A Lösungen der Übungsaufgaben

Im Folgenden sind Lösungen für die am Ende der einzelnen Kapitel formulierten Aufgaben zusammengestellt. Es handelt sich dabei nicht um „Musterlösungen“ mit vollständig ausformuliertem Lösungsweg, sondern nur um Lösungsansätze in knapper Form.

A.1 Kapitel 1

Lösung A.1.1: Aus der Beschränktheit der ersten Ableitung von f folgt die Lipschitz-Stetigkeit von f bzgl. x, und es ist

$$\|f\,(t,x)\,\|_\infty \leq \|f\,(t,x)\,f-(t,0)\,\|_\infty + \|f\,(t,0)\,\|_\infty \leq K\|x\|_\infty + \|f\,(t,0)\,\|_\infty$$

Somit existiert nach dem globalen Existenzsatz genau eine globale Lösung der AWA. Und diese ist nach dem Regularitätssatz aus C^∞.

Sei nun u diese Lösung. So ist

$$u^{(k)} = f^{(k-1)}$$

Und damit hat die Taylor-Reihe von u im Entwicklungspunkt t_0 gerade die angegebene Form. Es bleibt zu zeigen, dass die Taylorreihe für alle $t > t_0$ konvergiert. Dazu betrachten wir das $(n+1)$-te Restglied der Taylorentwicklung von u.

$$R_{n+1} = \frac{f^{(n)}\,(\xi, u\,(\xi))}{(n+1)!}\,(t-t_0)^{n+1}$$

mit $\xi \in (t_0, t)$. Aus der gleichmäßigen Beschränktheit der Ableitungen von f folgt

$$R_{n+1} \to 0 \quad (n \to \infty)\,.$$

und damit die Behauptung.

Lösung A.1.2: Wir setzen $p := \partial_x u,\ q := \partial_y u\,,\ r := \partial_x^2 u,\ s := \partial_x\partial_y u,\ t := \partial_y^2 u$ und

$$\alpha := \partial_x^3 u,\ \beta := \partial_x^2\partial_y u,\ \gamma = \partial_x\partial_y^2 u,\ \delta = \partial_y^3 u.$$

Differenzieren der Differentialgleichung ergibt

$$\begin{aligned} a_{11}\partial_x^3 u + 2a_{12}\partial_x^2\partial_y u + a_{22}\partial_x\partial_y^2 u &= \partial_x f - a_{01}\partial_x^2 u - a_{02}\partial_x\partial_y u - a_{00}\partial_x u \\ a_{11}\partial_x^2\partial_y u + 2a_{12}\partial_x\partial_y^2 u + a_{22}\partial_y^3 u &= \partial_y f - a_{01}\partial_x\partial_y u - a_{02}\partial_y^2 u - a_{00}\partial_y u \end{aligned}$$

Differenzieren von r, s, t entlang Γ ergibt

$$\begin{aligned} \partial_\tau r &= \partial_x r\partial_\tau x + \partial_y r\partial_\tau y = \alpha\partial_\tau x + \beta\partial_\tau y \\ \partial_\tau s &= \partial_x s\partial_\tau x + \partial_y s\partial_\tau y = \beta\partial_\tau x + \gamma\partial_\tau y \\ \partial_\tau t &= \partial_x t\partial_\tau x + \partial_y t\partial_\tau y = \gamma\partial_\tau x + \delta\partial_\tau y \end{aligned}$$

Zusammengenommen ergeben sich zwei 3×3-Gleichungssysteme für die gesuchten Ableitungen α, β, γ, δ:

$$\begin{aligned} a_{11}\alpha + 2a_{12}\beta + a_{22}\gamma &= \partial_x f - a_{01}r - a_{02}s - a_{00}p \\ \partial_\tau x\alpha + \partial_\tau y\beta &= \partial_\tau r \\ \partial_\tau x\beta + \partial_\tau y\gamma &= \partial_\tau s \end{aligned}$$

$$\begin{aligned} a_{11}\beta + 2a_{12}\gamma + a_{22}\delta &= \partial_y f - a_{01}s - a_{02}t - a_{00}q \\ \partial_\tau x\beta + \partial_\tau y\gamma &= \partial_\tau s \\ \partial_\tau x\gamma + \partial_\tau y\delta &= \partial_\tau t \end{aligned}$$

Beide haben dieselbe Koeffizientenmatrix B wie das entsprechende System zur Bestimmung der zweiten Ableitungen:

$$\begin{aligned} a_{11}r + 2a_{12}s + a_{22}t &= f - a_{01}p - a_{02}q - a_{00}u \\ \partial_\tau xr + \partial_\tau ys &= \partial_\tau p \\ \partial_\tau xs + \partial_\tau yt &= \partial_\tau q. \end{aligned}$$

Man überliegt sich leicht, dass im Falle $\det B \neq 0$ die durch die beiden Gleichunsgsysteme bestimmten vier dritten Ableitungen eindeutig (und widerspruchsfrei) bestimmt sind.

Lösung A.1.3: Der Differentialoperator $L = a_{11}\partial_x + 2a_{12}\partial_x\partial_y + a_{22}\partial_y^2 + \ldots$ ist „elliptisch“ für $a_{12}^2 - a_{11}a_{22} < 0$, „parabolisch'“ für $a_{12}^2 - a_{11}a_{22} = 0$ und „hyperbolisch“ für $a_{12}^2 - a_{11}a_{22} > 0$.

a) Der Operator $L = \partial_x\partial_y - \partial_x$ ist wegen $a_{12}^2 - a_{11}a_{22} = \frac{1}{4} > 0$ hyperbolisch.

b) Der Operator $L = \partial_x^2 + \partial_x\partial_y + y\partial_y^2 + 4$ ist wegen $a_{12}^2 - a_{11}a_{22} = \frac{1}{4} - y$ hyperbolisch für $y < \frac{1}{4}$, parabolisch für $y = \frac{1}{4}$ und elliptisch für $y > \frac{1}{4}$.

c) Der Operator $L = 2(\partial_x + \partial_y)^2 + \partial_y = 2\partial_x^2 + 4\partial_x\partial_y + 2\partial_y^2 + \partial_y$ ist wegen $a_{12}^2 - a_{11}a_{22} = 4 - 4 = 0$ parabolisch.

Lösung A.1.4: Wir wollen im Folgenden die ersten partiellen Ableitungen der u_i ($i = 1, 2$) bestimmen. Dazu führen wir folgende Abkürzungen ein:

$$r_i := \partial_x u_i \qquad s_i := \partial_y u_i$$

Durch Ableiten in Tangentialrichtung erhalten wir

$$\partial_\tau u_i = r_i\partial_\tau x + s_i\partial_\tau y$$

Zusammen mit den 2 Differentialgleichungen ergibt sich das folgende LGS:

$$\begin{pmatrix} b_{11}^1 & b_{12}^1 & b_{21}^1 & b_{22}^1 \\ b_{11}^2 & b_{12}^2 & b_{21}^2 & b_{22}^2 \\ \partial_\tau x & 0 & \partial_\tau y & 0 \\ 0 & \partial_\tau x & 0 & \partial_\tau y \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \partial_\tau u_1 \\ \partial_\tau u_2 \end{pmatrix}$$

Wir erhalten für die Determinante der Matrix

$$\begin{aligned} \begin{vmatrix} b_{11}^1 & b_{12}^1 & b_{21}^1 & b_{22}^1 \\ b_{11}^2 & b_{12}^2 & b_{21}^2 & b_{22}^2 \\ \partial_\tau x & 0 & \partial_\tau y & 0 \\ 0 & \partial_\tau x & 0 & \partial_\tau y \end{vmatrix} &= (\partial_\tau x)^2 \left(b_{21}^1 b_{22}^2 - b_{22}^1 b_{21}^2\right) \\ &\quad + (\partial_\tau y)^2 \left(b_{11}^1 b_{21}^2 - b_{11}^2 b_{21}^1\right) \\ &\quad + \partial_\tau x \partial_\tau y \left(b_{12}^1 b_{21}^2 - b_{21}^1 b_{12}^2 + b_{22}^1 b_{11}^2 - b_{11}^1 b_{22}^2\right). \end{aligned}$$

Durch Setzen von

$$\begin{aligned} \hat{a}_{11} &:= b_{21}^1 b_{22}^2 - b_{22}^1 b_{21}^2 \\ \hat{a}_{22} &:= b_{11}^1 b_{21}^2 - b_{11}^2 b_{21}^1 \\ \hat{a}_{12} &:= \frac{1}{2}\left(b_{12}^1 b_{21}^2 - b_{21}^1 b_{12}^2 + b_{22}^1 b_{11}^2 - b_{11}^1 b_{22}^2\right) \end{aligned}$$

ehen wir, dass die Lösbarkeits-Eigenschaften wieder von dem Vorzeichen des Terms

$$\hat{a}_{12}^2 - \hat{a}_{11}\hat{a}_{22}$$

abhängt. Im Vergleich dazu liefert die PDE 2. Ordnung

$$\begin{aligned} \hat{a}_{11} &:= a_{22} \\ \hat{a}_{22} &:= -a_{11} \\ \hat{a}_{12} &:= -\frac{a_{12} + a_{21}}{2} \end{aligned}$$

und somit im Fall $a_{12} = a_{21}$

$$\hat{a}_{12}^2 - \hat{a}_{11}\hat{a}_{22} = a_{12}^2 - a_{11}a_{22}.$$

Beide Vorgehensweisen liefern also die gleiche Typen-Einteilung.

