Technische Universität Chemnitz
Prof. Dr. R. Herzog, J. Blechschmidt

Chemnitz, 5. Dezember 2013

Optimierung mit partiellen Differentialgleichungen

Übung 6

Aufgabe 20: Newton-Differenzierbarkeit der Maximumsfunktion

Beweisen Sie die Aussage von Beispiel 9.15 (b) aus der Vorlesung: Zeigen Sie, dass

$$G(u)\,h = g_u \cdot h \quad \text{(punktweise Multiplikation)}$$

mit

$$g_u(x) = \begin{cases} 1, & \text{wo } u(x) > 0 \\ \delta, & \text{wo } u(x) = 0 \\ 0, & \text{wo } u(x) < 0 \end{cases}$$

eine Newton-Ableitung ($\delta \in \mathbb{R}$ beliebig) der Abbildung

$$L^q(\Omega) \ni u \mapsto F(u) := \max\{0, u\} \in L^p(\Omega)$$

(mit $1 \le p < q \le \infty$) auf ganz $L^q(\Omega)$ ist.

Beachte: Man braucht einen Sprung in den Normen („norm gap“), also $q > p$.

Lösung Aufgabe 20:

Es sei $1 \le p < q \le \infty$ und $s < \infty$, sodass

$$\frac{1}{q} + \frac{1}{s} = \frac{1}{p}.$$

Für $u, h \in L^q(\Omega)$ definieren wir das Restglied

$$\begin{aligned} r(u;h) &:= F(u+h) - F(u) - G(u+h)\,h \\ &= \max(0, u+h) - \max(0,u) - g_{u+h} \cdot h. \end{aligned}$$

Per Definition gilt

$$r(u;h)(x) = \begin{cases} (u+h)(x) - u(x) - \quad h(x) & \text{für } (u+h)(x) > 0, u(x) \geq 0 \\ (u+h)(x) \qquad\quad - \quad h(x) & \text{für } (u+h)(x) > 0, u(x) < 0 \\ \qquad\quad - u(x) - \delta\, h(x) & \text{für } (u+h)(x) = 0, u(x) \geq 0 \\ \qquad\qquad\quad - \delta\, h(x) & \text{für } (u+h)(x) = 0, u(x) < 0 \\ \qquad\quad - u(x) & \text{für } (u+h)(x) < 0, u(x) \geq 0 \\ \qquad\qquad 0 & \text{für } (u+h)(x) < 0, u(x) < 0 \end{cases}$$

$$= \begin{cases} \;\; u(x) & \text{für } (u+h)(x) > 0, u(x) < 0 \\ -u(x) & \text{für } (u+h)(x) < 0, u(x) > 0 \\ -u(x) - \delta\, h(x) & \text{für } (u+h)(x) = 0, u(x) > 0 \\ \qquad - \delta\, h(x) & \text{für } (u+h)(x) = 0, u(x) < 0 \\ 0 & \text{sonst} \end{cases}$$

Wir bezeichnen die Teilmengen von Ω in der Fallunterscheidung als I_1 bis I_4 (die Abhängigkeit dieser Mengen von h (und u) schleppen wir nicht mit). Nun schätzen wir

$$\|r(u;h)\|^p_{L^p(\Omega)} = \int_\Omega |r(u;h)(x)|^p \,\mathrm{d}x$$

auf diesen Teilmengen ab.

Teilmenge I_1:
Auf I_1 gilt

$$h(x) > -u(x) > 0$$

und

$$|r(u;h)(x)| = |u(x)| \leq |h(x)|.$$

Wir zeigen zunächst $\mu(I_1) \to 0$ für $\|h\|_{L^q(\Omega)} \to 0$.

Wir nehmen das Gegenteil an, also $\mu(I_1) \not\to 0$ für $\|h\|_{L^p(\Omega)} \to 0$. Das heißt, es gibt eine Folge h_n mit $\|h_n\|_{L^p(\Omega)} \to 0$ und $\mu(I_1) \geq \varepsilon > 0$.

