Humboldt-Universität zu Berlin
Institut für Mathematik
C. Wendl, M. Berghoff, A. Fauck,
C. Jurja, L. Kotan, M. Majchrzak

Analysis II*

SoSe 2019

Übungsblatt 5 (kommentiert)

Schriftliche Abgabe: Donnerstag 16. Mai 2019
Schreiben Sie jede Aufgabe bitte auf ein gesondertes Blatt, und schreiben Sie auf jedes Blatt ihren Namen, ihre Matrikelnummer und ihre Übungsgruppe (Wochentag + Übungsleiter + Zeit)

Aufgabe 5.1 (3 Punkte)
Sei $\mathcal{U} := (0,\infty) \times (0,\infty) \times (0,\infty) \subset \mathbb{R}^3$. Berechnen Sie die Jacobi-Matrix der Abbildung

$$f : \mathcal{U} \to \mathbb{R}^2 : (x,y,z) \mapsto \left(x\sqrt{z} + \sqrt{y}, \frac{z^2 y}{x} \right)$$

in einem beliebigen Punkt des Definitionsbereiches. Ist f überall differenzierbar?

Antwort:

$$J_f(x,y,z) = \begin{pmatrix} \sqrt{z} & \frac{1}{2\sqrt{y}} & \frac{x}{2\sqrt{z}} \\ -\frac{z^2 y}{x^2} & \frac{z^2}{x} & \frac{2zy}{x} \end{pmatrix}$$

Die partiellen Ableitungen in dieser Matrix sind alle stetige Funktionen auf $\mathcal{U}$, also ist f (laut einem Satz aus der Vorlesung) eine stetig differenzierbare Funktion auf $\mathcal{U}$.

Aufgabe 5.2 (2 + 2 Punkte)
Der Graph einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist die Teilmenge in $\mathbb{R}^{n+1}$ definiert durch

$$\Gamma_f := \left\{ (x_1, \ldots, x_n, z) \in \mathbb{R}^{n+1} \;\middle|\; z = f(x_1, \ldots, x_n) \right\}.$$

Wenn f differenzierbar ist, dann wird der *Tangentialraum* $T_p\Gamma_f$ an den Graphen Γ_f im Punkt $p := (a, f(a)) \in \Gamma_f$ definiert als die Menge der Vektoren $X \in \mathbb{R}^{n+1}$, die als $X = \gamma'(0)$ für beliebige differenzierbare Kurven $\gamma : (-\epsilon, \epsilon) \to \Gamma_f$ mit $\gamma(0) = p$ vorkommen. Beweisen Sie:

a) $T_p\Gamma_f = \left\{ (v,t) \in \mathbb{R}^{n+1} \;\middle|\; v \in \mathbb{R}^n,\ t = \langle \nabla f(a), v \rangle \right\}$.[1] Insbesondere ist $T_p\Gamma_f$ ein linearer Unterraum von $\mathbb{R}^{n+1}$.

Kommentar: Per Definition ist ein Vektor $(v,t) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ in $T_p\Gamma_f$ genau dann, wenn $(v,t) = \gamma'(0)$ für eine differenzierbare Kurve $\gamma : (-\epsilon,\epsilon) \to \Gamma_f$ mit $\gamma(0) = p = (a, f(a))$. Wir können γ dann als

$$\gamma(t) = (\beta(t), g(t)) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$$

für Funktionen $\beta : (-\epsilon,\epsilon) \to \mathbb{R}^n$ und $g : (-\epsilon,\epsilon) \to \mathbb{R}$ schreiben. Sei $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ die lineare Projektion $(v,t) \mapsto v$. Dann gilt

$$\beta(t) = \pi \circ \gamma(t),$$

und π ist als lineare Funktion differenzierbar, also folgt wegen der Kettenregel, dass β auch differenzierbar ist. Dass $\gamma(t)$ immer in Γ_f liegt, impliziert $g(t) = f(\beta(t))$,

[1] Das Symbol ∇f als Synonym für "grad f" kommt im Skript von Helga Baum nicht vor, ist jedoch eine Standardbezeichnung für den Gradienten einer Funktion und wird in dieser Vorlesung häufig verwendet.