Lösung A.1.5: (i) Wir verwenden den Beweisgang aus dem Text in leicht modifizierter Form. Für einen beliebigen Punkt $x = (x_1, x_2) \in Q$ ist

$$v(x) = v(x_1, x_2) - v(x_1, 0) = \int_0^{x_2} \partial_2 v(x_1, \xi)\, d\xi.$$

Mit Hilfe der Hölderschen Ungleichung folgt

$$|v(x)|^2 \le \Big(\int_0^{x_2} \partial_2 v(x_1,\xi)\, d\xi \Big)^2 \le \int_0^1 |\partial_2 v(x_1,\xi)|^2\, d\xi .$$

Wir integrieren diese Ungleichung unter Verwendung des Satzes von Fubini nacheinander bzgl. der Variablen x_1, x_2:

$$\begin{aligned} \int_Q |v(x)|^2\, dx &\le \int_0^1 \int_0^1 \left(\int_0^1 |\partial_2 v(x_1,\xi)|^2\, d\xi \right) dx_1 dx_2 \\ &= \int_0^1 \left(\int_0^1 \int_0^1 |\partial_2 v(x_1,\xi)|^2\, dx_1\, d\xi \right) dx_2 = \int_Q |\partial_2 v(x)|^2\, dx \end{aligned}$$

(ii) Für $\Gamma := \{(0,0)\}$ kann die Poincarésche Ungleichung *nicht* gelten. Zum Beweis konstruieren wir eine Folge von Funktionen $u_k \in V_0(\Gamma; Q)$ mit den Eigenschaften

$$\liminf_{k\to\infty} \int_Q |u_k|^2\, dx > 0, \qquad \int_Q \|\nabla u_k\|^2\, dx \to 0 \ \ (k\to\infty).$$

Dazu setzen wir unter Verwendung von Polarkoordinaten (r,θ):

$$u_k(r,\theta) := r^{1/k}.$$

Für diese Funktionen ist $\|\nabla u_k\| = |\partial_r u_k| = k^{-1} r^{-1+1/k}$ und folglich:

$$\begin{aligned} \int_Q |u_k|^2\, dx &\ge \int_0^{\pi/2} \int_0^1 r^{2/k}\, r\, dr d\omega = \frac{\pi}{2} \int_0^1 r^{1+2/k}\, dr \\ &= \frac{\pi}{2} \frac{1}{2/k+2} r^{2+2/k} \Big|_0^1 = \frac{\pi}{2} \frac{1}{2+2/k} \to \frac{\pi}{4} \quad (k\to\infty), \end{aligned}$$

sowie analog

$$\begin{aligned} \int_Q \|\nabla u_k\|^2\, dx &= \frac{1}{k^2} \int_Q r^{-2+2/k}\, dx \le \frac{\pi}{2k^2} \int_0^2 r^{-1+2/k}\, dr \\ &= \frac{\pi}{2k^2} \frac{k}{2} r^{2/k} \Big|_0^2 = \frac{\pi}{4} \frac{2^{2/k}}{k} \to 0 \quad (k\to\infty). \end{aligned}$$

Als Konsequenz dieses Resultats ist in diesem Fall die in dem Text verwendete „direkte Methode der Variationsrechnung“ (d. h. das „Minimalfolgenargument“) zum Nachweis der Existenz „schwacher“ Lösungen der zugehötigen 1. RWA des Laplace-Operators nicht anwendbar, da das Energiefunktional $J(u)$ auf $V_0(\Gamma; Q)$ nicht nach unten beschränkt ist. Tatsächlich bedeutet dies, dass in zwei (und höheren) Dimensionen die 1. RWA mit solchen einpunktigen Dirichlet-Randbedingungen *nicht* „wohl gestellt“ ist.

Lösung A.1.6: a) Seien $u, v \in C^2(\Omega) \cap C^1(\overline{\Omega})$ Lösungen derselben 2. RWA. Dann erfüllt $w := u - v$ die Gleichungen $-\Delta w + a w = 0$ in Ω und $\partial_n w_{|\partial\Omega} = 0$. Mit Hilfe partieller

Inegration folgt unter Aussnutzung der Randbedingung $\partial_n w_{|\partial\Omega} = 0$:

$$\int_\Omega \left\{ \|\nabla w\|^2 + a|w|^2 \right\} dx = \int_\Omega (-\Delta w + aw) w \, dx + \int_{\partial\Omega} \partial_n w w \, do = 0.$$

Dies impliziert $w \equiv 0$.

b) Mit denselben Bezeichnungen wie in (a) gilt nun $(\partial_n w + \alpha w)_{|\partial\Omega} = 0$. Damit ergibt sich dann wegen $\alpha \geq 0$:

$$\int_\Omega \left\{ \|\nabla w\|^2 + a|w|^2 \right\} dx = \int_\Omega (-\Delta w + aw) w \, dx + \int_{\partial\Omega} \partial_n w w \, do = -\int_{\partial\Omega} \alpha w^2 \, do \leq 0.$$

Dies impliziert wieder $w \equiv 0$.
Für $a = 0$ kann in beiden Fällen nur auf $\nabla w \equiv 0$ bzw. $w \equiv konst$ geschlossen werden. Es fehlt aber eine zusätzliche Bedingung, um hieraus $w \equiv 0$ folgern zu können. Eine solche Zusatzbedingung könnte z. B. die Forderung sein, daß nach Lösungen der RWAn mit verschwindendem Mittelwert gefragt ist: $\int_\Omega u \, dx = 0$.

Lösung A.1.7: Wir zeigen die beiden Ungleichungen einzeln:

i) Wir wollen zeigen $u \geq 0$. Es ist

$$-\Delta(-u) = -1 \leq 0 \quad \text{in } \Omega, \quad u = 0 \leq 0 \quad \text{auf } \partial\Omega$$

Nach dem Maximum-Prinzip ist also

$$-u \leq 0 \quad \text{oder} \quad \max_\Omega -u \leq \max_{\delta\Omega} -u = 0.$$

Also $u \geq 0$ auf $\overline{\Omega}$.

ii) Nun zeigen wir $u \leq \frac{1}{8}$. Dazu betrachten wir die Funktion $v = \frac{1}{4}(x(1-x) + y(1-y))$ mit

$$-\Delta v = \tfrac{1}{4}(2+2) = 1.$$

Wegen

$$\nabla v = \begin{pmatrix} 1-2x \\ 1-2y \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

sowie

$$v\left(\tfrac{1}{2}, \tfrac{1}{2}\right) = \tfrac{1}{8}, \qquad v = 0 \text{ auf } \partial Q_1$$

liegt in $(x, y) = (1/2, 1/2)$ ein Maximum von v, und es ist $v \geq 0$ in Q_1. Wegen $\Omega \subset Q_1$ folgt nun

$$-\Delta(u - v) = 0 \leq 0 \quad \text{in } \Omega$$

sowie

$$u - v = -v \leq 0 \quad \text{auf } \partial\Omega.$$

Nach dem Maximumprinzip ist also

$$u - v \leq 0 \quad \text{oder} \quad \max_{\Omega} u - v \leq \max_{\delta\Omega} u - v \leq 0$$

und somit

$$u \leq v \leq \tfrac{1}{8}.$$

Lösung A.1.8: Zur Wiederholung: Wir setzen $x_1 = r\cos\theta$, $x_2 = r\sin\theta$ und $u(x) = u(x_1, x_2) = u(r\cos\theta, r\sin\theta)$. Mit Hilfe der Kettenregel gilt dann:

$$\begin{aligned}
\partial_r^2 u(r,\theta) &= \partial_1^2 u(x)\cos^2\theta + \partial_2\partial_1 u(x)\sin\theta\cos\theta + \partial_1\partial_2 u(x)\cos\theta\sin\theta + \partial_2^2 u(x)\sin^2\theta,\\
\partial_\theta^2 u(r,\theta) &= \partial_1^2 u(x) r^2\cos^2\theta - \partial_2\partial_1 u(x) r^2\sin\theta\cos\theta - \partial_1\partial_2 u(x) r^2\cos\theta\sin\theta\\
&\quad - \partial_1 u(x) r\cos\theta - \partial_2 u(x) r\sin\theta + \partial_2^2 u(x) r^2\cos^2\theta.
\end{aligned}$$

Also ist $(\partial_r^2 + r^{-1}\partial_r + r^{-2}\partial_\theta^2)u(r,\theta) = (\partial_1^2 + \partial_2^2)u(x) = \Delta u(x)$.

i) Die Randbedingungen ließt man direkt ab. Wir setzen $\alpha := \pi/\omega$ und finden

$$\begin{aligned}
\Delta s_\omega(r,\theta) &= (\partial_r^2 + r^{-1}\partial_r + r^{-2}\partial_\theta^2)(r^\alpha \sin\theta\alpha)\\
&= (\alpha-1)\alpha r^{\alpha-2}\sin\theta\alpha + \alpha r^{\alpha-2}\sin\theta\alpha - r^{\alpha-2}\alpha^2\sin\theta\alpha = 0.
\end{aligned}$$

ii) Die Funktion s_ω ist im Innern des Sektorabschnitts G beliebig oft differenzierbar. Ihre ersten und zweiten Ableitungen verhalten sich dort wie

$$|\partial_i s_\omega(r,\theta)| \leq c\, r^{\pi/\omega-1}, \qquad |\partial_i\partial_j s_\omega(r,\theta)| \leq c\, r^{\pi/\omega-2}.$$

Zu überprüfen ist also die Existenz der (uneigentlichen) Integrale

$$\begin{aligned}
J_1(\omega) &:= \int_G r^{2\pi/\omega-2}\,dx = \int_0^\omega\int_0^1 r^{2\pi/\omega-1}\,dr\,d\theta = \omega\int_0^1 r^{2\pi/\omega-1}\,dr,\\
J_2(\omega) &:= \int_G r^{2\pi/\omega-4}\,dx = \int_0^\omega\int_0^1 r^{2\pi/\omega-3}\,dr\,d\theta = \omega\int_0^1 r^{2\pi/\omega-3}\,dr.
\end{aligned}$$

Für $\pi < \omega \leq 2\pi$ ist $2\pi/\omega - 1 \geq 0$ und folglich $J_1(\omega)$ existent, aber $2\pi/\omega - 3 < -1$ und folglich $J_2(\omega)$ *nicht* existent. Im Fall $\omega < \pi$ ist ∇s_ω beschränkt und somit (eigentlich) quadrat-integrierbar. Ferner ist $2\pi/\omega - 3 > -1$ und somit auch $\nabla^2 s_\omega$ wenigstens quadrat-integrierbar.