Nun existiert ein $\kappa > 0$, sodass für alle meßbaren Teilmengen $E \subset \Omega$ mit $\mu(E) \geq \varepsilon$ und $u > 0$ auf E

$$\int_E |u|^q \,\mathrm{d}x \geq \kappa^q$$

gilt. Ein Beweis findet sich in der Diplomarbeit von Gerd Wachsmuth, Lemma 3.5, oder so:
Sei $E \subset \Omega$ wie oben. Wir führen die Mengen

$$A = \{x \in \Omega \,:\, |u(x)| > 0\}$$
$$A_n = \{x \in \Omega \,:\, \tfrac{1}{n-1} > |u(x)| \geq \tfrac{1}{n}\}$$

für $n = 1, 2, \dots$ mit $1/0 = \infty$. Dann ist $A = \bigcup A_n$. Da $\mu(A) \leq \mu(\Omega) < \infty$ existiert ein $N \in \mathbb{N}$ mit $\sum_{n=N}^{\infty} \mu(A_n) \leq \varepsilon/2$. Dann gilt $\mu(E \cap \bigcup_{n=1}^{N-1} A_n) \leq \varepsilon/2$, woraus die Existenz von κ folgt.

$$\int_E |u|^q \,\mathrm{d}x \geq \int_{E \cap \bigcup_{n=1}^{N-1} A_n} |u|^q \,\mathrm{d}x \geq \frac{\varepsilon}{2\,(N-1)^q} =: \kappa > 0$$

Insbesondere gilt für I_1

$$\int_{I_1} |h_n|^q \,\mathrm{d}x \geq \int_{I_1} |u|^q \,\mathrm{d}x \geq \kappa^q.$$

Das ist ein Widerspruch zu $h_n \to 0$ in $L^q(\Omega)$.

Nun haben wir

$$\left(\int_{I_1} |r(u;h)(x)|^p \,\mathrm{d}x \right)^{1/p} \leq \left(\int_{I_1} |h(x)|^p \,\mathrm{d}x \right)^{1/p} \leq \|h\|_{L^q(\Omega)} \, \|\chi_{I_1}\|_{L^s(\Omega)} = \|h\|_{L^q(\Omega)} \, \mu(I_1)^{1/s}.$$

Die Teilmenge I_2 kann analog behandelt werden.

Teilmenge $I_3 \cup I_4$:
Analog zu I_1 zeigt man $\mu(I_3 \cup I_4) \to 0$. Damit gilt

$$\left(\int_{I_3 \cup I_4} |r(u;h)(x)|^p \,\mathrm{d}x \right)^{1/p} \leq (1+\delta)\, \|h\|_{L^q(\Omega)} \, \mu(I_3 \cup I_4)^{1/s}.$$

Insgesamt gilt also

$$\begin{aligned} \|r(u;h)\|_{L^p(\Omega)} &\leq \left(\int_{I_1} |r(u;h)(x)|^p \right)^{1/p} + \left(\int_{I_2} |r(u;h)(x)|^p \right)^{1/p} + \left(\int_{I_3 \cup I_4} |r(u;h)(x)|^p \right)^{1/p} \\ &\leq \|h\|_{L^q(\Omega)} \underbrace{\left(\mu(I_1)^{1/s} + \mu(I_2)^{1/s} + (1 + |\delta|)\, \mu(I_3 \cup I_4)^{1/s} \right)}_{\to 0 \text{ für } \|h\|_{L^q(\Omega)} \to 0} \end{aligned}$$

und somit

$$\frac{\|r(u;h)\|_{L^p(\Omega)}}{\|h\|_{L^q(\Omega)}} \to 0 \quad \text{für } \|h\|_{L^q(\Omega)} \to 0.$$

Aufgabe 21: Zur Newton-Differenzierbarkeit

(a) Zeigen Sie, dass die Fréchet-Ableitung einer stetig Fréchet-differenzierbaren Funktion auch eine Newton-Ableitung ist.

(b) Beweisen Sie eine Kettenregel für die Newton-Differenzierbarkeit: Es seien U, V, W Banachräume, $u \in U$ und $U_0 \subset U$ eine Umgebung von u. Ferner seien $g : U \to V$ stetig Fréchet-differenzierbar mit Ableitung g' und $f : V \to W$ Newton-differenzierbar in einer Umgebung V_0 von $g(u)$ mit Ableitung F, die in V_0 beschränkt ist. Dann ist $f \circ g$ Newton-differenzierbar in einer Umgebung von u mit der Newton-Ableitung

$$H(u) = F(g(u)) \circ g'(u).$$

Lösung Aufgabe 21:

(a) Siehe Beweis von (b).

(b) Wird in Lemma 9.17 gebraucht.