also ist $g = f \circ \beta$ auch differenzierbar (da f und β differenzierbar sind), und laut Kettenregel und der Definition des Gradienten gilt

$$g'(0) = Df(\beta(0))\beta'(0) = \langle \nabla f(a), \beta'(0) \rangle.$$

Der Vektor (v, t) hat also die Form

$$(v, t) = (\beta'(0), g'(0)) = (\beta'(0), \langle \nabla f(a), \beta'(0) \rangle,$$

also gilt $v = \beta'(0)$ und $t = \langle \nabla f(a), v \rangle$. Umgekehrt kann man für einen gegebenen Vektor $(v, t) \in \mathbb{R}^n \times \mathbb{R}$ mit $t = \langle \nabla f(a), v \rangle$, $\beta(t) := a + tv$ und $g(t) := f(\beta(t))$ definieren, und damit eine Kurve $\gamma = (\beta, g) : (-\epsilon, \epsilon) \to \Gamma_f$ mit $\gamma'(0) = (v, t)$ finden, um zu beweisen, dass $(v, t) \in T_p\Gamma_f$.

b) Der Vektor $(\nabla f(a), -1) \in \mathbb{R}^{n+1}$ ist orthogonal zu $T_p\Gamma_f$.

Kommentar: Orthogonal heißt, für $V := (\nabla f(a), -1) \in \mathbb{R}^{n+1}$ und alle $X \in T_p\Gamma_f$ mit $p = (a, f(a))$ gilt

$$\langle V, X \rangle = 0,$$

wobei $\langle\ ,\ \rangle$ das euklidische Skalarprodukt auf $\mathbb{R}^{n+1}$ bezeichnet. Nach Teil a) ist lässt sich dies leicht direkt beweisen, denn wir haben $X = (v, t)$ für $v \in \mathbb{R}^n$ und $t = \langle \nabla f(a), v \rangle \in \mathbb{R}$, also

$$\langle V, X \rangle = \langle (\nabla f(a), -1), (v, \langle \nabla f(a), v \rangle) \rangle = \langle \nabla f(a), v \rangle + (-1)\langle \nabla f(a), v \rangle = 0.$$

Etwas eleganter geht es aber auch so: der Graph Γ_f ist eine Niveaufläche $g^{-1}(0)$ für die Funktion

$$g : \mathbb{R}^{n+1} \to \mathbb{R}, \qquad g(x, t) := f(x) - t \quad \text{für } (x, t) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}.$$

Laut einem Satz aus der Vorlesung ist der Gradient $\nabla g(a, t)$ für $(a, t) \in \Gamma_f$ immer orthogonal zu $T_{(a,t)}\Gamma_f$. Dieser Gradient ist

$$\begin{aligned}\nabla g(a, t) &= \left(\frac{\partial g}{\partial x_1}(a, t), \quad \dots, \quad \frac{\partial g}{\partial x_n}(a, t), \quad \frac{\partial g}{\partial t}(a, t) \right) = \left(\frac{\partial f}{\partial x_1}(a), \quad \dots, \quad \frac{\partial f}{\partial x_n}(a), \quad -1 \right) \\ &= (\nabla f(a), -1).\end{aligned}$$

Aufgabe 5.3 (2 + 2 Punkte)
Wir betrachten die Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, definiert durch $f(x, y) := |xy|$.

a) Bestimmen Sie alle Punkte $a \in \mathbb{R}^2$, in denen die Funktion f die partiellen Ableitungen $\frac{\partial f}{\partial x}(a)$ und $\frac{\partial f}{\partial y}(a)$ besitzt.

Antwort: $\partial_x f(a)$ existiert im Punkt $(0, 0)$ und in allen Punkten $a = (a_1, a_2)$ mit $a_1 \neq 0$, aber nicht in den Punkten $(0, a_2)$ mit $a_2 \neq 0$. Es gilt:

$$\frac{\partial f}{\partial x}(a_1, a_2) = \begin{cases} |a_2| & \text{wenn } a_1 > 0, \\ -|a_2| & \text{wenn } a_1 < 0, \\ 0 & \text{für } (a_1, a_2) = (0, 0). \end{cases}$$

Die Antwort für $\partial_y f(a_1, a_2)$ ist genau gleich, aber mit den Rollen von a_1 und a_2 vertauscht.

b) Bestimmen Sie alle Punkte $a \in \mathbb{R}^2$, in denen die Funktion differenzierbar ist.