Lösung A.1.9: a) Es ist

$$\partial_x u\,(x,y) = \begin{cases} \frac{1}{2}\,(x-y)^{\frac{-1}{2}}, & x \geq y,\\ \frac{-1}{2}\,(y-x)^{\frac{-1}{2}}, & y < x, \end{cases}$$

und $\partial_y u = -\partial_x u$, so dass es genügt, $\partial_x u$ zu betrachten. Wir betrachten dazu die

Ausschöpfung $\Omega_\varepsilon := \{(x,y) \in \mathbb{R}^2 : x < y - \varepsilon \text{ oder } x > y + \varepsilon\}$ und wir erhalten

$$\int_{\Omega_\varepsilon} (\partial_x u)^2 \, d(x,y) = \int_0^1 \int_0^{y-\varepsilon} \frac{1}{4} (y-x)^{-1} \, dx \, dy + \int_0^1 \int_{y+\varepsilon}^1 \frac{1}{4} (x-y)^{-1} \, dx \, dy$$

Wir begnügen uns damit, das erste Integral auf der rechten Seite zu betrachten:

$$\begin{aligned} \int_0^1 \int_0^{y-\varepsilon} \frac{1}{4} (y-x)^{-1} \, dx \, dy &= \int_0^1 \frac{-\ln(y-x)}{4} \Big|_0^{y-\varepsilon} \, dy \\ &= \frac{-\ln(\varepsilon)}{4} + \int_0^1 \frac{\ln(y)}{4} \, dy \to \infty \ (\varepsilon \to 0) \end{aligned}$$

Also ist $\partial_x u$ nicht in $L^2(\Omega)$ und somit $u \notin H^1(\Omega)$.

b) Durch Anwendung der Kettenregel erhalten wir für die partiellen Ableitungen von u:

$$\partial_x u(x,y) = \cos\left(\ln\left(\frac{1}{r}\right)\right) \frac{-x}{r^2}, \qquad \partial_y u(x,y) = \cos\left(\ln\left(\frac{1}{r}\right)\right) \frac{-y}{r^2}.$$

Somit erhalten wir

$$\nabla u \nabla u = (\partial_x u)^2 + (\partial_y u)^2 = \cos\left(\ln\left(\frac{1}{r}\right)\right) \frac{x^2+y^2}{r^4} = \cos\left(\ln\left(\frac{1}{r}\right)\right)^2 \frac{1}{r^2}$$

Wir betrachten jetzt nur die Kreisbögen $S_\varepsilon = \{(x,y) \in \mathbb{R}^2 : \varepsilon < x^2 + y^2 < 1\}$. Und bestimmen das Integral

$$\begin{aligned} \int_{S_\varepsilon} \frac{\cos\left(\ln\left(\frac{1}{r}\right)\right)^2}{r^2} \, d(x,y) &= \int_\varepsilon^1 \int_0^{\frac{\pi}{2}} \frac{\cos\left(\ln\left(\frac{1}{r}\right)\right)^2}{r^2} r \, d\Theta \, dr \\ &= \frac{\pi}{2} \int_\varepsilon^1 \frac{\cos\left(\ln\left(\frac{1}{r}\right)\right)^2}{r} \, dr \\ &= \frac{\pi}{2} \int_{\ln(\varepsilon)}^0 \cos(x)^2 \, dx \\ &= \frac{\pi}{4} (\cos(x)\sin(x) + x) \Big|_{\ln(\varepsilon)}^0 \\ &= \frac{-\pi}{4} (\cos(\ln(\varepsilon))\sin(\ln(\varepsilon)) + \ln(\varepsilon)) \to \infty \ (\varepsilon \to 0) \end{aligned}$$

Also ist $u \notin H^1(\Omega)$.

Lösung A.1.10: Zunächst eine Feststellung, die Menge

$$M := \left\{ u \in H^1(\Omega) : \int_\Omega u(x) \, dx = 0 \right\}$$

ist konvex und abgeschlossen bzgl. der L^2-Topologie. Die Konvexität folgt sofort aus der

Linearität des Integrals, die Abgeschlossenheit, da

$$\left| \int_\Omega u(x)\ dx \right| \leq \int_\Omega |u(x)|\ dx \leq c \|u\|_\Omega .$$

Wir zeigen die Aussage durch ein Widerspruchsargument. Angenommen es gäbe keine Konstante c_Ω mit der Eigenschaft. Dann gibt es eine Folge u_n aus M, so dass

$$\|u_n\|_\Omega > n \|\nabla u_n\|_\Omega$$

bzw. $x_n := \|u_n\|_\Omega^{-1} \to 0\ (n \to \infty)$. Wir können also o.B.d.A. annehmen, daß $0 < x_n \leq 1,\ n \in \mathbb{N}$. Da $0 \in M$ und M konvex, ist auch $v_n := x_n u_n \in M$, und es gilt $\|v_n\|_\Omega = 1$. Aus der Ungleichung für u_n erhalten wir

$$\|\nabla v_n\|_\Omega = x_n \|\nabla u_n\|_\Omega < \frac{x_n}{n} \|u_n\|_\Omega = \frac{\|v_n\|_\Omega}{n} \to 0\ (n \to \infty).$$

Also ist v_n in $H^1(\Omega)$ beschränkt. Es gibt also eine Teilfolge, wir wollen o.B.d.A. annehmen dies wäre v_n, die schwach gegen $v \in H^1(\Omega)$ konvergiert. Insbesonder konvergiert ∇v_n schwach in L^2 gegen ∇v. Wegen obiger Ungleichung konvergiert aber ∇v_n stark in L^2 gegen 0, folglich ist $\nabla v = 0$. Da Ω ein Gebiet ist (also insbesondere zusammenhängend) folgt daraus $v =$ konst fast überall in Ω.

Nach dem Rellichschen Auswahlsatz gibt es eine stark in L^2 konvergente Teilfolge von v_n, die wir wieder mit v_n bezeichnen, wegen $\|v_n\|_\Omega = 1$ folgt also aus der starken Konvergenz $\|v\|_\Omega = 1$, und somit $v = 1$ fast überall auf Ω. Andererseits ist M bezüglich der Norm $\|\cdot\|_\Omega$ abgeschlossen, also folgt $v \in M$ im Widerspruch zu $v = 1$. Damit ist die Behauptung bewiesen.

Lösung A.1.11: Wir können im allgemeinen nur Existenz und Eindeutigkeit einer Lösung $u \in V$ der zugehörigen schwachen Formulierung, d. h. bezüglich

$$(\nabla u, \nabla \varphi) = (f, \varphi) \quad \forall \varphi \in V \tag{1.1.1}$$

erwarten. Wir zeigen zunächst die Existenz und Eindeutigkeit einer Lösung für den Raum

$$V := \Big\{ u \in H^1(\Omega) : \int_\Omega u(x)\, dx = 0 \Big\}.$$

Die Existenz einer solchen Lösung ist gegeben, wenn wir zeigen können, dass das Funktional

$$E(v) = \tfrac{1}{2} \|\nabla v\|_\Omega^2 - (f, v)_\Omega$$

in V ein Minimum annimmt. Zunächst zeigen wir, dass E in V nach unten beschränkt ist. Mithilfe der Poincaréschen Ungleichung aus der vorausgehenden Aufgabe

$$\int_\Omega f v dx \leq \|f\|_\Omega \|v\|_\Omega \leq c_\Omega \|f\|_\Omega \|\nabla v\|_\Omega .$$

Mithilfe der Youngschen Ungleichung folgt weiter

$$\int_\Omega fv\ dx \le \frac{c_\Omega^2}{2}\|f\|_\Omega^2 + \frac{1}{2}\|\nabla v\|_\Omega^2$$

und damit die Beschränktheit von $E(\cdot)$ nach unten:

$$E(v) \ge -\frac{c_\Omega^2}{2}\,\|f\|_\Omega^2\,.$$

Sei nun $(v_k)_{k\in\mathbb{N}}, v_k \in V$ eine Minimalfolge bezüglich $E(\cdot)$:

$$E(v_k) \to d := \inf_{v\in V} E(v).$$

Wie im Skript zeigen wir mithilfe der Parallelogramm-Identität, dass (v_k) Cauchy-Folge bezüglich der Energienorm

$$\|v\|_E := \|\nabla v\|_\Omega$$

ist:

$$\begin{aligned}\|v_n - v_m\|_E^2 &= 2\|v_n\|_E^2 + 2\|v_m\|_E^2 - 4\|\tfrac{1}{2}(v_n+v_m)\|_E^2\\ &\le 4E(v_n) + 4(f,v_n) + 4E(v_m) + 4(f,v_m) - 8E(\tfrac{1}{2}(v_n+v_m)) - 8(f,\tfrac{1}{2}(v_n+v_m))\\ &= 4E(v_n) + 4E(v_m) - 8E(\tfrac{1}{2}(v_n+v_m))\end{aligned}$$

Unter Berücksichtigung von $\lim_{k\to\infty} E(v_k) = d$ und $E\left(\frac{1}{2}(v_n+v_m)\right) \ge d$ folgt

$$\lim \sup_{n,m\to\infty} \|v_n - v_m\|_E^2 \le 0.$$

Aufgrund der Poincaréschen Ungleichung auf V gilt weiter

$$\|v_n - v_m\|_{H^1(\Omega)} \le C\|v_n - v_m\|_E,$$

d. h. $(v_k)_{k\in\mathbb{N}}$ ist Cauchy-Folge bezüglich der H^1-Norm und konvergiert folglich gegen einen Limes $v \in H^1(\Omega)$. Um $v \in V$ zu garantieren, müssen wir noch zeigen, dass auch die Mittelwertbedingung im Limes erhalten bleibt

$$\Big|\int_\Omega v dx\Big| = \Big|\int_\Omega v - v_k dx\Big| \le \|v - v_k\|_\Omega \to 0 (k\to\infty).$$

