7. Abbildungen von $\mathbb{R}^m$ nach $\mathbb{R}^n$: Differenzierbarkeit

"An element of surprise is also important. Beauty always has surprise."
Ingrid Daubechies

7.1 Def.: Sei $U \subset \mathbb{R}^m$ offen, $f: U \to \mathbb{R}^n$ eine Abbildung.

(i) f heißt differenzierbar in $\underline{x} \in U$, falls gilt:

Es gibt $A: \mathbb{R}^m \to \mathbb{R}^n$ linear und $\varphi: \mathbb{R}^m \to \mathbb{R}^n$ mit

$$f(\underline{x}+\underline{\xi}) = f(\underline{x}) + A\underline{\xi} + \varphi(\underline{\xi})$$

für alle ξ in einer Umgebung von 0 und mit

$$\lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{1}{\|\xi\|_2} \varphi(\xi) = 0.$$ [Es genügt, wenn φ in einer Umgebung von 0 definiert ist. (warum?)]

(ii) f heißt differenzierbar, falls f in jedem Punkt $\underline{x} \in U$ differenzierbar ist.

(iii) A ($= A(\underline{x})$) heißt Ableitung oder Differential von f in $\underline{x}$; wir schreiben auch $Df(\underline{x})$.

(Wir werden später sehen, dass A nicht von φ abhängt.)

Die zugehörige Matrix $(a_{ij}) \in M_{n \times m}(\mathbb{R})$ heißt auch Jacobimatrix oder Funktionalmatrix.

7.3 z.B.: $C \in M_{n \times n}(\mathbb{R})$ symmetrisch, d.h. $C = C^{tr}$.

Definiere $f : \mathbb{R}^n \to \mathbb{R}$ durch

$$f(\underline{x}) := \langle \underline{x}, C\underline{x} \rangle = \sum_{i,j=1}^{n} x_i c_{ij} x_j$$

(quadratische Form).

Für $\underline{x}, \underline{\xi} \in \mathbb{R}^n$ gilt

$$\begin{aligned} f(\underline{x} + \underline{\xi}) &= \langle \underline{x} + \underline{\xi}, C\underline{x} + C\underline{\xi} \rangle \\ &= \langle \underline{x}, C\underline{x} \rangle + \langle \underline{\xi}, C\underline{x} \rangle + \langle \underline{x}, C\underline{\xi} \rangle + \langle \underline{\xi}, C\underline{\xi} \rangle \\ &\overset{C \text{ symmetrisch}}{=} \langle \underline{x}, C\underline{x} \rangle + 2\langle C\underline{x}, \underline{\xi} \rangle + \langle \underline{\xi}, C\underline{\xi} \rangle \\ &= f(\underline{x}) + A(\underline{\xi}) + \varphi(\underline{\xi}) \end{aligned}$$

mit $A : \mathbb{R}^n \to \mathbb{R}$ gegeben durch $A(\underline{\xi}) := 2\langle C\underline{x}, \underline{\xi} \rangle$

und $\varphi : \mathbb{R}^n \to \mathbb{R}$ gegeben durch $\varphi(\underline{\xi}) := \langle \underline{\xi}, C\underline{\xi} \rangle$;

es gilt

$$\left| \frac{1}{\|\underline{\xi}\|_2} \cdot \varphi(\underline{\xi}) \right| \le \frac{1}{\|\underline{\xi}\|_2} \cdot \|C\| \cdot \|\underline{\xi}\|_2^2 = \|C\| \cdot \|\underline{\xi}\|_2 \xrightarrow{\underline{\xi} \to 0} 0,$$

also ist f differenzierbar.

7.4 **Satz:** Sei $U \subset \mathbb{R}^m$ offen, $f: U \to \mathbb{R}^n$ differenzierbar in $\underline{x} \in U$.

Dann gilt:

(i) f ist stetig in $\underline{x}$.

(ii) $$Df(\underline{x}) = \left(\frac{\partial f_i}{\partial x_j}(\underline{x}) \right)_{\substack{1 \le i \le n \\ 1 \le j \le m}} \in M_{n \times m}(\mathbb{R}).$$

Dabei benutzen wir die kanonische Identifizierung

$$L(\mathbb{R}^m, \mathbb{R}^n) \longleftrightarrow M_{n\times m}(\mathbb{R})$$

$$\left(\underline{\xi} \mapsto A\xi = \left(\sum_{j=1}^m a_{ij} \xi_j \right)_{1 \le i \le n} \right) \longleftarrow (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}}.$$

Bew.: (i) Es gilt $f(\underline{x} + \underline{\xi}) = f(\underline{x}) + Df(\underline{x})\underline{\xi} + \varphi(\underline{\xi})$ ($\underline{\xi}$ in einer Umgebung von $\underline{0}$),

für ein $\varphi: \mathbb{R}^m \to \mathbb{R}^n$ mit $\lim\limits_{\substack{\xi \to 0 \\ \xi \ne 0}} \frac{1}{\|\xi\|_2} \cdot \varphi(\xi) = 0$.

Wir erhalten $Df(\underline{x})\underline{\xi} \xrightarrow{\xi \to 0} 0$ und $\varphi(\underline{\xi}) \xrightarrow{\xi \to 0} 0$,

also $f(\underline{x} + \underline{\xi}) \xrightarrow{\xi \to 0} f(\underline{x}) + 0 + 0$.