Zuerst stellen wir fest, dass g Lipschitzstetig ist (die stetige Ableitung ist in einer Umgebung beschränkt). Damit gilt $\|g(u+h) - g(u)\|_V \leq L \|h\|_U$. Insbesondere folgt aus $\|h\|_U \to 0$ auch $\|g(u+h) - g(u)\|_V \to 0$ und damit (dies ist die Newton-differenzierbarkeit von f an der Stelle $g(u)$ in Richtung $g(u+h) - g(u)$)

$$\frac{\big\| f(g(u+h)) - f(g(u)) - F(g(u+h))\,(g(u+h) - g(u)) \big\|_W}{\|g(u+h) - g(u)\|_V} \to 0 \quad \text{für } \|h\|_U \to 0.$$

Die Lipschitzstetigkeit von g impliziert nun

$$\frac{\big\| f(g(u+h)) - f(g(u)) - F(g(u+h))\,(g(u+h) - g(u)) \big\|_W}{\|h\|_U} \to 0 \quad \text{für } \|h\|_U \to 0. \tag{$*$}$$

Aus der stetigen Fréchetdifferenzierbarkeit von g folgt

$$g(u+h) - g(u) = \int_0^1 g'(u + \theta\, h) h \,\mathrm{d}\theta,$$

und aus der Stetigkeit von $g' : U \to \mathcal{L}(U, V)$ folgt

$$\left\| \int_0^1 g'(u + \theta\, h)\,\mathrm{d}\theta - g'(u+h) \right\|_{\mathcal{L}(U,V)} \to 0 \quad \text{für } \|h\|_U \to 0.$$

Damit haben wir die Newton-differenzierbarkeit von g

$$\frac{\|g(u+h) - g(u) - g'(u+h)\,h\|_V}{\|h\|_U} \to 0 \quad \text{für } \|h\|_U \to 0.$$

Da nun F in einer Umgebung von $g(u)$ beschränkt ist, folgt

$$\frac{\big\| F(g(u+h))\big(g(u+h) - g(u) - g'(u+h)\,h\big) \big\|_W}{\|h\|_U} \to 0 \quad \text{für } \|h\|_U \to 0.$$

Dies setzt man nun in (**??**) ein und erhält

$$\frac{\big\| f(g(u+h)) - f(g(u)) - F(g(u+h))\, g'(u+h)\,h \big\|_W}{\|h\|_U} \to 0 \quad \text{für } \|h\|_U \to 0.$$

Aufgabe 22: Voraussetzungen des Newton-Verfahrens

Zeigen Sie, dass das verallgemeinerte Newton-Verfahren angewendet auf das KKT-System (9.34) der Aufgabe (2.4), also

$$F(u,\mu) := \begin{bmatrix} \nabla f(u) + \mu \\ \mu - \min\{0, \mu + c\,(u - u_a)\} - \max\{0, \mu + c\,(u - u_b)\} \end{bmatrix} = 0 \quad \text{f. ü. in } \Omega_{\text{ctrl}}$$

mit $u, \mu \in L^2(\Omega_{\text{ctrl}})$ und beliebigem $c > 0$, *nicht* die Voraussetzungen von Satz 9.16 erfüllt. Betrachten Sie dabei die Fälle, dass

(a) F als Abbildung $F : L^2(\Omega_{\text{ctrl}}) \times L^2(\Omega_{\text{ctrl}}) \to L^2(\Omega_{\text{ctrl}}) \times L^2(\Omega_{\text{ctrl}})$ definiert wird,

(b) F als Abbildung $F : L^2(\Omega_{\text{ctrl}}) \times L^2(\Omega_{\text{ctrl}}) \to L^2(\Omega_{\text{ctrl}}) \times L^p(\Omega_{\text{ctrl}})$ definiert wird mit $1 \le p < 2$.

Lösung Aufgabe 22:

(a) Hier liegt das Problem darin, dass die zweite Komponente nicht die notwendige Regularität besitzt. Genauer heißt das, dass der notwendige Norm-Gap für die Differenzierbarkeit von $\max\{0, \ldots\}$ fehlt.