Wir haben also eine schwache Lösung $v \in V$ gefunden. Aufgrund der Mittelwertbedingung an f bleibt (1.1.1) auch gültig, wenn man den Testraum auf $H^1(\Omega)$ erweitert (Für $\varphi \in H^1(\Omega)$ liegt $\hat\varphi := \varphi - |\Omega|^{-1}\int_\Omega \varphi$ in M und

$$(\nabla u, \nabla\varphi) = (\nabla u, \nabla\hat\varphi), \quad (f,\varphi) = (f,\hat\varphi).$$

Wir haben also eine schwache Lösung $v \in H^1(\Omega)$ für das Problem

$$(\nabla u, \nabla\varphi) = (f,\varphi) \quad \forall\varphi \in H^1(\Omega)$$

gefunden. Seien nun v_1, v_2 zwei Lösungen, so gilt für $w = v_1 - v_2$

$$(\nabla w, \nabla w) = 0$$

und daher $v_1 = v_2 + const.$

Lösung A.1.12: a) Die RWA auf der gepunkteten Kreisscheibe ist *nicht* wohl gestellt, da das Energiefunktional

$$E(v) = \tfrac{1}{2}\|\nabla v\|_\Omega^2 - (f, v)_\Omega$$

auf $H_0^1(\Omega)$ kein Minimum besitzt. Es existiert eine Folge $(v_k)_{k\in\mathbb{N}} \subset H_0^1(\Omega)$, welche bzgl. der H^1-Norm eine Cauchy-Folge ist, für deren Limes aber $u(0) \neq 0$ ist. Eine reguläre schwache Lösung genügt also nicht der geforderten Randbedingung in $x = 0$.

b) Die RWA auf der geschlitzten Kreisscheibe ist wohl gestellt, denn zu jeder Cauchy-Folge $(v_k)_{k\in\mathbb{N}} \subset H_0^1(\Omega)$ gehört aufgrund der Spurabschätzung

$$\|v\|_{L^2(\Gamma)} \leq c\|v\|_{H^1(\Omega)},$$

eine Cauchy-Folge $(v_{k|\Gamma})_{k\in\mathbb{N}} \subset L^2(\Gamma)$ von Spuren. Der H^1-Limes $v = \lim_{k\to\infty} v_k$ hat also automatisch die Spur $v_{|\Gamma} \equiv 0$. Jede reguläre schwache Lösung genügt also der geforderten homogenen Dirichlet-Randbedingung auf Γ.

Lösung A.1.13: Gesucht ist $u \in H^1(\Omega)$ mit der Eigenschaft

$$(\nabla u, \nabla\varphi)_\Omega + (u, \varphi)_\Omega + (u, \varphi)_{\partial\Omega} = (f, \varphi)_\Omega + (g, \varphi)_{\partial\Omega} \quad \forall\varphi \in H^1(\Omega).$$

Für eine hinreichend reguläre Lösung folgt durch partielle Integration

$$(-\Delta u + u - f, \varphi)_\Omega + (\partial_n u + u - g, \varphi)_{\partial\Omega} = 0 \quad \forall\varphi \in H^1(\Omega),$$

und damit die Gültigkeit der gegebenen Differentialgleichung sowie der Randbedingung.

Lösung A.1.14: Für die Ungleichungen gilt:

$$\begin{array}{lll}
a) & \|u\|_{L^\infty(\Omega)} \leq c\,\|u\|_{H^2(\Omega)}, & u \in H^2(\Omega),\ \Omega \subset \mathbb{R}^3; \quad (richtig) \\
b) & \|u\|_{L^\infty(\Omega)} \leq c\,\|u\|_{H^{1,1}(\Omega)}, & u \in H^{1,1}(\Omega),\ \Omega \subset \mathbb{R}^1; \quad (richtig) \\
c) & \|u\|_{L^\infty(\Omega)} \leq c\,\|u\|_{H^1(\Omega)}, & u \in H^1(\Omega),\ \Omega \subset \mathbb{R}^2; \quad (falsch) \\
d) & \|u\|_{L^1(\partial\Omega)} \leq c\,\|u\|_{H^{1,1}(\Omega)}, & u \in H^{1,1}(\Omega),\ \Omega \subset \mathbb{R}^2 \quad (richtig).
\end{array}$$

Lösung A.1.15: Wir betrachten zunächst die rechte Seite der Ungleichung. Sei $u^0 \in H^1(\Omega)$. Wir verwenden die L^2-Orthonormalbasis aus Eigenfunktionen des Laplace-Operators $(v_k)_{k\in\mathbb{N}}$. Für u^0 gelte die Darstellung

$$u^0(x) = \sum_{j=0}^{\infty} u_j^0 v_j.$$

Die n-te Partialsumme

$$s_n := \sum_{j=0}^{n} u_j^0 v_j(x).$$

konvergiert in $H^1(\Omega)$ gegen u^0. Insbesondere gilt

$$\|\nabla u^0\|_\Omega^2 = \lim_{n\to\infty} \|\nabla s_n\|_\Omega^2.$$

Für die Partialsumme können wir Integration und Differentiation mit der Summation vertauschen

$$(\nabla s_n, \nabla s_n)_\Omega = \sum_{j,k=0}^{n} u_j^0 u_k^0 \left(\nabla v_j(x), \nabla v_k(x)\right)_\Omega$$

Wir nutzen aus, dass v_j auch schwache Lösungen des Eigenwertproblems des Laplace-Operators sind, d. h.:

$$(\nabla v_j, \nabla \varphi)_\Omega = \lambda_j (v_j, \varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

Wir bekommen

$$(\nabla s_n, \nabla s_n)_\Omega = \sum_{j,k=0}^{n} u_j^0 u_k^0 \lambda_j \left(v_j(x), v_k(x)\right)_\Omega .$$

Schließlich nutzen wir die Orthonormalitätseigenschaft der v_j. Insgesamt haben wir gezeigt

$$\|\nabla u^0\|_\Omega^2 = \sum_{j=0}^{\infty} \left(u_j^0\right)^2 \lambda_j. \tag{1.1.2}$$

Für die Lösung der Wärmeleitungsgleichung gilt die Darstellung

$$u(x,t) = \sum_{j=1}^{\infty} u_j^0 v_j(x) e^{-\lambda_j t}.$$

Differenzieren ergibt

$$\partial_t u(x,t) = \Delta u(x,t) = -\sum_{j=1}^{\infty} u_j^0 \lambda_j v_j(x) e^{-\lambda_j t}.$$

Mit der Parsevalschen Identität gilt

$$\|\partial_t u(x,t)\|_\Omega^2 = \|\Delta u(x,t)\|_\Omega^2 = \sum_{j=1}^{\infty} (u_j^0)^2 (\lambda_j)^2 e^{-2\lambda_j t}.$$

Die Funktion xe^{-2x} nimmt ihr Maximum auf $[0,\infty)$ in $x = 0.5$ an. Es gilt

$$xe^{-2x} \leq \tfrac{1}{2}e^{-1}.$$

Mit diesem Resultat können wir weiter abschätzen

$$\|\partial_t u(x,t)\|_\Omega^2 = \|\Delta u(x,t)\|_\Omega^2 = \frac{1}{2t}e^{-1}\sum_{j=1}^{\infty}(u_j^0)^2\lambda_j.$$

Zusammen mit (1.1.2) haben wir gezeigt

$$\|\partial_t u(x,t)\|_\Omega^2 = \|\Delta u(x,t)\|_\Omega^2 \leq \frac{1}{2t}e^{-1}\|\nabla u^0\|_\Omega^2.$$

Wegen

$$\sqrt{\tfrac{1}{2}e^{-1}} = 0.42888... < 0.5$$

folgt daraus die Behauptung

Lösung A.1.16: Wir machen den Ansatz $u(x,t) = \psi(t)v(x)$. Damit ergibt sich

$$\partial_t^2 u - \Delta u = \psi''(t)v(x) - \psi(t)\Delta v(x) = 0$$

bzw.

$$\frac{\psi''(t)}{\psi(t)} = \frac{\Delta v(x)}{v(x)} = \text{const} = -\lambda.$$

Dies führt auf die Eigenwertprobleme

$$-\Delta v = \lambda v, \quad -\psi''(t) = \lambda\psi(t)$$

mit den in der Aufgabenstellung gegebenen Randwerten. Die Eigenfunktionen v_j für das Laplace-Problem mit Nullrandbedingungen definieren ein Orthonormalsystem in $L^2(\Omega)$. Wir wissen, dass die zugehörigen Eigenwerte λ_j positiv sind. Die Differentialgleichung in der Zeit hat dann die Fundamentallösung

$$\psi(t) = \sum_{j=0}^{\infty} a_j \sin(\sqrt{\lambda_j}t) + b_j \cos(\sqrt{\lambda_j}t).$$

Zusammen ergibt sich

$$u(x,t) = v(x)\psi(t) = \sum_{j=0}^{\infty} v_j(x)\left(a_j \sin(\sqrt{\lambda_j}t) + b_j \cos(\sqrt{\lambda_j}t)\right).$$

Machen wir für die Funktionen $u^0(x), u^1(x)$ den Ansatz

$$u^0(x) = \sum_{j=0}^{\infty} u_j^0 v_j(x), \quad u^1(x) = \sum_{j=0}^{\infty} u_j^1 v_j(x)$$

erhalten wir durch Koeffizientenvergleich $b_j = u_j^0$ sowie nach Ableiten $a_j = \frac{u_j^1}{\sqrt{\lambda_j}}$. Wir erhalten die Darstellung

$$u(x,t) = \sum_{j=0}^{\infty} v_j(x)\Big(\frac{u_j^1}{\sqrt{\lambda_j}}\sin(\sqrt{\lambda_j}t) + u_j^0\cos(\sqrt{\lambda_j}t)\Big).$$

Es bleibt zu untersuchen, welche Regularitätsanforderungen wir an u^0 und u^1 stellen müssen, um

$$\partial_t^2 u, \Delta u \in L^2(\Omega \times (0,T))$$

garantieren zu können. Differenzieren führt zu

$$\partial_t^2 u = \Delta u = -\sum_{j=0}^{\infty} v_j(x)\left(u_j^1\lambda_j^{\frac{1}{2}}\sin(\sqrt{\lambda_j}t) + u_j^0\lambda_j\cos(\sqrt{\lambda_j}t)\right).$$