(ii) Sei $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}}$ die zu $Df(\underline{x})$ gehörige Matrix,

d. h. $Df(\underline{x})\underline{\xi} = \left(\sum_{j=1}^m a_{ij} \xi_j \right)_{1 \le i \le n}$.

Sei φ wie oben, dann gilt

$f_i(\underline{x} + \underline{\xi}) = f_i(\underline{x}) + \sum_{j=1}^m a_{ij} \xi_j + \varphi_i(\underline{\xi})$, $i = 1, \dots, n$, also

$f_i(\underline{x} + h \cdot \underline{e}_j) = f_i(\underline{x}) + h a_{ij} + \varphi_i(h \cdot \underline{e}_j)$. Weiter gilt

$$\frac{\partial f_i}{\partial x_j}(\underline{x}) = \lim_{\substack{h \to 0 \\ h \ne 0}} \frac{f_i(\underline{x} + h \cdot \underline{e}_j) - f_i(\underline{x})}{h} = a_{ij} + \lim_{\substack{h \to 0 \\ h \ne 0}} \frac{\varphi_i(h \cdot \underline{e}_j)}{h} = a_{ij}.$$ □

7.5 Satz: Sei $U \subset \mathbb{R}^m$ offen, $f: U \to \mathbb{R}^n$ stetig partiell differenzierbar.
Dann ist f differenzierbar (und insbesondere stetig).

Bew.: Sei $\underline{x} \in U$. U offen $\Rightarrow \exists\, \delta > 0 : B_{\|\cdot\|_{\max}}(\underline{x}, \delta) \subset U$.

Für $\underline{\xi} \in \mathbb{R}^m$ mit $\|\underline{\xi}\|_{\max} < \delta$ setze

$$\underline{z}^{(i)} := \underline{x} + \sum_{\ell=1}^{i} \xi_\ell \cdot \underline{e}_\ell , \quad i = 0, \dots, m,$$

dann ist $\underline{z}^{(0)} = \underline{x}$ und $\underline{z}^{(m)} = \underline{x} + \underline{\xi}$.

Definiere $g^{(i)} : [0,1] \to \mathbb{R}$ durch

$$g^{(i)}(t) := f(\underline{z}^{(i-1)} + t \cdot \xi_i \cdot \underline{e}_i),$$

dann ist $g^{(i)}$ differenzierbar (warum?) mit

$$g^{(i)\prime}(t) = D_i f(\underline{z}^{(i-1)} + t \cdot \xi_i \cdot \underline{e}_i) \cdot \xi_i .$$

Nach MWS existiert $\theta_i \in [0,1]$ mit

$$\begin{array}{ccc} g^{(i)}(1) - g^{(i)}(0) & = (1-0) & g^{(i)\prime}(\theta_i) \\ \| & & \| \\ f(\underline{z}^{(i)}) - f(\underline{z}^{(i-1)}) & & D_i f(\underline{z}^{(i-1)} + \theta_i \cdot \xi_i \cdot \underline{e}_i) \cdot \xi_i . \end{array}$$

Wir erhalten $f(\underline{x} + \underline{\xi}) - f(\underline{x}) = \sum_{i=1}^m D_i f(\underline{z}^{(i-1)} + \theta_i \cdot \xi_i \cdot \underline{e}_i) \cdot \xi_i$,

also
$$f(\underline{x} + \underline{\xi}) = f(\underline{x}) + \underbrace{\sum_{i=1}^m D_i f(\underline{x}) \xi_i}_{=: A\underline{\xi}} + \underbrace{\sum_{i=1}^m \left(D_i f(\underline{z}^{(i-1)} + \theta_i \cdot \xi_i \cdot \underline{e}_i) - D_i f(\underline{x}) \right) \xi_i}_{=: \varphi(\underline{\xi})}$$

$$= f(\underline{x}) + A \underline{\xi} + \varphi(\underline{\xi}).$$

Dabei ist $A: \mathbb{R}^m \to \mathbb{R}$ linear und für φ gilt

$$\frac{1}{\|\underline{\xi}\|_2} \cdot |\varphi(\underline{\xi})| \le \frac{1}{\|\underline{\xi}\|_2} \cdot m \cdot \max_i \left| D_i f(\underline{z}^{(i-1)} + \theta_i \cdot \xi_i \cdot \underline{e}_i) - D_i f(\underline{x}) \right| \cdot \max_i |\xi_i|$$

$$\le \frac{1}{\|\underline{\xi}\|_2} \cdot m \cdot \max_i \left| D_i f(\underline{z}^{(i-1)} + \theta_i \cdot \xi_i \cdot \underline{e}_i) - D_i f(\underline{x}) \right| \cdot \|\underline{\xi}\|_2$$

$$= m \cdot \max_i \left| D_i f(\underbrace{\underline{z}^{(i-1)} + \theta_i \cdot \xi_i \cdot \underline{e}_i}_{\downarrow \, \underline{\xi} \to 0 \atop \underline{x}}) - D_i f(\underline{x}) \right|$$

$$\xrightarrow{\underline{\xi} \to 0} 0$$

wegen Stetigkeit von $D_i f$. $\square$

7.6 __Satz__ (Kettenregel): Seien $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ offen, $g: U \to V$, $f: V \to \mathbb{R}^k$ Abbildungen. g sei differenzierbar in $\underline{x} \in U$, f sei differenzierbar in $\underline{y} := g(\underline{x}) \in V$.
Dann ist $f \circ g: U \to \mathbb{R}^k$ differenzierbar in $\underline{x}$ mit

$$D(f \circ g)(\underline{x}) = Df(g(\underline{x})) \circ Dg(\underline{x}) \quad (: \mathbb{R}^m \to \mathbb{R}^k).$$

