mx + Mh = f(x) + Mh + 0.$$

Speziell: Für $f: \mathbb{R}^2 \to \mathbb{R}$, $\binom{x_1}{x_2} \mapsto x_1 + x_2$ ist $f'(x)h = (1,1)\binom{h_1}{h_2}$

c) $f: \mathbb{R}^2 \to \mathbb{R}$, $\binom{x_1}{x_2} \mapsto x_1 x_2$. Dann ist

$$f(x+h) = (x_1+h_1)(x_2+h_2) = x_1x_2 + x_2h_1 + x_1h_2 + h_1h_2$$
$$= f(x) + (x_2, x_1)\binom{h_1}{h_2} + h_1h_2.$$

Da

$$\frac{|h_1h_2|}{|h|} \le \frac{1}{|h|}\left(\frac{h_1^2}{2} + \frac{h_2^2}{2}\right) = \frac{1}{2}\frac{|h|^2}{|h|} = \frac{1}{2}|h| \to 0, \quad |h| \to 0,$$

gilt, folgt

$$f'(x) = (x_2, x_1) \in \mathbb{R}^{1\times 2}.$$

d) $f: \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \langle x, x\rangle = |x|^2$. Dann ist $f'(x) = 2\langle x, \cdot\rangle$. Denn:

$$f(x+h) = \langle x+h, x+h\rangle = \langle x,x\rangle + \langle h,x\rangle + \langle x,h\rangle + \langle h,h\rangle$$
$$= f(x) + 2\langle x,h\rangle + |h|^2 = f(x) + 2\langle x,h\rangle + o(|h|).$$

Wegen $2\langle x,h\rangle = 2x^T h$ ist $f'(x) = 2x^T$ (Zeilenvektor). ◁

<u>Satz 3.4.</u> Ist f in x differenzierbar, so ist f stetig.

<u>Definition 3.5.</u> a) Seien $x \in U$ und $v \in \mathbb{R}^n$ mit $|v| = 1$. Falls

$$(D_v f)(x) := \lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{f(x+tv) - f(x)}{t} \in \mathbb{R}^m$$

existiert, heißt f an der Stelle x <u>in Richtung v differenzierbar</u> und $(D_v f)(x)$ die <u>Richtungsableitung</u> in Richtung v an der Stelle x. Speziell für $v = e_i$ heißt $(D_{e_i} f)(x) \in \mathbb{R}^m$ die <u>i-te partielle Ableitung</u> oder die <u>partielle Ableitung nach x_i</u>.

b) Sei nun $m = 1$, d.h., $f: \mathbb{R}^n \to \mathbb{R}$. Dann schreibt man

$$\frac{\partial f}{\partial x_i}(x) := (D_{e_i} f)(x) \in \mathbb{R}.$$

Andere Schreibweisen sind $\frac{\partial}{\partial x_i} f(x)$, $\partial_{x_i} f(x)$, $\partial_i f(x)$.

f heißt <u>partiell differenzierbar</u>, falls f nach allen x_i partiell differenzierbar, falls f nach allen x_i partiell differenzierbar ist.

Für $f: \mathbb{R}^n \to \mathbb{R}$ partiell differenzierbar heißt

$$\operatorname{grad} f(x) := \nabla f(x) := \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right)^T$$

der <u>Gradient</u> von f an der Stelle x.

<u>Satz 3.6.</u> Sei f in $x \in U$ differenzierbar. Dann ist f in jeder Richtung $v \in \mathbb{R}^n$, $|v| = 1$, differenzierbar und es gilt $(D_v f)(x) = f'(x)v \in \mathbb{R}^m$.