Mithilfe der Parsevalschen Identität erhalten wir

$$\begin{aligned}\|\partial_t^2 u\|^2_{L^2(\Omega\times(0,T))} &= \int_0^T \int_\Omega \left(\sum_{j=0}^{\infty} v_j(x)\left(u_j^1\lambda_j^{\frac{1}{2}}\sin(\sqrt{\lambda_j}t) + u_j^0\lambda_j\cos(\sqrt{\lambda_j}t)\right)\right)^2 dxdt \\ &= \int_0^T \left(\sum_{j=0}^{\infty}\left(u_j^1\lambda_j^{\frac{1}{2}}\sin(\sqrt{\lambda_j}t) + u_j^0\lambda_j\cos(\sqrt{\lambda_j}t)\right)\right)^2 dt.\end{aligned}$$

Wir ziehen Betragsstriche unter die Summe und schätzen die trigonometrischen Terme ab

$$\|\partial_t^2 u\|^2_{L^2(\Omega\times(0,T))} \le T\sum_{j=0}^{\infty}(\lambda_j u_j^0)^2 + \lambda_j\left(u_j^1\right)^2 + 2\lambda_j^{\frac{3}{2}}u_j^0u_j^1.$$

Ausnutzen der Youngschen Ungleichung führt auf

$$\|\partial_t^2 u\|^2_{L^2(\Omega\times(0,T))} \le T\sum_{j=0}^{\infty}(\lambda_j u_j^0)^2 + \lambda_j\left(u_j^1\right)^2. \tag{1.1.3}$$

Wir wollen nun $u^0 \in H^2(\Omega)$ annehmen. Wir definieren die n-te Partialsumme

$$s_n := \sum_{j=0}^{n} u_j^0 v_j(x).$$

Diese konvergiert dann in $H^2(\Omega)$ gegen u^0. Damit folgt insbesondere

$$\infty > \|\Delta u^0\|_\Omega^2 = \lim_{n\to\infty}\|\Delta s_n\|_\Omega^2.$$

Für die Partialsumme können wir Integration und Differentiation mit der Summation

vertauschen

$$\begin{aligned}(\Delta s_n, \Delta s_n)_\Omega &= \sum_{j,k=0}^{n} u_j^0 u_k^0 \left(\Delta v_j(x), \Delta v_k(x)\right)_\Omega \\ &= \sum_{j,k=0}^{n} u_j^0 u_k^0 \left(\lambda_j v_j(x), \lambda_k v_k(x)\right)_\Omega .\end{aligned}$$

Schließlich folgt aufgrund der Orthonormalitätseigenschaft der v_n

$$\infty > \lim_{n\to\infty} \sum_{j=0}^{n} \left(u_j^0 \lambda_j\right)^2 = \sum_{j=0}^{\infty} \left(u_j^0 \lambda_j\right)^2 .$$

Dies ist genau der erste Teil der Summe in (1.1.3). Sei nun $u^1 \in H^1(\Omega)$. Die n-te Partialsumme

$$r_n := \sum_{j=0}^{n} u_j^1 v_j(x).$$

konvergiert in $H^1(\Omega)$ gegen u^1 und wir haben

$$\infty > \|\nabla u^1\|_\Omega^2 = \lim_{n\to\infty} \|\nabla r_n\|_\Omega^2.$$

Nutzen wir aus, dass v_j auch schwache Lösung des Eigenwertproblems sind, erhalten wir mit derselben Argumentation wie oben

$$\begin{aligned}(\nabla r_n, \nabla r_n)_\Omega &= \sum_{j,k=0}^{n} u_j^1 u_k^1 \left(\nabla v_j(x), \nabla v_k(x)\right)_\Omega \\ &= \sum_{j,k=0}^{n} u_j^1 u_k^1 \lambda_j \left(v_j(x), v_k(x)\right)_\Omega .\end{aligned}$$

Schließlich folgt wieder mit der Orthonormalitätseigenschaft der v_n

$$\infty > \sum_{j=0}^{\infty} \left(u_j^1\right)^2 \lambda_j.$$

Wir haben gezeigt, dass (1.1.3) und damit sowohl $\partial_t^2 u$ als auch Δu in $L^2(\Omega \times (0,T))$ wohldefiniert sind. Unter den getroffenen Regularitätsvoraussetzungen löst das oben konstruierte $u(x,t)$ die Wellengleichung also in einem (starken) L^2-Sinne.

A.2 Kapitel 2

Lösung A.2.1: Wir bestimmen den Abschneidefehler

$$\tau_h(x,y) := -\Delta_h^{(9)} u(x,y) - f_h(x,y) = -\Delta_h^{(9)} u_h - f - \tfrac{1}{12}h^2\Delta f.$$

Taylor-Entwicklung von $u(x\pm h,y), u(x,y\pm h)$ und $u(x\pm h,y\pm h)$ ergibt:

$$\begin{aligned}
u(x\pm h,y) &= \big(1 \pm h\partial_x + \tfrac{1}{2}h^2\partial_x^2 \pm \tfrac{1}{6}h^3\partial_x^3 + \tfrac{1}{24}h^4\partial_x^4 \pm \tfrac{1}{120}h^5\partial_x^5\big)u(x,y)\\
&\quad + \tfrac{1}{720}h^6\partial_x^6 u(\xi,\eta),\\
u(x,y\pm h) &= \big(1 \pm h\partial_y + \tfrac{1}{2}h^2\partial_y^2 \pm \tfrac{1}{6}h^3\partial_y^3 + \tfrac{1}{24}h^4\partial_y^4 \pm \tfrac{1}{120}h^5\partial_y^5\big)u(x,y)\\
&\quad + \tfrac{1}{720}h^6\partial_y^6 u(\xi,\eta),\\
u(x\pm h,y\pm h) &= \big(1 \pm h\partial_x \pm h\partial_y + \tfrac{1}{2}h^2\partial_x^2 \pm h^2\partial_x\partial_y + \tfrac{1}{2}h^2\partial_y^2 \pm \tfrac{1}{6}h^3\partial_x^3 \pm \tfrac{1}{2}h^3\partial_x^2\partial_y\\
&\quad \pm \tfrac{1}{2}h^3\partial_x\partial_y^2 \pm \tfrac{1}{6}h^3\partial_y^3 + \tfrac{1}{24}h^4\partial_x^4 \pm \tfrac{1}{6}h^4\partial_x^3\partial_y + \tfrac{1}{4}h^4\partial_x^2\partial_y^2 \pm \tfrac{1}{6}h^4\partial_x\partial_y^3\\
&\quad + \tfrac{1}{24}h^4\partial_y^4 \pm \tfrac{1}{120}h^5\partial_x^5 \pm \tfrac{1}{24}h^5\partial_x\partial_y^4 \pm \tfrac{1}{12}h^5\partial_x^2\partial_y^3 \pm \tfrac{1}{12}h^5\partial_x^3\partial_y^2 \pm \tfrac{1}{24}h^5\partial_x^4\partial_y\\
&\quad \pm \tfrac{1}{120}h^5\partial_y^5\big)u(x,y) + \big(\tfrac{1}{720}h^6\partial_x^6 \pm \tfrac{1}{120}h^6\partial_x^5\partial_y + \tfrac{1}{48}h^6\partial_x^4\partial_y^2 \pm \tfrac{1}{36}h^6\partial_x^3\partial_y^3\\
&\quad + \tfrac{1}{48}h^6\partial_x^4\partial_y^2 \pm \tfrac{1}{120}h^6\partial_x\partial_y^5 + \tfrac{1}{720}h^6\partial_y^6\big)u(\xi,\eta).
\end{aligned}$$

Damit folgt für den kompakten 9-Punkte-Operator:

$$\begin{aligned}
\Delta_h^{(9)} u_h(x,y) &= \frac{1}{6h^2}\Big\{4u(x\pm h,y) + 4u(x,y\pm h) + u(x\pm h,y\pm h) - 20u(x,y)\Big\}\\
&= \frac{1}{6h^2}\Big((8+8+4-20) + h(0+0)\partial_x + h(0+0)\partial_y\\
&\quad + \tfrac{1}{2}h^2(8+4)\partial_x^2 + h^2(0)\partial_x\partial_y + \tfrac{1}{2}h^2(8+4)\partial_y^2\\
&\quad + \tfrac{1}{6}h^3(0+0)\partial_x^3 + \tfrac{1}{2}h^3(0)\partial_x^2\partial_y + \tfrac{1}{2}h^3(0)\partial_x\partial_y^2 + \tfrac{1}{6}h^3(0+0)\partial_y^3\\
&\quad + \tfrac{1}{24}h^4(8+4)\partial_x^4 + \tfrac{1}{6}h^4(0)\partial_x^3\partial_y + \tfrac{1}{4}h^4(4)\partial_x^2\partial_y^2 + \tfrac{1}{6}h^4(0)\partial_x\partial_y^3 + \tfrac{1}{4}h^4(8+4)\partial_y^4\\
&\quad + \tfrac{1}{120}h^5(0+0+0+0+0+0)\partial_x^5\Big)u(x,y) + \mathcal{O}(M_6(u)h^4)\\
&= \big(\partial_x^2+\partial_y^2\big)u(x,y) + \tfrac{1}{12}h^2\big(\partial_x^4 + 2\partial_x^2\partial_y^2 + \partial_y^4\big)u(x,y) + \mathcal{O}(M_6(u)h^4)\\
&= \Delta u(x,y) + \tfrac{1}{12}h^2\Delta^2 u(x,y) + \mathcal{O}(M_6(u)h^4)\\
&= -f(x,y) - \tfrac{1}{12}h^2\Delta f(x,y) + \mathcal{O}(M_6(u)h^4).
\end{aligned}$$

Also ist $\tau_h(x,y) = \mathcal{O}(h^4)$, d. h.: Die Diskretisierung hat die Konsistenzordnung $m = 4$.