Bew.: Nach Voraussetzung existieren

$$\varphi : \mathbb{R}^m \to \mathbb{R}^n, \quad \psi : \mathbb{R}^n \to \mathbb{R}^k$$

mit

$$\lim_{\substack{\underline{\xi} \to 0 \\ \underline{\xi} \neq 0}} \frac{1}{\|\underline{\xi}\|_2} \cdot \varphi(\underline{\xi}) = 0, \quad (*) \qquad \lim_{\substack{\underline{\eta} \to 0 \\ \underline{\eta} \neq 0}} \frac{1}{\|\underline{\eta}\|_2} \cdot \psi(\underline{\eta}) = 0 \quad (**)$$

und

$$g(\underline{x} + \underline{\xi}) = g(\underline{x}) + Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi}),$$

$$f(\underline{y} + \underline{\eta}) = f(\underline{y}) + Df(\underline{y})\underline{\eta} + \psi(\underline{\eta}).$$

Definiere $\chi : \mathbb{R}^m \to \mathbb{R}^k$ durch

$$\chi(\underline{\xi}) := Df(\underline{y})\varphi(\underline{\xi}) + \psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})),$$

dann gilt

$$\begin{aligned}
f \circ g(\underline{x} + \underline{\xi}) &= f(\underbrace{g(\underline{x})}_{\underline{y}} + \underbrace{(g(\underline{x}+\underline{\xi}) - g(\underline{x}))}_{\underline{\eta}}) \\
&= f(g(\underline{x})) + Df(g(\underline{x}))(g(\underline{x}+\underline{\xi}) - g(\underline{x})) + \psi(g(\underline{x}+\underline{\xi}) - g(\underline{x})) \\
&= f(g(\underline{x})) + Df(g(\underline{x}))(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})) + \psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})) \\
&= f(g(\underline{x})) + Df(g(\underline{x}))Dg(\underline{x})\underline{\xi} + \chi(\underline{\xi}).
\end{aligned}$$

Wegen (*) existieren $K > 0$ und $\delta > 0$ mit

$\|\varphi(\underline{\xi})\|_2 \leq K \cdot \|\underline{\xi}\|_2$ für alle $\underline{\xi} \in B(0,\delta)$. (Warum?)

Dann gilt

$$\frac{1}{\|\underline{\xi}\|_2}\|\psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi}))\|_2 = \frac{\|Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})\|_2}{\|\underline{\xi}\|_2 \cdot \|Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})\|_2} \cdot \|\psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi}))\|_2$$

$$\leq \frac{\|Dg(\underline{x})\underline{\xi}\|_2 + \|\varphi(\underline{\xi})\|_2}{\|\underline{\xi}\|_2 \cdot \|Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})\|_2} \cdot \|\psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi}))\|_2$$

$$\leq (\|Dg(\underline{x})\| + K)\frac{\|\underline{\xi}\|_2}{\|\underline{\xi}\|_2} \cdot \underbrace{\frac{1}{\|Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi})\|_2}\underbrace{\|\psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi}))\|_2}_{\downarrow\, \underline{\xi}\to 0 \atop 0}}_{(**)\ \downarrow\, \underline{\xi}\to 0 \atop 0}$$

$$\xrightarrow{\underline{\xi}\to 0} 0.$$

Außerdem gilt

$$\frac{1}{\|\underline{\xi}\|_2}\|Df(g(\underline{x}))\varphi(\underline{\xi})\|_2 \leq \|Df(g(\underline{x}))\| \cdot \frac{1}{\|\underline{\xi}\|_2} \cdot \|\varphi(\underline{\xi})\|_2 \underset{*}{\xrightarrow{\underline{\xi}\to 0}} 0,$$

also

$$\frac{1}{\|\underline{\xi}\|_2} \cdot \|\chi(\underline{\xi})\|_2 \leq \frac{1}{\|\underline{\xi}\|_2}\left(\|\psi(Dg(\underline{x})\underline{\xi} + \varphi(\underline{\xi}))\|_2 + \|Df(g(\underline{x}))\varphi(\underline{\xi})\|_2\right) \xrightarrow{\underline{\xi}\to 0} 0.$$

$\square$

7.7 Z.B.: $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ offen, $g : U \to V$, $f : V \to \mathbb{R}$ differenzierbar.

Dann gilt $\dfrac{\partial (f \circ g)}{\partial x_i}(\underline{x}) = \sum_{j=1}^{n} \dfrac{\partial f}{\partial y_j}(g(\underline{x}))\, \dfrac{\partial g_j}{\partial x_i}(\underline{x})$ für $i \in \{1, \dots, m\}$,

denn

$$\left(\frac{\partial (f \circ g)}{\partial x_1}(\underline{x}), \dots, \frac{\partial (f \circ g)}{\partial x_m}(\underline{x}) \right) = D(f \circ g)(\underline{x}) \overset{7.6}{=} Df(g(\underline{x}))\, Dg(\underline{x})$$

$$= \left(\frac{\partial f}{\partial y_1}(g(\underline{x})), \dots, \frac{\partial f}{\partial y_n}(g(\underline{x})) \right) \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\underline{x}) & \cdots & \frac{\partial g_1}{\partial x_m}(\underline{x}) \\ \vdots & & \\ \frac{\partial g_n}{\partial x_1}(\underline{x}) & \cdots & \frac{\partial g_n}{\partial x_m}(\underline{x}) \end{pmatrix}.$$

7.8 Def.: Sei $U \subset \mathbb{R}^m$ offen, $f : U \to \mathbb{R}$, $\underline{x} \in U$, $\underline{v} \in \mathbb{R}^m$ mit $\|\underline{v}\| = 1$.