Kommentar: In jedem Punkt (a_1, a_2) mit $a_1 \neq 0$ und $a_2 \neq 0$ existieren beide partielle Ableitungen $\partial_x f$ und $\partial_y f$ und sind stetig auf einer Umgebung von (a_1, a_2), also folgt, dass f in (a_1, a_2) differenzierbar ist.[2] *In den Punkten $(a_1, 0)$ mit $a_1 \neq 0$ und $(0, a_2)$ mit $a_2 \neq 0$ kann f nicht differenzierbar sein, weil die eine oder andere partielle Ableitung nicht existiert. Der Fall $a = (0,0)$ muss individuell geprüft werden; hier wird die Frage durch keinen Satz aus der Vorlesung beantwortet, denn $(0,0)$ hat keine Umgebung $\mathcal{U} \subset \mathbb{R}^2$, so dass für jedes $a \in \mathcal{U}$ sowohl $\partial_x f(a)$ als auch $\partial_y f(a)$ existiert. Aber wegen $\partial_x f(0,0) = \partial_y f(0,0) = 0$ können wir folgendes sagen: wenn f in $(0,0)$ differenzierbar ist, dann gilt $Df(0,0) = 0$. Differenzierbarkeit würde dann bedeuten, dass die Approximationsformel*

$$f(a) = f(0) + Df(0)a + o(\|a\|) = o(\|a\|)$$

gilt. Die Formel gilt tatsächlich, denn $|x| \leq \sqrt{x^2+y^2}$ und $|y| \leq \sqrt{x^2+y^2}$ implizieren für $a = (x,y) \in \mathbb{R}^2$,

$$f(a) = |x||y| \leq x^2 + y^2 = \|a\|^2,$$

und $\|a\|^2 = o(\|a\|)$.

Bemerkung: Die Aufgabe widerlegt die Vermutung, dass die Umkehrung von Satz 6.9 im Skript (über die Differenzierbarkeit von Funktionen mit stetigen partiellen Ableitungen) gilt. Eine Funktion kann auch in einem Punkt differenzierbar sein, der keine Umgebung hat, auf der die partiellen Ableitungen existieren.

Aufgabe 5.4 (3 + 3 Punkte)
Sei $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ein beliebiges Skalarprodukt auf $\mathbb{R}^n$ und $\phi : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung. Zeigen Sie, dass die folgenden Abbildungen differenzierbar sind und geben Sie jeweils das Differential in einem beliebigen Punkt des Definitionsbereiches an.

a) $f : \mathbb{R}^n \to \mathbb{R}$ gegeben durch $f(\mathbf{x}) := \langle \phi(\mathbf{x}), \mathbf{x} \rangle$.

b) $g : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ gegeben durch $g(\mathbf{u}, \mathbf{v}, \mathbf{w}) := \det \begin{pmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{pmatrix}$, wobei $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ als die Spalten einer 3×3-Matrix zu verstehen sind. Was ist insb. $Dg(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ für die Standard-Basisvektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in \mathbb{R}^3$?

Kommentar: Im Skript von Helga Baum gibt es einen Satz über Ableitungen von multilinearen Funktionen (Beispiel 3 auf Seite 186), den man für diese Aufgabe direkt anwenden könnte. In der Vorlesung wurde nicht genau dieser Satz sondern der folgende allgemeinere Satz präsentiert: Seien $E_1, \ldots, E_k$ und E normierte Vektorräume, $\mu : E_1 \times \ldots \times E_k \to E$ eine stetige multilineare Abbildung, und $f_i : \mathcal{U} \to E_i$ Funktionen, die in $a \in \mathcal{U} \subset \mathbb{R}^n$ differenzierbar sind. Dann ist $f : \mathcal{U} \to E$, definiert durch $f(x) := \mu(f_1(x), \ldots, f_k(x))$, auch differenzierbar in a, mit

$$\begin{aligned} Df(a)h = {} & \mu(Df_1(a)h, f_2(a), \ldots, f_k(a)) + \mu(f_1(a), Df_2(a)h, f_3(a), \ldots, f_k(a)) \\ & + \ldots + \mu(f_1(a), \ldots, f_{k-1}(a), Df_k(a)h). \end{aligned}$$

[2] Hier ist Satz 6.9 im Skript von Helga Baum anwendbar, aber nicht direkt, denn der Definitionsbereich von f wurde als $\mathbb{R}^2$ festgelegt, und die partiellen Ableitungen von f existieren nicht überall auf $\mathbb{R}^2$. Aber Differenzierbarkeit in einem Punkt $a \in \mathbb{R}^2$ hängt nur von einer Umgebung des Punktes ab, d.h. f ist genau dann in a differenzierbar, wenn es eine offene Umgebung $\mathcal{U} \subset \mathbb{R}^2$ von a gibt, so dass die eingeschränkte Funktion $f|_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}$ in a differenzierbar ist. Deswegen kann man eine Umgebung $\mathcal{U}$ wählen, auf der beide partielle Ableitungen existieren, und f mit $f|_{\mathcal{U}}$ ersetzen, um Satz 6.9 anzuwenden.