Lösung A.2.2: a) Zur Rechtfertigung der Formel schreiben wir

$$\frac{\|e_h\|_h}{\|e_h\|_{h/2}} = \frac{h^\alpha}{(h/2)^\alpha} = 2^\alpha$$

und erhalten durch Logarithmieren

$$\alpha = \frac{\log(\|e_h\|_h / \|e_h\|_{h/2})}{\log(2)}.$$

Wenn die Lösung u unbekannt ist, machen wir mit einer h-unabhängigen Verteilungsfunktion $c(x)$ den (heuristischen) Ansatz

$$\begin{aligned} u_h - u_{h/2} &= u_h - u + u - u_{h/2} = e_{h/2} - e_h = c(h/2)^\alpha - ch^\alpha \\ &= ch^\alpha(2^{-\alpha} - 1) \\ u_{h/2} - u_{h/4} &= u_{h/2} - u + u - u_{h/4} = e_{h/4} - e_{h/2} = c(h/4)^\alpha - c(h/2)^\alpha \\ &= ch^\alpha(4^{-\alpha} - 2^{-\alpha}) = ch^\alpha 2^{-\alpha}(2^{-\alpha} - 1). \end{aligned}$$

Folglich gilt

$$\frac{\|u_h - u_{h/2}\|_h}{\|u_{h/2} - u_{h/4}\|_h} \approx 2^\alpha \quad \text{bzw.} \quad \alpha = \frac{\log(\|u_h - u_{h/2}\|_h / \|u_{h/2} - u_{h/4}\|_h)}{\log(2)}$$

b) Die inhärenten Konvergenzordnungen der gegegebenen Folgen sind:

$$\begin{aligned} \alpha &= \frac{\log(\frac{33.627-30.318}{30.318-29.100})}{\log(2)} = \frac{\log(\frac{3.309}{1.218})}{\log(2)} \approx \frac{\log(2.716)}{\log(2)} \approx \frac{0.9994}{0.6931} \approx 1.44 \\ \alpha &= \frac{\log(\frac{30.318-29.100}{29.100-28.586})}{\log(2)} = \frac{\log(\frac{1.218}{0.514})}{\log(2)} \approx \frac{\log(2.369)}{\log(2)} \approx \frac{0.8624}{0.6931} \approx 1.24 \\ \alpha &= \frac{\log(\frac{29.100-28.586}{28.586-28.351})}{\log(2)} = \frac{\log(\frac{0.514}{0.235})}{\log(2)} \approx \frac{\log(2.187)}{\log(2)} \approx \frac{0.7826}{0.6931} \approx 1.12 \end{aligned}$$

und

$$\begin{aligned} \alpha &= \frac{\log(\frac{26.570-27.008}{27.008-27.883})}{\log(2)} = \frac{\log(\frac{-0.438}{-0.875})}{\log(2)} \approx \frac{\log(0.500)}{\log(2)} \approx \frac{-0.6920}{0.6931} \approx -0.998 \\ \alpha &= \frac{\log(\frac{27.008-27.883}{27.883-28.072})}{\log(2)} = \frac{\log(\frac{-0.875}{-0.189})}{\log(2)} \approx \frac{\log(4.629)}{\log(2)} \approx \frac{1.532}{0.6931} \approx 2.21 \\ \alpha &= \frac{\log(\frac{27.883-28.072}{28.072-28.117})}{\log(2)} = \frac{\log(\frac{-0.189}{-0.045})}{\log(2)} \approx \frac{\log(4.2)}{\log(2)} \approx \frac{1.435}{0.6931} \approx 2.07 \end{aligned}$$

Lösung A.2.3: a) Wir verwenden wieder die Greensche Identität

$$v_h(P) = h^2 \sum_{Q \in \Omega_h} G_h(P,Q) L_h v_h(Q) + \sum_{Q \in \partial\Omega_h} G_h(P,Q) v_h(Q)$$

mit der durch

$$L_h G_h(P,Q) = h^2 \delta(P,Q), \quad P \in \Omega_h, \quad G_h(P,Q) = \delta(P,Q), \quad P \in \partial\Omega_h, \quad Q \in \overline{\Omega}_h,$$

definierten diskreten Greenschen Funktion $G_h : \overline{\Omega}_h \times \overline{\Omega}_h \to \mathbb{R}$. Da der kompakte 9-Punkte-Operator konsistent ist und die Bedingungen (B1), (B2) und (B3) erfüllt, gilt für ihn das diskrete Maximumprinzip sowie die Abschätzungen

$$0 \le h^2 \sum_{Q \in \Omega_h} G_h(P,Q) \le \frac{d_\Omega^2}{4}.$$

Anwendung der Greenschen Identität für die Fehlerfunktion $e_h := u - u_h$ ergibt dann wegen $e_h = 0$ auf $\partial\Omega_h$:

$$\begin{aligned}\max_{\overline{\Omega}_h} |e_h| &\le \frac{d_\Omega^2}{4} \max_{\Omega_h} |L_h e_h| = \frac{d_\Omega^2}{4} \max_{\Omega_h} |L_h u - L_h u_h| \\ &= \frac{d_\Omega^2}{4} \max_{\Omega_h} |L_h u - f - \tfrac{1}{12} h^2 \Delta f| \le c\, M_6(u) h^4.\end{aligned}$$

c) Wir können denselben Ansatz wie in Teil (b) verwenden, da auch das modifizierte 9-Punkte-Schema konsistent ist und den Bedingungen (B1), (B2) und (B3) genügt. Ausgehend von der diskreten Greenschen Identität erhalten wir

$$e_h(P) = h^2 \sum_{Q \in \Omega_h^0} G_h(P,Q) L_h e_h(Q) + h^2 \sum_{Q \in \partial\Omega_h^*} G_h(P,Q) L_h e_h(Q)$$

Mit Hilfe der auch hier gültigen Abschätzungen

$$0 \le h^2 \sum_{Q \in \Omega_h} G_h(P,Q) \le \frac{d_\Omega^2}{4}, \qquad \sum_{Q \in \partial\Omega_h^*} G_h(P,Q) \le \frac{1}{2}$$

folgt

$$\begin{aligned}\max_{P \in \overline{\Omega}_h} |e_h| &\le \frac{d_\Omega^2}{4} \max_{Q \in \Omega_h^0} |-\Delta_h^{(9)} u - f_h| + \frac{1}{2} \max_{Q \in \partial\Omega_h^*} |-\Delta_h^* - f| \\ &\le c M_6(u) h^6 + c M_3(u) h^3.\end{aligned}$$

Lösung A.2.4: a) Wir prüfen zunächst die Formel für die Eigenwerte von A_h nach. Es gilt (nachrechnen):

$$A_h w^{\nu\mu} = ...$$

b) Der Darstellung der Eigenwerte von A_h und der Taylor-Entwicklung des Cosinus

$$\cos(x) = 1 - \frac{x^2}{2} + \mathcal{O}(x^4)$$

entnehmen wir, dass

$$\lambda_{\max}(A_h) = h^{-2}\big(4 - 4\cos((1-h)\pi)\big) = h^{-2}\big(4 + 4\cos(h\pi)\big) = 8h^{-2} + \mathcal{O}(1)$$
$$\lambda_{\min}(A_h) = h^{-2}\big(4 - 4\cos(h\pi)\big) = h^{-2}\big(4 - 4 + 2\pi^2h^2 + \mathcal{O}(h^4)\big)$$

und somit

$$\mathrm{cond}_2(A_h) =: \frac{\lambda_{\max}(A_h)}{\lambda_{\min}(A_h)} = \frac{8 + \mathcal{O}(h^2)}{2\pi^2h^2 + \mathcal{O}(h^4)} = \frac{4}{\pi^2h^2}\frac{1 + \mathcal{O}(h^2)}{1 + \mathcal{O}(h^2)} = \frac{4}{\pi^2h^2} + \mathcal{O}(1).$$

Lösung A.2.5: a) Seien $x, y \in \mathbb{R}^N$ mit der Eigenschaft $x \geq y$ gegeben. Dann gilt für beliebige nicht-negative Zahlen $a_{ij}^{(-1)}$, dass:

$$\sum_j a_{ij}^{(-1)} x_i \geq \sum_j a_{ij}^{(-1)} y_i.$$

Angewendet auf die Multiplikation mit A^{-1} folgt:

$$Av \geq Aw \quad \Rightarrow \quad A^{-1}Av \geq A^{-1}Aw \quad \Rightarrow \quad v \geq w$$

und somit die Behauptung.

b) Für den Vektor $w \in \mathbb{R}^N$ sei $A_h w \geq (1, \ldots, 1)^T$. Die Matrix A_h ist invers-monoton, d.h. $A_h^{-1} \geq 0$. Also ist

$$w = A_h^{-1}A_h w \geq A_h^{-1}(1, \ldots, 1)^T$$

d. h.: Der Vektor w ist komponentenweise größer oder gleich den jeweiligen (nicht-negativen) Zeilensummen von A_h^{-1}. Folglich gilt für die Maximale-Zeilensummen-Norm von A_h^{-1}:

$$\|A_h^{-1}\|_\infty \leq \|w\|_\infty.$$

c) Für die Maximale-Zeilensummen-Norm der Matrix A_h gilt offenbar $\|A_h\|_\infty \leq 8h^{-2}$. Für die Funktion $w = x(1-x)/4 + y(1-y)/4$ gilt

$$-\Delta_h^{(5)} w = -\Delta w = 1.$$

Dies ist gleichbedeutend mit $A_h w \geq (1, \ldots, 1)^T$. Nach Teil (b) folgt daraus

$$\|A_h^{-1}\|_\infty \leq \|w\|_\infty = \frac{1}{8}$$

Dies impliziert

$$\mathrm{cond}_\infty(A_h) := \|A_h\|_\infty \|A_h^{-1}\|_\infty \leq h^{-2}.$$

Lösung A.2.6: Wir bezeichnen die Anzahl der inneren Gitterpunkt mit der y-Koordinate h mit n. Insgesammt gibt es dann gerade $N = \frac{1}{2}(n^2 + n)$ innere Gitterpunkte.