Falls $D_{\underline{v}} f(\underline{x}) := \lim\limits_{\substack{t \to 0 \\ t \neq 0}} \dfrac{f(\underline{x} + t\underline{v}) - f(\underline{x})}{t}$ existiert, so heißt

$D_{\underline{v}} f(\underline{x})$ Richtungsableitung von f in $\underline{x}$ in Richtung $\underline{v}$.

(Es gilt $D_{\underline{e}_i} f(\underline{x}) = D_i f(\underline{x})$ – warum?)

7.9 Satz: Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ stetig (partiell) differenzierbar.
Dann gilt für $\underline{x} \in U$ und $\underline{v} \in \mathbb{R}^n$ mit $\|\underline{v}\| = 1$

$$D_{\underline{v}} f(\underline{x}) = \langle \underline{v}, \operatorname{grad} f(\underline{x}) \rangle \left(= \|\operatorname{grad} f(\underline{x})\| \cos \theta \text{ falls } \operatorname{grad} f(\underline{x}) \neq 0\right)$$

Winkel zwischen $\underline{v}$ und $\operatorname{grad} f(\underline{x})$

Bew.: Definiere $g: \mathbb{R} \to \mathbb{R}^n$ durch $g(t) := \underline{x} + t \cdot \underline{v}$,
dann existiert $\varepsilon > 0$ mit $g((-\varepsilon, \varepsilon)) \subset U$ und
$f \circ g : (-\varepsilon, \varepsilon) \to \mathbb{R}$ ist definiert.
Es gilt

$$D_{\underline{v}} f(\underline{x}) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(\underline{x} + t \cdot \underline{v}) - f(\underline{x})}{t} = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{(f \circ g)(t) - (f \circ g)(0)}{t}$$

$$= D(f \circ g)(0) = Df(g(0)) \cdot Dg(0) = \langle \operatorname{grad} f(\underline{x}), \underline{v} \rangle. \quad \square$$

7.10 Def.: Sei $A: [a, b] \to M_{n \times m}(\mathbb{R})$ stetig.
Dann sind $A_{ij}: [a, b] \to \mathbb{R}$, $1 \leq i \leq n$, $1 \leq j \leq m$, stetig; wir setzen

$$\int_a^b A(t)\,dt := \left(\int_a^b A_{ij}(t)\,dt \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in M_{n \times m}(\mathbb{R}).$$

7.11 Satz: Sei $U \subset \mathbb{R}^m$ offen und $f: U \to \mathbb{R}^n$ stetig (partiell) differenzierbar. Sei $\underline{x} \in U$ und $\underline{\xi} \in \mathbb{R}^m$ so dass $\{\underline{x} + t\cdot\underline{\xi} \mid t \in [0,1]\} \subset U$.

Dann gilt

$$f(\underline{x}+\underline{\xi}) - f(\underline{x}) = \left(\int_0^1 Df(\underline{x}+t\cdot\underline{\xi})\,dt \right) \cdot \underline{\xi} \, ;$$

außerdem gilt

$$\|f(\underline{x}+\underline{\xi}) - f(\underline{x})\|_2 \le \sup\{\|Df(\underline{x}+t\cdot\underline{\xi})\| \mid t\in[0,1]\} \cdot \|\underline{\xi}\|_2 .$$

Bew.: Für $1 \le i \le n$ definieren $g_i : [0,1] \to \mathbb{R}$ durch

(Kettenregel)

$g_i(t) := f_i(\underline{x}+t\cdot\underline{\xi})$, dann gilt $g_i'(t) = Df_i(\underline{x}+t\cdot\underline{\xi})\underline{\xi}$ und

$$f_i(\underline{x}+\underline{\xi}) - f_i(\underline{x}) = g_i(1) - g_i(0) = \int_0^1 g_i'(t)\,dt$$

$$= \int_0^1 Df_i(\underline{x}+t\cdot\underline{\xi})\underline{\xi}\,dt$$

$$= \int_0^1 \left(\sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(\underline{x}+t\underline{\xi})\, \xi_j \right) dt$$

$$= \sum_{j=1}^m \left(\int_0^1 \frac{\partial f_i}{\partial x_j}(\underline{x}+t\cdot\underline{\xi})\,dt \right) \xi_j ,$$

also $f(\underline{x}+\underline{\xi}) - f(\underline{x}) = \left(\int_0^1 \frac{\partial f_i}{\partial x_j}(\underline{x}+t\cdot\underline{\xi})\,dt \right)_{\substack{1\le i\le n\\ 1\le j\le m}} \underline{\xi} = \left(\int_0^1 Df(\underline{x}+t\underline{\xi})\,dt \right) \underline{\xi}$.