In Teil a) kann man das mit $E_1 = E_2 = \mathbb{R}^n$, $E = \mathbb{R}$, $\mu = \langle\ ,\ \rangle$, $f_1 = \phi : \mathbb{R}^n \to \mathbb{R}^n$ und $f_2 = \mathbb{1} : \mathbb{R}^n \to \mathbb{R}^n$ anwenden, und bekommt

$$Df(\mathbf{x})\mathbf{h} = \langle \phi(\mathbf{h}), \mathbf{x} \rangle + \langle \phi(\mathbf{x}), \mathbf{h} \rangle.$$

(Hier wird auch ein Satz aus der Vorlesung über Ableitungen linearer Funktionen verwendet: wenn $f : \mathbb{R}^n \to E$ linear ist, dann ist f in jedem Punkt $\mathbf{x} \in \mathbb{R}^n$ differenzierbar und es gilt $Df(\mathbf{x})\mathbf{h} = f(\mathbf{h})$. Anders gesagt, die Ableitung einer linearen Funktion $f : \mathbb{R}^n \to E$ ist die konstante Funktion $Df : \mathbb{R}^n \to \mathscr{L}(\mathbb{R}^n, E)$ mit wert $f \in \mathscr{L}(\mathbb{R}^n, E)$ in jedem Punkt.)

Bei Teil b) kann man $\mu = \det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ betrachten und die Funktionen $f_1, f_2, f_3 : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ als Projektionen wie z.B. $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \mapsto \mathbf{u}$ definieren. Dabei muss man diese Projektionen differenzieren—das ist nicht schwierig, weil sie lineare Funktionen sind. (Hier könnte man diesen Schritt umgehen, wenn man den Satz im Skript lieber direkt anwenden möchte.) Die Antwort im Allgemeinen ist

$$Dg(\mathbf{u}, \mathbf{v}, \mathbf{w})(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \det \begin{pmatrix} \mathbf{a} & \mathbf{v} & \mathbf{w} \end{pmatrix} + \det \begin{pmatrix} \mathbf{u} & \mathbf{b} & \mathbf{w} \end{pmatrix} + \det \begin{pmatrix} \mathbf{u} & \mathbf{v} & \mathbf{c} \end{pmatrix}.$$

Hier fällt der Fall $(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ auf, weil es als Spur der Matrix mit Spalten $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ interpretiert werden kann:

$$Dg(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathrm{Tr} \begin{pmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \end{pmatrix}.$$

Aufgabe 5.5 (2 + 2 + 4 Punkte)
Der Raum $\mathbb{R}^{n \times n}$ von reellen $n \times n$-Matrizen ist ein n^2-dimensionaler Vektorraum, den wir in dieser Aufgabe mit $\mathbb{R}^{n^2}$ identifizieren werden, damit wir differenzierbare Funktionen auf offenen Teilmengen in $\mathbb{R}^{n \times n}$ betrachten können. Eine wichtige offene Teilmenge[3] ist

$$\mathrm{GL}(n, \mathbb{R}) = \left\{ \mathbf{A} \in \mathbb{R}^{n \times n} \;\middle|\; \mathbf{A} \text{ ist invertierbar} \right\}.$$

In Aufgabe 2.2 wurde auf $\mathbb{R}^{n \times n}$ eine Norm definiert, womit $\mathbb{R}^{n \times n}$ ein Banachraum wird, und diese Norm erfüllt auch die Ungleichung $\|\mathbf{AB}\| \le \|\mathbf{A}\| \cdot \|\mathbf{B}\|$ für alle $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Beweisen Sie:

a) Für alle $\mathbf{A} \in \mathbb{R}^{n \times n}$ mit $\|\mathbf{A}\| < 1$ konvergiert die Reihe $\sum_{k=0}^\infty \mathbf{A}^k$ in $\mathbb{R}^{n \times n}$.