a) Bei zeilenweiser Nummerierung, erhalten wir folgende Nummerierung

$$\begin{array}{ccccc} N & & & & \\ N-2 & N-1 & & & \\ \vdots & & \ddots & & \\ n+1 & n+2 & \cdots & 2n-1 & \\ 1 & 2 & \cdots & n-1 & n \end{array}$$

dabei ist die Differenz der Nummer zwischen zwei benachbarten Punkten kleiner oder gleich n. Man beachte, das dieser Abstand für Punkte mit grösseren Nummern kleiner wird. Die Systemmatrix hat damit die Form

$$\begin{pmatrix} B_n & -I_{n-1} & 0 & \cdots & 0 \\ -I_{n-1} & B_{n-1} & -I_{n-2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -I_2 & B_2 & -I_1 \\ 0 & \cdots & 0 & -I_1 & B_1 \end{pmatrix}.$$

Dabei sind

$$I_i = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \in \mathbb{R}^{i\times i} \quad B_i = \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix} \in \mathbb{R}^{i\times i}$$

und an der oberen linken Ecke sind entsprechen 0-en zu ergänzen. Damit ist der Speicheraufwand $O(Nn) = O(n^3)$ und der Rechenaufwand ist $O(Nn^2) = O(N^2)$.

b) Mit der Nummerierung

$$\begin{array}{ccccc} N-n+1 & & & & \\ \vdots & \ddots & & & \\ 4 & & \ddots & & \\ 2 & 5 & & N-1 & \\ 1 & 3 & 6 & \cdots & N \end{array}$$

erhalten wir die folgende Systemmatrix:

$$\begin{pmatrix} 4I_1 & M_1 & & & \\ M_1^t & 4I_2 & \ddots & & \\ & M_2^t & \ddots & \ddots & \\ & & \ddots & 4I_{n-1} & M_{n-1} \\ & & & M_{n-1}^t & 4I_n \end{pmatrix}.$$

Wobei M_i durch

$$M_i = \begin{pmatrix} -1 & -1 & & \\ & \ddots & \ddots & \\ & & -1 & -1 \end{pmatrix} \in \mathbb{R}^{i\times(i+1)}$$

gegeben ist. Insgesammt liegen zwischen den Nummern der 4 benachbarten Felder höchstens die Nummern entlang einer Diagonalen, die Bandbreite ist also $n+1$. Damit ist entsprechend zu Teil (a) der Speicheraufwand $O(Nn) = O(n^3)$ und der Rechenaufwand ist durch $O(Nn^2) = O(N^2)$ gegeben.

c) Im Falle der schachbrettartigen Nummerierung liegen zwischen den Indizes benachbarter Elemente höchstens $N/2 + n/2$ Zahlen. Die entstehende Systemmatrix hat also die Bandbreite $M/2 + n/2$. Der Speicherbedarf ist demnach $O(NN) = O(n^4)$ und der Rechenaufwand $O(NN^2) = O(N^3)$. Die Systemmatrix hat die Form

$$\begin{pmatrix} 4I_{\frac{N}{2}} & * \\ * & 4I_{\frac{N}{2}} \end{pmatrix}$$

wobei $*$ eine Matrix mit höchstens vier von Null verschiedenen Eintraegen pro Zeile ist. Wir erkennen, das die Zerlegung nur auf den mit $*$ gekennzeicheten Einträgen operieren muss, so dass sich der Rechenaufwand auf $O(n^5)$ verringert.

Lösung A.2.7: a) Die Eigenwerte der Iterationsmatrix $B_\theta = I - \theta A$ sind $\mu = 1 - \theta\lambda$ und folglich $\mu_{\max} = 1 - \theta\lambda_{\min}$ sowie $\mu_{\min} = 1 - \theta\lambda_{\max}$ bzw.

$$\rho(B_\theta) = \max\{|\mu|\} = \max\{|1 - \theta\lambda_{\max}|, |1 - \theta\lambda_{\min}|\}.$$

b) Im Falle $0 < \lambda_{\min} \leq \lambda_{\max}$ gilt

$$0 < \theta < \frac{2}{\lambda_{\max}} \quad \Leftrightarrow \quad 0 < \theta\lambda_{\min} \leq \theta\lambda_{\max} < 2.$$

Dies wiederum ist äquivalent mit $\max\{|1 - \theta\lambda_{\min}|, |1 - \theta\lambda_{\max}|\} < 1$.

c) Ein einfaches geometrische Argument ergibt

$$\theta_{\text{opt}} = \frac{2}{\lambda_{\min} + \lambda_{\max}}.$$

Lösung A.2.8: Durch Nachrechnen erhalten wir das für $1 \leq k, l \leq N$ die Funktionen

$$\varphi^{kl} := \sin\left(\frac{k\pi x}{N+1}\right) \sin\left(\frac{l\pi x}{N+1}\right)$$

Eigenfunktionen von A_x zum Eigenwert $\lambda_x^{kl} = 2 - 2\cos\left(\frac{k\pi}{N+1}\right) = 4\sin^2\left(\frac{k\pi}{2(N+1)}\right)$, und Eigenfunktionen von A_y zum Eigenwert $\lambda_y^{kl} = 2 - 2\cos\left(\frac{l\pi}{N+1}\right) = 4\sin^2\left(\frac{l\pi}{2(N+1)}\right)$ sind. Wir haben also N^2 Eigenfunktionen von A_x bzw. A_y aus Dimensionsgründen sind diese gerade ein vollständiges System von Eigenvektoren. Insbesondere ist $A_x A_y = A_y A_x$. Die zum ADI-Verfahren gehöhrige Iterationsmatrix hat gerade die Form

$$B_\sigma = (\sigma + A_y)^{-1} (\sigma - A_y) (\sigma + A_x)^{-1} (\sigma - A_x)$$

und hat die Eigenwerte

$$\lambda_\sigma^{kl} = \frac{\sigma - \lambda_y^{kl}}{\sigma + \lambda_y^{kl}} \frac{\sigma - \lambda_x^{kl}}{\sigma + \lambda_x^{kl}}.$$

Diese sind für $\sigma > 0$ wohldefiniert, da die Eigenwerte von A_x und A_y positiv sind. Nun ist

$$\frac{|\sigma - c|}{|\sigma + c|} < 1$$

für jedes feste $c > 0$ und beliebieges $\sigma > 0$, so dass für den Spektralradius von B_σ gilt:

$$\rho(B_\sigma) < 1.$$

Das ADI-Verfahren ist also unabhängig von $\sigma > 0$ konvergent. Die Konvergenz ist am schnellsten falls $\rho(B_\sigma)$ minimal ist. Wir suchen also

$$\min_{\sigma>0} \max_{k,l} \left| \frac{\sigma - \lambda_y^{kl}}{\sigma + \lambda_y^{kl}} \frac{\sigma - \lambda_x^{kl}}{\sigma + \lambda_x^{kl}} \right|.$$

Da λ_x^{kl} von l unabhängig und λ_y^{kl} von k unabhängig ist, sowie $\lambda_x^{kl} = \lambda_y^{lk}$ ist

$$\max_{k,l} \left| \frac{\sigma - \lambda_y^{kl}}{\sigma + \lambda_y^{kl}} \frac{\sigma - \lambda_x^{kl}}{\sigma + \lambda_x^{kl}} \right| = \max_k \frac{|\sigma - \lambda_x^{k1}|^2}{|\sigma + \lambda_x^{k1}|^2}$$

Wir betrachten jetzt den Graphen der Funktion $x \mapsto \frac{|\sigma - x|}{\sigma + x}$, und erkennen, dass das Maximum der Beziehung

$$\max_{\lambda_x^1 \leq x \leq \lambda_x^N} \frac{|\sigma - x|^2}{|\sigma + x|^2} = \max\left(\frac{|\sigma - \lambda_x^1|^2}{|\sigma + \lambda_x^1|^2}, \frac{|\sigma - \lambda_x^N|^2}{|\sigma + \lambda_x^N|^2}\right)$$

genügt. Damit erhalten wir, dass für den optimalen Parameter die Beziehung

$$\frac{|\sigma - \lambda_x^1|^2}{|\sigma + \lambda_x^1|^2} = \frac{|\sigma - \lambda_x^N|^2}{|\sigma + \lambda_x^N|^2}$$

gilt. Aus Monotoniegründen kann dies nur für $\sigma \in \left[\lambda_x^1, \lambda_x^N\right]$ der Fall sein. Wir erhalten:

$$\begin{aligned} & \frac{\sigma - \lambda_x^1}{\sigma + \lambda_x^1} = \frac{\lambda_x^N - \sigma}{\sigma + \lambda_x^N} \\ \Leftrightarrow \quad & \sigma^2 - \sigma\lambda_x^1 + \sigma\lambda_x^N - \lambda_x^1\lambda_x^N = -\sigma^2 - \sigma\lambda_x^1 + \sigma\lambda_x^N + \lambda_x^1\lambda_x^N \\ & \Leftrightarrow \quad \sigma = \sqrt{\lambda_x^1\lambda_x^N}. \end{aligned}$$

Im speziellen Fall erhalten wir also

$$\sigma = 4\sin\left(\frac{\pi}{2(N+1)}\right)\sin\left(\frac{N\pi}{2(N+1)}\right)$$

und mit diesem σ ist

$$\rho\left(B_\sigma\right) = \frac{\cos^2\left(\frac{\pi}{N+1}\right)}{\left(1+\sin\left(\frac{\pi}{N+1}\right)\right)^2}$$

A.3 Kapitel 3

Lösung A.3.1: a) Wir suchen zunächst eine geeignete schwache Formulierung. Durch Multiplikation mit einer Funktion $\varphi \in V = H^1$, Integration über Ω, und anschließende partielle Integration erhalten wir:

$$(\nabla u, \nabla\varphi)_\Omega - \langle \partial_n u, \varphi\rangle_{\partial\Omega} = (f, \varphi)_\Omega \quad \forall\varphi \in V.$$

Indem wir die Neumann-Randwerte einsetzen, folgt die schwache Formulierung:

$$(\nabla u, \nabla\varphi)_\Omega = (f, \varphi)_\Omega + \langle g, \varphi\rangle_{\partial\Omega} \quad \forall\varphi \in V.$$