<u>Satz 3.7.</u> Sei $f: \mathbb{R}^n \to \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist f in Richtung e_i $(i = 1, \dots, n)$ differenzierbar, und es gilt

$$(D_{e_i} f)(x) = f'(x) e_i \quad (i\text{-te Spalte der Matrix } f'(x)).$$

Für $f = (f_1, \dots, f_m)^T$ gilt

$$f'(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_m(x)}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} =: (Df)(x).$$

Die Matrix $(Df)(x)$ heißt Jacobi-Matrix von f an der Stelle x (andere Schreibweise $J_f(x)$). Insbesondere gilt für $f: \mathbb{R}^n \to \mathbb{R}$, f in x differenzierbar:

$$f'(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right) = \nabla f(x)^T.$$

Bemerkung 3.8. Aus der partiellen Differenzierbarkeit folgt nicht die Differenzierbarkeit, ja nicht einmal die Stetigkeit. ◊

Beispiel 3.9. Sei $f: \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto \begin{pmatrix} x_1^2 - x_2^2 \\ x_1 x_2 \end{pmatrix}$. Dann ist

$$(Df)(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{pmatrix}.$$ ◊

Zu $x \in U$ fest ist $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$, d.h. $f': U \to L(\mathbb{R}^n, \mathbb{R}^m)$. Man kann f' auch als Abbildung $f': U \times \mathbb{R}^n \to \mathbb{R}^m$, $(x,h) \mapsto f'(x)h$, sehen. Da $L(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{m \times n} \sim \mathbb{R}^{m \cdot n}$ wieder ein endlich-dimensionaler Raum ist, sind alle Normen auf $L(\mathbb{R}^n, \mathbb{R}^m)$ äquivalent (kanonische Norm = Operatornorm). Damit ist klar, wie Stetigkeit einer Funktion $U \to L(\mathbb{R}^n, \mathbb{R}^m)$ definiert ist. Eine Funktion $f: U \to \mathbb{R}^m$ heißt stetig differenzierbar, falls f in U differenzierbar ist und $f': U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig ist.

Wie üblich setzen wir $C^1(U, \mathbb{R}^m) = \{ f: U \to \mathbb{R}^m \mid f \text{ stetig differenzierbar} \}$.

Satz 3.10 (Kettenregel) Seien $f: \mathbb{R}^n \supset U \to \mathbb{R}^m$, $g: \mathbb{R}^m \supset V \to \mathbb{R}^k$ differenzierbar, und

$$\underbrace{(g \circ f)'(x)}_{\in \mathbb{R}^{k \times n}} = \underbrace{(g' \circ f)(x)}_{\in \mathbb{R}^{k \times m}} \cdot \underbrace{f'(x)}_{\in \mathbb{R}^{m \times n}} \in L(\mathbb{R}^n, \mathbb{R}^k).$$

Beispiel 3.11. Seien $I \subset \mathbb{R}$, $x, y, z: I \to \mathbb{R}$ differenzierbar an $a \in I$. Seien

$$f(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, \quad t \in I,$$

und $g: \mathbb{R}^3 \supset U \to \mathbb{R}$ differenzierbar in $b = f(a)$. Für $h := g \circ f: I \to \mathbb{R}$ gilt dann

$$\underbrace{h'(a)}_{\in \mathbb{R}} = \underbrace{g'(b)}_{\in \mathbb{R}^{1\times 3}} \underbrace{f'(a)}_{\in \mathbb{R}^{3\times 1}} = (\partial_1 g(b), \partial_2 g(b), \partial_3 g(b)) \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}$$

$$= \partial_1 g(b)\, x'(t) + \partial_2 g(b)\, y'(t) + \partial_3 g(b)\, z'(t).$$

In älterer Literatur wird oft nicht zwischen g und h unterschieden, und man schreibt

$$\frac{\partial g}{\partial t} = \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial t}$$

$\lightning$

Bemerkung 3.12. Es gelten die üblichen Regeln über die Ableitung von Summen / Produkten / Quotienten etc. Etwa für $f, g: \mathbb{R}^n \supset U \to \mathbb{R}$:

$$(f\cdot g)'(x) = \underbrace{g(x)}_{\in \mathbb{R}} \cdot \underbrace{f'(x)}_{\in \mathbb{R}^n} + \underbrace{f(x)}_{\in \mathbb{R}} \underbrace{g'(x)}_{\in \mathbb{R}^n} \quad \text{etc.} \qquad \text{(Zeilenvektoren).}$$

$\Diamond$

Bemerkung 3.13. (Interpretation des Gradienten). Seien $f: \mathbb{R}^n \supset U \to \mathbb{R}$ differenzierbar und $v \in \mathbb{R}^n$ mit $|v| = 1$. Nach Satz 3.6 ist

$$|(D_v f)(x)| = |f'(x)v| = |\langle \nabla f(x), v\rangle| \leq |\nabla f(x)|\,|v| = |\nabla f(x)|$$

Damit ist $|D_v f(x)|$ maximal, falls $v = \pm \nabla f(x) / |\nabla f(x)|$ (falls $|\nabla f(x)| \neq 0$). Der Gradient ist die Richtung des steilsten Anstiegs von f. Der negative Gradient zeigt in Richtung der Falllinie.