Normabschätzung:

$$\|f(\underline{x}+\underline{\xi}) - f(\underline{x})\|_2 = \left\| \int_0^1 Df(\underline{x}+t\underline{\xi})\underline{\xi}\,dt \right\|_2$$

$$\le \int_0^1 \|Df(\underline{x}+t\cdot\underline{\xi})\underline{\xi}\|_2\,dt$$

$$\le \int_0^1 \|Df(\underline{x}+t\cdot\underline{\xi})\|\,\|\underline{\xi}\|_2\,dt$$

$$\le \sup_{t\in[0,1]} \|Df(\underline{x}+t\cdot\underline{\xi})\|\,\|\underline{\xi}\|_2$$

Bemerkung: $\left\| \int_0^1 \underline{v}(t)\,dt \right\|_2 \le \int_0^1 \|\underline{v}(t)\|_2\,dt$ (Übung)

$\square$

7.12 <u>Notation:</u> Für $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$ setzen wir

$$|\alpha| := \alpha_1 + \dots + \alpha_m,$$

$$\alpha! := \alpha_1! \cdot \dots \cdot \alpha_m!.$$

Für $\underline{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$ sei

$$\underline{x}^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_m^{\alpha_m}.$$

Für $f : U \to \mathbb{R}$ ($U \subset \mathbb{R}^m$ offen) stetig differenzierbar sei

$$D^\alpha f := D_1^{\alpha_1} \dots D_m^{\alpha_m} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_m^{\alpha_m}}.$$

7.13 <u>Prop.:</u> Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}$ k-mal stetig differenzierbar.
Sei $\underline{x} \in U$ und $\underline{\xi} \in \mathbb{R}^m$ so dass $\{\underline{x} + t\cdot\underline{\xi} \mid t \in [0,1]\} \subset U$.
Dann ist $g : [0,1] \to \mathbb{R}$, $g(t) := f(\underline{x} + t \cdot \underline{\xi})$,
k-mal stetig differenzierbar mit

$$\frac{d^k g}{dt^k}(t) = \sum_{\substack{\alpha \in \mathbb{N}^m, \\ |\alpha| = k}} \frac{k!}{\alpha!} D^\alpha f(\underline{x} + t \cdot \underline{\xi}) \underline{\xi}^\alpha.$$

Bew.: B.h.: $\displaystyle\frac{d^k g}{dt^k}(t) = \sum_{i_1, \dots, i_k \in \{1, \dots, m\}} D_{i_k} \dots D_{i_1} f(\underline{x} + t \cdot \underline{\xi}) \xi_{i_1} \dots \xi_{i_k}.$

IA: $\displaystyle\frac{d^1 g}{dt^1}(t) = \frac{dg}{dt}(t) = Df(\gamma(t))\,\underline{\xi} = \sum_{i=1}^m D_i f(\underline{x} + t\underline{\xi}) \xi_i.$

IS: Es gilt

$$\frac{d^{\bar{k}} g}{dt^{\bar{k}}}(t) = \sum_{i_1,\dots,i_{\bar{k}} \in \{1,\dots,m\}} D_{i_{\bar{k}}} \dots D_{i_1} f(\underline{x} + t\cdot\underline{\xi})\, \xi_{i_1} \dots \xi_{i_{\bar{k}}}$$

für ein $\bar{k} < k$. Dann gilt

$$\frac{d^{\bar{k}+1} g}{dt^{\bar{k}+1}}(t) = \frac{d}{dt}\left(\frac{d^{\bar{k}} g}{dt^{\bar{k}}}\right)(t)$$

$$= \sum_{i \in \{1,\dots,m\}} D_i \left(\sum_{i_1,\dots,i_{\bar{k}} \in \{1,\dots,m\}} D_{i_{\bar{k}}} \dots D_{i_1} f(\underline{x} + t\cdot\underline{\xi})\, \xi_{i_1} \dots \xi_{i_{\bar{k}}} \right) \xi_i$$

$$= \sum_{i_1,\dots,i_{\bar{k}+1} \in \{1,\dots,m\}} D_{i_{\bar{k}+1}} \dots D_{i_1} f(\underline{x} + t\cdot\underline{\xi})\, \xi_{i_1} \dots \xi_{i_{\bar{k}+1}}.$$

Induktion $\Rightarrow$ Beh.

Falls $(i_1,\dots,i_k) \in \{1,\dots,m\}^k$ den Index i α_i-mal enthält ($i = 1,\dots,m$), so gilt nach Satz 6.6

$$D_{i_k} \dots D_{i_1} f(\underline{x} + t\cdot\underline{\xi})\, \xi_{i_1} \dots \xi_{i_k} = D_1^{\alpha_1} \dots D_m^{\alpha_m} f(\underline{x} + t\cdot\underline{\xi})\, \xi_1^{\alpha_1} \dots \xi_m^{\alpha_m}.$$

[$= D^\alpha f(\underline{x} + t\cdot\underline{\xi})\, \underline{\xi}^\alpha$ mit $\alpha = (\alpha_1,\dots,\alpha_m)$; dabei ist $|\alpha| = \sum \alpha_i = k$.]