(b) Hier können wir zwar die Newton-Ableitung

$$G(u,\mu)\,(\delta u, \delta\mu) = \begin{bmatrix} A\,\delta u + \delta\mu \\ \chi_{\mathcal{I}(u,\mu)}\,\delta\mu - c\,\chi_{\mathcal{A}^-(u,\mu)}\,\delta u - c\,\chi_{\mathcal{A}^+(u,\mu)}\,\delta u \end{bmatrix} \in L^2(\Omega_{\text{ctrl}}) \times L^p(\Omega_{\text{ctrl}})$$

angeben, allerdings erhalten wir noch nicht einmal die Existenz der Inversen, da $G(u,\mu)$ nicht surjektiv ist. Dabei sind die aktiven Mengen wie in der Vorlesung durch

$$\begin{aligned} \mathcal{A}^-(u,\mu) &:= \big\{x \in \Omega_{\text{ctrl}} : \mu + c\,(u - u_a) < 0\big\} \\ \mathcal{A}^+(u,\mu) &:= \big\{x \in \Omega_{\text{ctrl}} : \mu + c\,(u - u_b) > 0\big\} \\ \mathcal{A}(u,\mu) &:= \mathcal{A}^+(u,\mu) \cup \mathcal{A}^-(u,\mu) \\ \mathcal{I}(u,\mu) &:= \Omega_{\text{ctrl}} \setminus \mathcal{A}(u,\mu). \end{aligned}$$

gegeben.

Aufgabe 23: Transfer der punktweisen in koeffizientenweise Beschränkungen

Es seien $U_h^{(0)}$ der Raum der stückweise konstanten Funktionen und $U_h^{(1)}$ der Raum der stückweise linearen Funktionen und $u_h \in U_h^{(1)}$ bzw. $u_h \in U_h^{(0)}$ dargestellt durch

$$u_h(x) = \sum_{i=1}^{N_{\text{vertices}}} u_i\,\varphi_i(x) \quad \text{bzw.} \quad u_h(x) = \sum_{i=1}^{N_{\text{cells}}} u_i\,\psi_i(x).$$

Dann gilt:

$$u_a \le u_h(x) \le u_b \quad \text{in } \Omega \quad \Leftrightarrow \quad u_a \le u_i \le u_b \quad \text{für alle } i = 1, \ldots, N_{\text{vertices}} \text{ bzw. } N_{\text{cells}}.$$

Lösung Aufgabe 23:

Sei $K \in \mathcal{T}$ eine Zelle und $\varphi_{i,K}$ bzw. $u_{i,K}$ die zugehörigen globalen Basisfunktionen bzw. die Koeffizienten. Wir nutzen in beiden Fällen ($U_h = U_h^{(0)}$ und $U_h = U_h^{(1)}$), dass alle Basisfunktionen φ_i bzw. ψ_i

(a) nicht-negativ sind,

(b) als Maximum den Wert 1 annehmen

(c) und sich auf jeder Zelle K zu 1 addieren.

Wir betrachten im Folgenden nur den schwierigeren Fall $U_h^{(1)}$. Es folgt

$$u_h(x)_{|K} = \sum_{i=1}^{3} u_{i,K}\,\varphi_{i,K}(x) \overset{(a)}{\leq} \max_{i=1,2,3}\{u_{i,K}\} \sum_{i=1}^{3} \varphi_{i,K}(x) \overset{(c)}{=} \max_{i=1,2,3}\{u_{i,K}\}.$$

Wegen (b) wird beim Knoten $x = a_{i,K}$ tatsächlich der Wert $\max_{i=1,2,3}\{u_{i,K}\}$ angenommen. Es folgt also

$$u_h(x) \leq u_b \quad \text{in } K \quad \Leftrightarrow \quad \max_{x\in K} \sum_{i=1}^{3} u_{i,K}\,\varphi_{i,K}(x) \leq u_b \quad \Leftrightarrow \quad \max_{i=1,2,3} u_{i,K} \leq u_b.$$

Analog zeigt man

$$u_h(x) \geq u_a \quad \text{in } K \quad \Leftrightarrow \quad \min_{x\in K} \sum_{i=1}^{3} u_{i,K}\,\varphi_{i,K}(x) \geq u_a \quad \Leftrightarrow \quad \min_{i=1,2,3} u_{i,K} \geq u_a.$$

Aufgabe 24: Zur Anwendung des verallgemeinerten Newton-Verfahrens

(a) Zeigen Sie, dass die Matrix (9.59)

$$S = \begin{bmatrix} \mathbf{A} & \mathbf{E}_{\text{ctrl}}^{-1}\,\mathbf{D}_{\mathcal{A}}\,\mathbf{E}_{\text{ctrl}} \\ \mathbf{D}_{\mathcal{A}} & -c^{-1}\,\mathbf{D}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}} \end{bmatrix}$$

aus der Vorlesung selbstadjungiert bezüglich des Skalarprodukts

$$M = \begin{bmatrix} \mathbf{E}_{\text{ctrl}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{\text{ctrl}} \end{bmatrix}$$

und indefinit ist.