   Kommentar: wegen $\|\mathbf{A}^k\| \le \|\mathbf{A}\|^k$ und $\|\mathbf{A}\| < 1$ konvergiert die Reihe $\sum_{k=0}^\infty \|\mathbf{A}^k\|$, also folgt die Beahauptung von einem Satz aus der Vorlesung Analysis I über absolut konvergente Reihen in Banachräumen.

b) Für alle $\mathbf{A} \in \mathbb{R}^{n \times n}$ mit $\|\mathbf{A}\| < 1$ ist $\mathbb{1} + \mathbf{A} \in \mathbb{R}^{n \times n}$ invertierbar, und

$$(\mathbb{1} + \mathbf{A})^{-1} = \sum_{k=0}^\infty (-1)^k \mathbf{A}^k.$$

   Bemerkung: Dies verallgemeinert die Formel $\frac{1}{1+x} = \sum_{k=0}^\infty (-1)^k x^k$ für $|x| < 1$.

   Kommentar: man betrachte die Partialsummen $\mathbf{S}_N := \sum_{k=0}^N (-1)^k \mathbf{A}^k$ und beweist, dass $(\mathbb{1} + \mathbf{A})\mathbf{S}_N$ und $\mathbf{S}_N(\mathbb{1} + \mathbf{A})$ gegen $\mathbb{1}$ konvergieren, wenn $N \to \infty$.

[3] $\mathrm{GL}(n, \mathbb{R})$ ist das Urbild der offenen Teilmenge $\mathbb{R} \setminus \{0\} \subset \mathbb{R}$ unter der stetigen Abbildung $\det : \mathbb{R}^{n \times n} \to \mathbb{R}$, und ist daher offen.

c) Die Funktion $\iota : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^{n\times n}$, definiert durch $\iota(\mathbf{A}) = \mathbf{A}^{-1}$, ist differenzierbar, mit Ableitung $D\iota(\mathbf{A}) \in \mathscr{L}(\mathbb{R}^{n\times n}, \mathbb{R}^{n\times n})$ im Punkt $\mathbf{A} \in \mathbb{R}^{n\times n}$ gegeben durch

$$D\iota(\mathbf{A})\mathbf{H} = -\mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1}.$$

Bemerkung: Dies verallgemeinert die Formel $\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$.

Kommentar: zu zeigen ist die Approximationsformel

$$\iota(\mathbf{A} + \mathbf{H}) = \iota(\mathbf{A}) + D\iota(\mathbf{A})\mathbf{H} + o(\|\mathbf{H}\|)$$

für alle $\mathbf{A} \in \mathrm{GL}(n, \mathbb{R})$ *und alle* $\mathbf{H} \in \mathbb{R}^{n\times n}$ *hinreichend klein, so dass* $\mathbf{A} + \mathbf{H}$ *auch invertierbar ist. Mit* $D\iota(\mathbf{A})\mathbf{H} = -\mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1}$ *heißt das konkret,*

$$(\mathbf{A} + \mathbf{H})^{-1} - \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1} = o(\|\mathbf{H}\|). \tag{1}$$

Da $\mathbf{A}$ *invertierbar ist, können wir* $\mathbf{A} + \mathbf{H} = \mathbf{A}(\mathbb{1} + \mathbf{A}^{-1}\mathbf{H})$ *schreiben. Gilt* $\|\mathbf{H}\| < \frac{1}{\|\mathbf{A}^{-1}\|}$, *dann folgt* $\|\mathbf{A}^{-1}\mathbf{H}\| \le \|\mathbf{A}^{-1}\| \cdot \|\mathbf{H}\| < 1$, *also mit dem Resultat von Teil b),*

$$\begin{aligned}(\mathbf{A} + \mathbf{H})^{-1} &= \left[\mathbf{A}(\mathbb{1} + \mathbf{A}^{-1}\mathbf{H})\right]^{-1} = (\mathbb{1} + \mathbf{A}^{-1}\mathbf{H})^{-1}\mathbf{A}^{-1} = \sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{H})^k\mathbf{A}^{-1} \\ &= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1} + (\mathbf{A}^{-1}\mathbf{H})^2\sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{H})^k\mathbf{A}^{-1}.\end{aligned}$$

Die Norm der linken Seite von (1) *ist also*

$$\begin{aligned}\left\|(\mathbf{A}^{-1}\mathbf{H})^2\sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{H})^k\mathbf{A}^{-1}\right\| &\le \|\mathbf{A}^{-1}\|^2 \cdot \|\mathbf{H}\|^2 \cdot \left\|\sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{H})^k\mathbf{A}^{-1}\right\| \\ &=: \|\mathbf{H}\| \cdot R(\mathbf{H}),\end{aligned}$$

und es bleibt zu zeigen, dass die Funktion

$$R(\mathbf{H}) := \|\mathbf{H}\| \cdot \|\mathbf{A}^{-1}\|^2 \cdot \left\|\sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{H})^k\mathbf{A}^{-1}\right\|$$