Eine Höhenlinie (Niveaulinie) ist gegeben durch

$$N_f(c) := \{x \in U : f(x) = c\}$$

für $c \in \mathbb{R}$. Eine Parametrisierung der Höhenlinie ist eine Abbildung $x: \mathbb{R} \to N_f(c)$. In diesem Fall gilt für alle $t \in \mathbb{R}$: $f(x(t)) = c$, d.h.,

$$0 = \frac{d}{dt} f(x(t)) = f'(x(t))\, x'(t) = \langle \nabla f(x(t)), x'(t)\rangle.$$

Somit stehen die Vektoren $\nabla f(x(t))$ und $x'(t)$ aufeinander senkrecht.

Eine Parametrisierung der Falllinie ist eine Abbildung $\beta: \mathbb{R} \to \mathbb{R}^n$ mit $\beta'(t) = \nabla f(\beta(t))$. Höhen- und Falllinien stehen aufeinander senkrecht.

Beispiel 3.14. $f: \mathbb{R}^2 \to \mathbb{R}$, $x \mapsto |x|^2 = x_1^2 + x_2^2$. Dann ist f in $\mathbb{R}^2$ differenzierbar und $\nabla f(x) = (2x_1, 2x_2)^T = 2x$. Die Falllinien sind Geraden durch den Ursprung, die Höhenlinien konzentrische Kreise. ◊

Definition 3.15. a) Sei $f: \mathbb{R}^n \to \mathbb{R}^n$ differenzierbar. Dann heißt

$$\operatorname{div} f: \mathbb{R}^n \to \mathbb{R},\ x \mapsto \operatorname{div} f(x) := \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) = \sum_{i=1}^n \partial_i f_i(x)$$

die Divergenz von f. Manchmal wird auch $\operatorname{div} f(x) = \nabla \cdot f(x)$ geschrieben.

b) Sei $f: \mathbb{R}^3 \to \mathbb{R}^3$ differenzierbar. Dann heißt $\operatorname{rot} f: \mathbb{R}^3 \to \mathbb{R}^3$

$$x \longmapsto \operatorname{rot} f(x) = \begin{pmatrix} \partial_2 f_3(x) - \partial_3 f_2(x) \\ \partial_3 f_1(x) - \partial_1 f_3(x) \\ \partial_1 f_2(x) - \partial_2 f_1(x) \end{pmatrix}$$

die Rotation von f. Manchmal schreibt man auch $\operatorname{rot} f(x) = \nabla \times f(x)$ (Kreuzprodukt).

c) Die Abbildung $\Delta: C^2(\mathbb{R}^n; \mathbb{R}) \to C(\mathbb{R}^n; \mathbb{R})$,

$$f \mapsto \Delta f := \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$$

heißt Laplaceoperator. Es gilt "$\Delta = \operatorname{div} \operatorname{grad}$". Für $m \in \mathbb{N}$ definiere $\Delta: C^2(\mathbb{R}^n; \mathbb{R}^m) \to C(\mathbb{R}^n; \mathbb{R}^m)$ durch

$$\begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = f \longmapsto \Delta f = \begin{pmatrix} \Delta f_1 \\ \vdots \\ \Delta f_m \end{pmatrix}.$$