(b) Zeigen Sie für den Fall einer diagonalen Matrix $\mathbf{E}_{\text{ctrl}}$, dass sich die linke Seite des reduzierten Systems (9.60) erheblich vereinfacht, bzw. genauer, dass

$$(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{I}}^\top)^{-1}(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top) = \mathbf{P}_{\mathcal{I}}\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top$$

gilt. Wie vereinfacht sich die rechte Seite $-(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{I}}^\top)^{-1}\,\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\big(\mathbf{A}(\mathbf{u}+\delta\mathbf{u}_{\mathcal{A}})-\mathbf{b}\big)$?

Lösung Aufgabe 24:

(a) Wir zeigen sie Selbstadjungiertheit durch Nachrechnen von $S^\top M = M\,S$:

$$\begin{aligned}
S^\top M &= \begin{bmatrix} \mathbf{A} & \mathbf{E}_{\text{ctrl}}^{-1}\,\mathbf{D}_{\mathcal{A}}\,\mathbf{E}_{\text{ctrl}} \\ \mathbf{D}_{\mathcal{A}} & -c^{-1}\,\mathbf{D}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}} \end{bmatrix}^\top \begin{bmatrix} \mathbf{E}_{\text{ctrl}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{\text{ctrl}} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}^\top & \mathbf{D}_{\mathcal{A}} \\ \mathbf{E}_{\text{ctrl}}\,\mathbf{D}_{\mathcal{A}}\,\mathbf{E}_{\text{ctrl}}^{-1} & -c^{-1}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{D}_{\mathcal{I}} \end{bmatrix} \begin{bmatrix} \mathbf{E}_{\text{ctrl}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{\text{ctrl}} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}^\top\,\mathbf{E}_{\text{ctrl}} & \mathbf{D}_{\mathcal{A}}\,\mathbf{E}_{\text{ctrl}} \\ \mathbf{E}_{\text{ctrl}}\,\mathbf{D}_{\mathcal{A}} & -c^{-1}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{D}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{E}_{\text{ctrl}}\,\mathbf{A} & \mathbf{E}_{\text{ctrl}}\,\mathbf{E}_{\text{ctrl}}^{-1}\,\mathbf{D}_{\mathcal{A}}\,\mathbf{E}_{\text{ctrl}} \\ \mathbf{E}_{\text{ctrl}}\,\mathbf{D}_{\mathcal{A}} & -c^{-1}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{D}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{E}_{\text{ctrl}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{\text{ctrl}} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{E}_{\text{ctrl}}^{-1}\,\mathbf{D}_{\mathcal{A}}\,\mathbf{E}_{\text{ctrl}} \\ \mathbf{D}_{\mathcal{A}} & -c^{-1}\,\mathbf{D}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}} \end{bmatrix} = M\,S
\end{aligned}$$

(b) Der Trick besteht darin, an geeigneter Stelle eine produktive Null einzufügen. Es gilt

$$\begin{aligned}
&(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{I}}^\top)^{-1}(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top) \\
&= (\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{I}}^\top)^{-1}(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,[\mathbf{P}_{\mathcal{I}}^\top\,\mathbf{P}_{\mathcal{I}} + \mathbf{P}_{\mathcal{A}}^\top\,\mathbf{P}_{\mathcal{A}}]\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top) \\
&= \mathbf{P}_{\mathcal{I}}\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top + (\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{I}}^\top)^{-1}(\underbrace{\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{A}}^\top}_{=0}\,\mathbf{P}_{\mathcal{A}}\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top) \quad = \mathbf{P}_{\mathcal{I}}\,\mathbf{A}\,\mathbf{P}_{\mathcal{I}}^\top.
\end{aligned}$$

Mit demselben Trick kürzt sich auch auf der rechten Seite die Multiplikation mit der Inversen des Skalarprodukts heraus. Wir erhalten

$$-(\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\mathbf{P}_{\mathcal{I}}^\top)^{-1}\,\mathbf{P}_{\mathcal{I}}\,\mathbf{E}_{\text{ctrl}}\,\big(\mathbf{A}(\mathbf{u} + \delta\mathbf{u}_{\mathcal{A}}) - \mathbf{b}\big) = -\mathbf{P}_{\mathcal{I}}\,\big(\mathbf{A}(\mathbf{u} + \delta\mathbf{u}_{\mathcal{A}}) - \mathbf{b}\big)$$