Weiter gilt

$$\#\{(i_1,\dots,i_k) \in \{1,\dots,m\}^k \mid \text{Index } i \text{ kommt } \alpha_i\text{-mal vor}, i = 1,\dots,m\} = \frac{k!}{\alpha_1! \dots \alpha_m!},$$

$\left[= \frac{k!}{\alpha!}\right]$

also

$$\sum_{(i_1,\dots,i_k) \in \{1,\dots,m\}^k} D_{i_k} \dots D_{i_1} f(\underline{x} + t\cdot\underline{\xi})\, \xi_{i_1} \dots \xi_{i_k}$$

$$= \sum_{\substack{(\alpha_1,\dots,\alpha_m) \in \mathbb{N}^m,\\ \alpha_1 + \dots + \alpha_m = k}} \frac{k!}{\alpha_1! \dots \alpha_m!} \cdot D_1^{\alpha_1} \dots D_m^{\alpha_m} f(\underline{x} + t\cdot\underline{\xi})\, \xi_1^{\alpha_1} \dots \xi_m^{\alpha_m}$$

$$= \sum_{\alpha \in \mathbb{N}^m, |\alpha| = k} \frac{k!}{\alpha!} D^\alpha f(\underline{x} + t\cdot\underline{\xi})\, \underline{\xi}^\alpha$$

□

7.14 **Satz** (Taylorformel): Sei $U \subset \mathbb{R}^n$ offen, $\underline{x} \in U$ und $\delta > 0$ mit $\overline{B}(\underline{x}, \delta) \subset U$. Sei $f: U \to \mathbb{R}$ $(k+1)$-mal stetig differenzierbar.

Dann existiert $R_{k+1}: \mathbb{R}^n \to \mathbb{R}$ mit

$$f(\underline{x} + \underline{\xi}) = \sum_{\substack{\alpha \in \mathbb{N}^n, \\ |\alpha| \le k}} \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha + R_{k+1}(\underline{\xi}) \text{ für } \underline{\xi} \in \mathbb{R}^n \text{ mit } \|\underline{\xi}\|_2 < \delta$$

und

$$\lim_{\substack{\underline{\xi} \to 0, \\ \underline{\xi} \neq 0}} \frac{R_{k+1}(\underline{\xi})}{\|\underline{\xi}\|^k} = 0.$$

[Hierfür genügt es, wenn f k-mal stetig differenzierbar ist.]

Für jedes $\underline{\xi} \in \mathbb{R}^n$ mit $\|\underline{\xi}\|_2 < \delta$ existiert $\Theta \in [0,1]$ mit

$$R_{k+1}(\underline{\xi}) = \sum_{\substack{\alpha \in \mathbb{N}^n, \\ |\alpha| = k+1}} \frac{D^\alpha f(\underline{x} + \Theta \underline{\xi})}{\alpha!} \underline{\xi}^\alpha. \quad (*)$$

Bew.: Sei $\underline{\xi} \in B(0, \delta)$.

Definiere $g: [0,1] \to \mathbb{R}$ durch $g(t) := f(\underline{x} + t \cdot \underline{\xi})$.

Nach Prop. 7.13 ist g $(k+1)$-mal stetig differenzierbar, nach Satz ~~1.2~~ 1.12 existiert $\Theta \in [0,1]$ mit

$$g(1) = g(0+1) \overset{1.12}{=} \sum_{l=0}^{k} \frac{1}{l!} \frac{d^l}{dt^l} g(0) \cdot 1^l + \frac{1}{(k+1)!} \frac{d^{k+1}}{dt^{k+1}} g(\Theta) \cdot 1^{k+1}$$

$$\overset{7.13}{=} \sum_{l=0}^{k} \sum_{\substack{\alpha \in \mathbb{N}^n, \\ |\alpha| = l}} \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha + \sum_{\substack{\alpha \in \mathbb{N}^n, \\ |\alpha| = k+1}} \frac{D^\alpha f(\underline{x} + \Theta \underline{\xi})}{\alpha!} \underline{\xi}^\alpha.$$

$\Rightarrow (*)$

Für $\alpha \in \mathbb{N}^m$ mit $|\alpha| = k+1$ ist $D^\alpha f = D_1^{\alpha_1} \dots D_m^{\alpha_m} f : U \to \mathbb{R}$ stetig, also beschränkt auf $\overline{B}(x, \delta) \subset U$ (kompakt). Es gilt

$$\frac{1}{\|\underline{\xi}\|_2^k} \cdot |\underline{\xi}^\alpha| = \frac{1}{\|\underline{\xi}\|_2^k} \cdot |\xi_1^{\alpha_1}| \cdot \ldots \cdot |\xi_m^{\alpha_m}| \le \frac{1}{\|\underline{\xi}\|_2^k} \|\underline{\xi}\|_2^{\alpha_1} \cdot \ldots \cdot \|\underline{\xi}\|_2^{\alpha_m} = \|\underline{\xi}\|_2 \xrightarrow{\xi \to 0} 0,$$

also

$$\frac{1}{\|\underline{\xi}\|^k} R_{k+1}(\underline{\xi}) \xrightarrow{\xi \to 0} 0. \qquad \square$$

7.15 Cor.: Sei $U \subset \mathbb{R}^m$ offen, $\underline{x} \in U$, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt

$$f(\underline{x} + \underline{\xi}) = f(\underline{x}) + \langle \operatorname{grad} f(\underline{x}), \underline{\xi} \rangle + \frac{1}{2} \langle \underline{\xi}, \operatorname{Hess} f(\underline{x}) \underline{\xi} \rangle + R_3(\underline{\xi}),$$

$\underline{\xi}$ in einer Umgebung von 0.