Bemerkung 3.16 Sei f differenzierbar in $a \in U$. Dann gilt

$$f(a+h) = f(a) + f'(a)h + r(a,h)|h|$$
$$= f(a) + f'(a)h + R(a,h)\cdot h = f(x) + F(a,h)\cdot h$$

mit $\lim_{h \to 0} R(a,h) = 0$ bzw. $\lim_{h \to 0} F(a,h) = f'(a)$. Denn sei zunächst $f: U \to \mathbb{R}$ (d.h. $m = 1$). Dann ist $|h|^2 = h^T h$, d.h. für $h \neq 0$ ist $r(a,h)\cdot|h| = \underbrace{\left(\frac{r(a,h)}{|h|} \cdot h^T\right)}_{\text{Zeilenvektor}} \underbrace{h}_{\in \mathbb{R}^n}$, und man setzt $R(a,h) := \frac{r(a,h)}{|h|} h^T$

bzw. $F(a,h) = f'(a) + R(a,h)$. Für $f: U \to \mathbb{R}^m$, d.h., $r(a,h) = (r_1(a,h), \dots, r_m(a,h))^T$ führt man dies für jede Komponente durch und erhält die Matrix

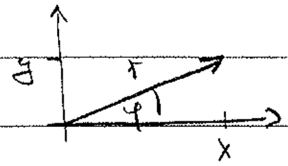
$$R(a,h) = \begin{pmatrix} \frac{r_1(a,h)}{|h|} h^T \\ \vdots \\ \frac{r_m(a,h)}{|h|} h^T \end{pmatrix} = \frac{r(a,h)}{|h|} h^T \in \mathbb{R}^{m \times n}.$$

◊

Satz 3.17. (Differenzierbarkeit der Umkehrabbildung). Seien $U \subset \mathbb{R}^n$ offen, $a \in U$ und $f: U \to \mathbb{R}^n$ injektiv sowie in a differenzierbar mit $\det f'(x) \neq 0$. Sei $b := f(a)$ ein innerer Punkt von $f(U)$, und sei $g := f^{-1}: f(U) \to \mathbb{R}^n$ stetig in b. Dann ist g in b differenzierbar, und $g'(b) = f'(a)^{-1}$.

Beispiel 3.18. (Polarkoordinaten) Definiere $f: (0,\infty) \times (0, 2\pi) \to \mathbb{R}^2$ durch

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \longmapsto f(r,\varphi) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$$



Dann ist f injektiv, und $g := f^{-1}$ ist gegeben durch $\begin{pmatrix} r \\ \varphi \end{pmatrix} = g(x,y)$ mit

$$r = \sqrt{x^2+y^2} \quad \text{und} \quad \varphi = \begin{cases} \arctan \frac{y}{x} & \text{falls } x,y > 0 \\ \pi/2 & \text{falls } x = 0, y > 0 \\ \pi + \arctan \frac{y}{x} & \text{falls } x < 0 \\ 3\pi/2 & \text{falls } x = 0, y < 0 \\ 2\pi + \arctan \frac{y}{x} & \text{falls } x > 0, y < 0 \end{cases}$$

Denn: $\varphi \in (0, \pi/2)$, $\tan\varphi = \frac{\sin\varphi}{\cos\varphi} = \frac{r\sin\varphi}{r\cos\varphi} = \frac{x}{y} \iff \varphi = \arctan\frac{x}{y}$

$\varphi \in (\pi/2, 3\pi/2)$: $\varphi' = \varphi - \pi \in (-\pi/2, \pi/2) \Rightarrow x/y = \tan\varphi = \frac{-x}{-y} = \tan\varphi'$

$$\Rightarrow \arctan\frac{x}{y} = \varphi - \pi$$

$\varphi \in (3\pi/2, 2\pi)$: $\varphi' = \varphi - 2\pi \in (-\pi/2, \pi/2)$.

Es gilt

$$f'(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \Rightarrow \det f'(r,\varphi) = r > 0$$

Damit ist g überall differenzierbar (außer in $(0,0)$) und

$$g'(x,y) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y/x^2}{\sqrt{x^2+y^2}} & \frac{1/x}{\sqrt{x^2+y^2}} \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{\sin\varphi}{r} & \frac{\cos\varphi}{r} \end{pmatrix} = f'(r,\varphi)^{-1}$$

◊

<u>Definition 3.19.</u> Seien $U, V \subset \mathbb{R}^n$ offen. Eine Funktion $f: U \to \mathbb{R}^n$ heißt Diffeomorphismus von U auf V, falls f bijektiv ist und $f \in C^1(U;\mathbb{R}^n)$, $f^{-1} \in C^1(V;\mathbb{R}^n)$.