gegen 0 *konvergiert, wenn* $\mathbf{H} \to 0$. *Das folgt, weil die Norm der Summierung beschränkt bleibt: für alle* $\mathbf{H}$ *mit* $\|\mathbf{H}\| < \frac{1}{2\|\mathbf{A}^{-1}\|}$ *gilt* $\|(\mathbf{A}^{-1}\mathbf{H})^k\| < 1/2^k$ *und deswegen*

$$\left\|\sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{H})^k\mathbf{A}^{-1}\right\| \le \left(\sum_{k=0}^{\infty}\frac{1}{2^k}\right) \cdot \|\mathbf{A}^{-1}\| < \infty.$$

Insgesamt: **25 Punkte**

Schriftliche Zusatzaufgabe 5.Z (3 Punkte)
Finden Sie eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit den folgenden Eigenschaften: f ist stetig, alle Richtungsableitungen von f im Punkt $(0, 0)$ existieren, $(\partial_x f(0, 0), \partial_y f(0, 0)) \in \mathbb{R}^2$ ist nicht null, aber dieser Vektor gibt *nicht* diejenige Richtung an, in der f im Punkt $(0, 0)$ am schnellsten wächst.[4]

[4] Die Botschaft dieser Aufgabe ist, dass der Vektor $(\partial_1 f, \dots, \partial_n f)$ keine besondere Bedeutung haben muss, wenn f nicht differenzierbar ist. Deswegen definiert man den Gradienten ∇f im Allgemeinen nur für *differenzierbare* Funktionen—die Existenz der partiellen Ableitungen allein ist nicht ausreichend.

Hinweis: Die Funktion in Aufgabe 4.3(c) könnte helfen, müsste aber modifiziert werden.

Kommentar: z.B. klappt

$$f(x,y) := \begin{cases} \epsilon x + \frac{x|y|}{\sqrt{x^2+y^2}} & \text{falls } (x,y) \neq (0,0), \\ 0 & \text{falls } (x,y) = (0,0), \end{cases}$$

wenn $\epsilon > 0$ hinreichend klein ist. Stetigkeit folgt leicht von Aufgabe 4.3(c), dann muss man nur die Richtungsableitungen ausrechnen und vergleichen mit dem vermeintlichen "Gradienten".

In der Vorlesung wurde $\nabla f(a)$ bzw. $\operatorname{grad} f(a)$ für eine Funktion $f : \mathcal{U} \to \mathbb{R}$ differenzierbar in $a \in \mathcal{U}$ als der eindeutige Vektor definiert, der $Df(a) = \langle \nabla f(a), \cdot \rangle$ erfüllt. Danach haben wir die Formel $\nabla f(a) = (\partial_1 f(a), \ldots, \partial_n f(a))$ ausgerechnet. Diese Aufgabe ist für diejenige Studierenden, die sich gefragt haben, warum das nicht in der umgekehrten Reihenfolge ging: da $\nabla f(a) := (\partial_1 f(a), \ldots, \partial_n f(a))$ auch für einige nichtdifferenzierbare Funktionen definiert werden könnte, wollte hiermit kommuniziert werden, warum das ohne Differenzierbarkeit nicht sinnvoll ist.

Die folgenden Aufgaben werden teilweise in den Übungen besprochen.

Aufgabe 5.A
Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ gegeben durch

$$f(x,y) := \begin{cases} (x^2+y^2) \sin \frac{1}{\sqrt{x^2+y^2}} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{für } (x,y) = (0,0). \end{cases}$$

Beweisen Sie, dass die partielle Ableitung $\frac{\partial f}{\partial x}$ in $(0,0)$ nicht stetig ist, aber f trotzdem in $(0,0)$ differenzierbar ist.

Kommentar: So wie Aufgabe 5.3 gibt dies einen Widerspruch zur Umkehrung des Satzes, dass Funktionen mit stetigen partiellen Ableitungen differenzierbar sind.

Aufgabe 5.B
Sei $f : \mathcal{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ eine Funktion, die auf $\mathcal{U}$ alle partiellen Ableitungen besitzt. Zeigen Sie: Sind alle partiellen Ableitungen von f auf $\mathcal{U}$ beschränkt, so ist f stetig.

Kommentar: Der Beweis dieses Resultats ist analog zum Beweis des Satzes aus der Vorlesung, dass Funktionen mit stetigen partiellen Ableitungen differenzierbar sind.