Dabei ist $\operatorname{Hess} f(\underline{x}) = \left(D_i D_j f(\underline{x}) \right)_{\substack{1 \le i \le m \\ 1 \le j \le m}} \in M_{m \times m}(\mathbb{R})$ die Hessematrix von f in $\underline{x}$.

Bew.: Nach 7.14 gilt

$$f(\underline{x} + \underline{\xi}) = \sum_{\substack{\alpha \in \mathbb{N}^m \\ |\alpha| \le 2}} \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha + R_3(\underline{\xi}).$$

$$|\alpha| = 0 \Rightarrow \alpha = (0, \dots, 0) \text{ und } \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha = \frac{D^{(0)} f(\underline{x})}{0!} \xi_1^0 \cdot \ldots \cdot \xi_m^0 = f(\underline{x}).$$

$|\alpha| = 1 \Rightarrow \alpha = (0, \dots, \underset{i}{1}, \dots, 0)$, $i = 1, \dots, m$, und

$$\sum_{\substack{\alpha \in \mathbb{N}^m \\ |\alpha|=1}} \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha = \sum_{i=1}^m \frac{D_i f(\underline{x})}{1!} \cdot \xi_i = \langle \operatorname{grad} f(\underline{x}), \underline{\xi} \rangle .$$

$|\alpha| = 2 \Rightarrow \alpha = (0, \dots, \underset{i}{2}, \dots, 0)$ oder $\alpha = (0, \dots, 0, \underset{i}{1}, 0, \dots, 0, \underset{j}{1}, 0, \dots, 0)$

für $1 \le i \le m$ $\qquad$ für $1 \le i < j \le m$

$\Downarrow \qquad \Downarrow$

$$\frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha = \frac{D_i^2 f(\underline{x}) \xi_i^2}{2} \qquad \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha = \frac{D_i D_j f(\underline{x})}{1} \xi_i \xi_j$$

$$\Rightarrow \sum_{\substack{\alpha \in \mathbb{N}^m \\ |\alpha|=2}} \frac{D^\alpha f(\underline{x})}{\alpha!} \underline{\xi}^\alpha = \sum_{i \in \{1, \dots, m\}} \frac{D_i^2 f(\underline{x}) \xi_i^2}{2} + \sum_{\substack{i,j \in \{1, \dots, m\} \\ i<j}} D_i D_j f(\underline{x}) \xi_i \xi_j$$

$$= \frac{1}{2} \sum_{\substack{i,j \in \{1, \dots, m\} \\ i=j}} D_i D_j f(\underline{x}) \xi_i \xi_j + \frac{1}{2} \sum_{\substack{i,j \in \{1, \dots, m\} \\ i \neq j}} D_i D_j f(\underline{x}) \xi_i \xi_j$$

$$= \frac{1}{2} \sum_{i,j \in \{1, \dots, m\}} D_i D_j f(\underline{x}) \xi_i \xi_j$$

$$= \frac{1}{2} \langle \underline{\xi}, \operatorname{Hess} f(\underline{x}) \underline{\xi} \rangle . \qquad \square$$

7.16 Def.: Sei $U \subset \mathbb{R}^m$, $f: U \to \mathbb{R}$, $\underline{x} \in U$.
f besitzt in $\underline{x}$ ein lokales Maximum (Minimum), falls $\underline{x} \in V \underset{\text{offen}}{\subseteq} U$ existiert mit $f(\underline{y}) \le f(\underline{x})$ $(f(\underline{y}) \ge f(\underline{x}))$ für alle $\underline{y} \in V$. Ein lokales Extremum ist ein lokales Max. oder Min. Das Max. (Min.) ist isoliert, falls $f(\underline{y}) < f(\underline{x})$ $(f(\underline{y}) > f(\underline{x}))$ für alle $\underline{x} \neq \underline{y} \in V$.

7.17 Prop.: Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ partiell differenzierbar. Falls f in $\underline{x} \in U$ ein lokales Extremum besitzt, so gilt $\operatorname{grad} f(\underline{x}) = 0$.

Bew.: Für $\varepsilon > 0$ klein genug und $i \in \{1, \dots, n\}$ definiere $g_i : (-\varepsilon, \varepsilon) \to \mathbb{R}$ durch $g_i(t) := f(\underline{x} + t \cdot \underline{e}_i)$; dann ist g_i differenzierbar und hat in 0 ein lokales Extremum, es gilt nach Ana I $0 = g_i'(0) = Df(\underline{x} + 0 \cdot \underline{e}_i)\, \underline{e}_i = D_i f(\underline{x})$, also $\operatorname{grad} f(\underline{x}) = (D_1 f(\underline{x}), \dots, D_n f(\underline{x})) = 0$. $\square$

7.18 Erinnerung: Sei $A \in M_{n \times n}(\mathbb{R})$ symmetrisch, dann existiert $T \in M_{n \times n}(\mathbb{R})$ unitär mit

$$T^{-1} A T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_1, \dots, \lambda_n \in \mathbb{R} \text{ Eigenwerte}$$

(zu Eigenvektoren $T\underline{e}_1, \dots, T\underline{e}_n$).

A heißt positiv (negativ) definit, falls alle λ_i echt positiv (negativ) sind. Es gilt dann $\langle \underline{\xi}, A\underline{\xi} \rangle > 0$ (< 0) für alle $\underline{\xi} \in \mathbb{R}^n$.

7.19 Satz: Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ zweimal stetig differenzierbar, $\underline{x} \in U$ mit $\operatorname{grad} f(\underline{x}) = 0$.

(i) $(\operatorname{Hess} f)(\underline{x})$ positiv definit $\Rightarrow$ f hat in $\underline{x}$ ein isoliertes Minimum

(ii) $(\operatorname{Hess} f)(\underline{x})$ negativ definit $\Rightarrow$ f hat in $\underline{x}$ ein isoliertes Maximum

(iii) Falls $(\operatorname{Hess} f)(\underline{x})$ echt positive und negative Eigenwerte besitzt, so besitzt f in $\underline{x}$ kein lokales Extremum.

Bew.: (i) Nach Corollar 7.15 gilt für $\underline{\xi}$ in einer Umgebung von 0

$$f(\underline{x}+\underline{\xi}) = f(\underline{x}) + \langle \underbrace{\operatorname{grad} f(\underline{x})}_{0}, \underline{\xi} \rangle + \frac{1}{2} \langle \underline{\xi}, \operatorname{Hess} f(\underline{x}) \underline{\xi} \rangle + R_3(\underline{\xi}).$$

Wegen $\lim\limits_{\substack{\underline{\xi} \to 0 \\ \underline{\xi} \neq 0}} \frac{R_3(\underline{\xi})}{\|\underline{\xi}\|_2^2} = 0$ existiert zu $\varepsilon > 0$ ein $\delta > 0$

mit: $\|\underline{\xi}\|_2 < \delta \Rightarrow |R_3(\underline{\xi})| < \varepsilon \cdot \|\underline{\xi}\|_2^2$.

$S^{n-1} := \{ \underline{\eta} \in \mathbb{R}^n \mid \|\underline{\eta}\|_2 = 1 \} \subset \mathbb{R}^n$ ist kompakt und

$\underline{\eta} \mapsto \langle \underline{\eta}, \operatorname{Hess} f(\underline{x}) \underline{\eta} \rangle$ ist stetig, nimmt also auf

S^{n-1} ihr Minimum an. Da $\langle \underline{\eta}, \operatorname{Hess} f(\underline{x}) \underline{\eta} \rangle > 0$

für $\underline{\eta} \in S^{n-1}$ (warum?), gilt

$$\alpha := \inf \{ \langle \underline{\eta}, \operatorname{Hess} f(\underline{x}) \underline{\eta} \rangle \mid \underline{\eta} \in S^{n-1} \} = \min \{ \langle \underline{\eta}, \operatorname{Hess} f(\underline{x}) \underline{\eta} \rangle \mid \underline{\eta} \in S^{n-1} \} > 0.$$

Für $\underline{\xi} \in \mathbb{R}^n$ gilt nun

$\langle \underline{\xi}, \operatorname{Hess} f(\underline{x}) \underline{\xi} \rangle \geq \alpha$ (klar für $\underline{\xi} = 0$; für $\underline{\xi} \neq 0$ beachte $\frac{1}{\|\underline{\xi}\|_2} \underline{\xi} \in S^{n-1}$).

Sei nun $\varepsilon := \frac{\alpha}{4}$ und wähle δ wie oben, so dass

$$|R_3(\underline{\xi})| < \frac{\alpha}{4} \|\underline{\xi}\|_2^2 \quad \text{für } \underline{\xi} \in B(0, \delta),$$

dann gilt

$$f(\underline{x}+\underline{\xi}) \geq f(\underline{x}) + \underbrace{\langle \operatorname{grad} f(\underline{x}), \underline{\xi} \rangle}_{0} + \frac{\alpha}{2} \|\underline{\xi}\|_2^2 - \frac{\alpha}{4} \|\underline{\xi}\|_2^2$$

$$\geq f(\underline{x}) + \frac{\alpha}{4} \|\underline{\xi}\|_2^2 > f(\underline{x}) \quad \text{falls } 0 < \|\underline{\xi}\|_2 < \delta.$$

$\Rightarrow$ isoliertes Minimum in $\underline{x}$. (ii) analog. (iii) Übung. □

7.20 z.B.: $f: \mathbb{R}^2 \to \mathbb{R}$, $f\begin{pmatrix} x \\ y \end{pmatrix} = c + x^2 + y^2$, $c \in \mathbb{R}$.

Taylorformel:

$$f\begin{pmatrix} \xi \\ \eta \end{pmatrix} = f\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \left\langle \nabla f\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \mathrm{Hess} f\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle + R_3\begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

$$\nabla f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} D_x f\begin{pmatrix} x \\ y \end{pmatrix} \\ D_y f\begin{pmatrix} x \\ y \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \qquad \nabla f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mathrm{Hess} f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} D_x D_x f\begin{pmatrix} x \\ y \end{pmatrix} & D_x D_y f\begin{pmatrix} x \\ y \end{pmatrix} \\ D_x D_y f\begin{pmatrix} x \\ y \end{pmatrix} & D_y D_y f\begin{pmatrix} x \\ y \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\mathrm{Hess} f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{positiv definit}$$

$$\Rightarrow f\begin{pmatrix} \xi \\ \eta \end{pmatrix} = c + 2\xi^2 + 2\eta^2 + R_3\begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

und f hat in $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ein isoliertes Minimum.