Wir wissen bereits, dass obiges Problem nicht eindeutig lösbar ist, wir können also auch keine Bestapproximationseigenschaft erwarten. Fordern wir zusätzlich, dass $\int_\Omega u = 0$ für alle $u \in V$, um eine eindeutige Lösung zu erhalten, so ist aufgrund der Poincaréschen Ungleichung für $u \in V$

$$\|u\|_{H^1} \leq c\sqrt{(\nabla u, \nabla u)_\Omega}.$$

Also ist $\sqrt{(\nabla u, \nabla u)_\Omega}$ eine zu $\|u\|_{H^1}$ äquivalente Norm auf V. Sei nun V_h ein endlichdimensionaler Teilraum von V, dann lautet das Ritz-Verfahren gegeben durch

$$(\nabla u_h, \nabla\varphi_h)_\Omega = (f, \varphi_h)_\Omega + \langle g, \varphi_h\rangle_{\partial\Omega} \quad \forall\varphi_h \in V_h.$$

Wir erhalten durch Galerkinorthogonalität für beliebiges $v_h \in V_h$:

$$\begin{aligned} \left(\nabla\left(u-u_h\right), \nabla\left(u-u_h\right)\right)_\Omega &= \left(\nabla\left(u-u_h\right), \nabla\left(u-v_h\right)\right)_\Omega + \left(\nabla\left(u-u_h\right), \nabla\left(v_h-u_h\right)\right)_\Omega \\ &= \left(\nabla\left(u-u_h\right), \nabla\left(u-v_h\right)\right)_\Omega \\ &\leq \left\|\nabla\left(u-u_h\right)\right\|_\Omega \left\|\nabla\left(u-v_h\right)\right\|_\Omega. \end{aligned}$$

Und somit, da $v_h \in V_h$ beliebig war,

$$\|\nabla (u - u_h)\|_\Omega \leq \inf_{v_h \in V_h} \|\nabla (u - v_h)\|_\Omega .$$

b) Um eine variationelle Formulierung zu erhalten, betracheten wir den Rayleigh-Quotienten

$$R(u) = \frac{(\nabla u, \nabla u)_\Omega}{(u,u)_\Omega}$$

Es kann gezeigt werden, dass $R(v) \geq \lambda$, falls λ der kleinste Eigenwert des Laplace-Operators ist. Für einen Eigenvektor u zum Eigenwert λ ist $R(u) = \lambda$. Das Eigenwertproblem zum kleinsten Eigenwert lässt sich, mit $V := H_0^1$, schreiben als

$$\min_{u \in V} R(u) =: \lambda.$$

Entsprechend ist das Rayleigh-Ritz-Verfahren mit einem endlich-dimensionalen Teilraum $V_h \subset V$

$$\min_{u_h \in V_h} R(u_h) =: \lambda_h.$$

Wir sehen hierhaus sofort:

$$\lambda \leq \lambda_h$$

Für einen beliebigen Eigenwert λ_l gilt das min-max-Prinzip

$$\lambda_l = \min_{S_l \subset V, \dim(S_l)=l} \max_{v \in S_l} R(v)$$

sowie im Endlichdimensionalen

$$\lambda_l^h = \min_{S_l \subset V_h, \dim(S_l)=l} \max_{v_h \in S_l} R(v_h).$$

Mit Mitteln, die über den Stoff dieses Textes hinausgehen, zeigt man für einen konformen Finite-Elemente-Ansatz der Ordnung k die Fehlerabschätzung

$$|\lambda_l - \lambda_l^h| \leq C h^{2k} \lambda_l^k.$$

c) Wir erhalten wieder durch Multiplikation mit einer Funktion v und partieller Integration

$$\begin{aligned} \left(\Delta^2 u, v\right) &= -(\nabla \Delta u, \nabla v) + \langle \partial_n \Delta u, v \rangle \\ &= (\Delta u, \Delta v) + \langle \partial_n \Delta u, v \rangle - \langle \Delta u, \partial_n v \rangle \\ &= (\Delta u, \Delta v), \end{aligned}$$

und somit als schwache Formulierung

$$(\Delta u, \Delta v) = (-f, v) \quad \forall v \in H_0^2(\Omega).$$

Entsprechend ist das Ritzverfahren mit $V_h \subset H_0^2(\Omega)$ endlichdimensional:

$$(\Delta u_h, \Delta v_h) = (-f, v_h) \quad \forall v_h \in V_h.$$

Um die Bestapproximationseigenschaft zu erhalten benötigen wir Stetigkeit und Koerzitivität von $(\Delta\cdot, \Delta\cdot)$ auf $H_0^2(\Omega)$. Die Stetigkeit ist klar. Um die Koerzitivität zu zeigen, beachten wir, dass aufgrund der Poincaréschen Ungleichung

$$\|u\|_{H^2} \leq c\,|u|_{H^2}$$

gilt. Um nun die 2. Ableitungen durch den Laplace-Operator abzuschätzen betrachten wir das Vektorfeld

$$w = \begin{pmatrix} u_1 u_{22} \\ -u_1 u_{12} \end{pmatrix}.$$

Es gilt nach dem Gaussschen-Integralsatz unter Berücksichtigung der Randwerte

$$\int_\Omega u_{11}u_{22} - u_{12}^2 = \int_\Omega \nabla \cdot w = \int_{\partial\Omega} n \cdot w = 0.$$

Zusammengenommen ergibt sich

$$(\Delta u, \Delta u) = \int_\Omega |\Delta u| - 2\left(u_{11}u_{22} - u_{12}^2\right) = \int_\Omega u_{11}^2 + u_{22}^2 + u_{12}^2 = |u|_{H^2}^2$$

und somit die Koerzivität.

Lösung A.3.2: a) Die Galerkin-Gleichungen entsprechen einem quadratischen linearen Gleichungssystem der Dimension $N = \dim(V_h)$ für die Entwicklungskoeffizienten von u_h bzgl. einer beliebigen Basis von V_h. Zum Nachweis der eindeutigen Lösbarkeit genügt also der Nachweis der Eindeutigkeit. Für zwei Lösungen $u_h^{(1)}, u_h^{(2)}$ erfüllt die Differenz $w_h := u_h^{(1)} - u_h^{(2)}$ die Gleichung

$$a(w_h, \varphi_h) = 0 \quad \forall \varphi_h \in V_h.$$

Bei Wahl von $\varphi_h = w_h$ folgt daher mit Hilfe der V-Elliptizität

$$0 = |a(w_h, w_h)| \geq \kappa \|w_h\|^2 \quad \Rightarrow \quad w_h = 0.$$

Zum Nachweis der Fehleranschätzung verwenden wir zunächst die V-Elliptizität, dann die Galerkin-Orthogonalität mit einem beliebigen $\varphi_h \in V_h$ und schließlich die Beschränktheit:

$$\|u - u_h\|^2 \leq \frac{1}{\kappa}|a(u - u_h, u - u_h)| = \frac{1}{\kappa}|a(u - u_h, u - \varphi_h)| \leq \frac{\alpha}{\kappa}\|u - u_h\|\|u - \varphi_h\|.$$

Dies impliziert offensichtlich die behauptete Ungleichung.

bi) Wir diskutieren zunächst die Lösbarkeit der Variationsaufgabe

$$a(u,\varphi) = l(\varphi) \quad \forall \varphi \in V,$$

unter der Voraussetzung, daß $a(\cdot,\cdot)$ folgende Koerzitivitätseigenschaften besitzt:

$$\sup_{\varphi\in V} \frac{a(v,\varphi)}{\|\varphi\|} \geq \gamma\|v\|, \qquad \sup_{\varphi\in V} \frac{a(\varphi,v)}{\|\varphi\|} \geq \gamma\|v\|, \quad v \in V.$$

Dies bedeutet, daß $a(\cdot,\cdot)$ und die durch $a^*(v,\varphi) := a(\varphi,v)$ definierte zugehörige „adjungierte" Bilinearform koerzitiv sind. Die Bilinearformen $a(\cdot,\cdot)$ und $a^*(\cdot,\cdot)$ definieren mit Hilfe des Rieszschen Darstellungssatzes durch

$$(Av,\varphi) = a(v,\varphi), \quad (A^*v,\varphi) = a^*(v,\varphi) = (\varphi,v), \quad \varphi \in V,$$

lineare Operatoren $A : V \to V$ und $A^* : V \to V$. Diese sind wegen der Koerzitivität der definierenden Bilinearformen injektiv:

$$Av = 0 \quad \Rightarrow \quad 0 = \sup_{\varphi\in V} \frac{a(v,\varphi)}{\|\varphi\|} \geq \gamma\|v\| \quad \Rightarrow \quad v = 0,$$

und analog für A^*. Wir definieren die symmetrische und positive Bilinearform

$$[v,\varphi] := (A^*v, A^*\varphi), \quad v,\varphi \in V.$$

Mit der zugehörigen Norm $|v| := [v,v]^{1/2} = \|A^*v\|$ gilt dann:

$$\gamma\|v\| \leq \sup_{\varphi\in V} \frac{a(\varphi,v)}{\|\varphi\|} = \sup_{\varphi\in V} \frac{(\varphi,A^*v)}{\|\varphi\|} \leq \|A^*v\| = |v| \leq \gamma'\|v\|$$

mit

$$\gamma' := \sup_{\varphi\in V} \frac{\|A^*\varphi\|}{\|\varphi\|}.$$

Folglich definiert $[\cdot,\cdot]$ ein Skalarprodukt auf V, welches zu dem auf V gegebenen Skalarprodukt $(\cdot,\cdot)$ äquivalent ist. Nach dem Rieszschen Darstellungssatz existiert daher zu jedem linearen Funktional $l \in V^*$ ein $w \in V$, so dass

$$l(\varphi) = [\varphi,v] = (A^*\varphi, A^*w) = (\varphi, AA^*w) = (\varphi, Av) \quad \forall \varphi \in V,$$

mit $v := A^*w$. Also ist v (eindeutige) Lösung der obigen Variationsaufgabe.

(bii) Die Koerzitivität von $a(\cdot,\cdot)$ auf V reicht allein nicht aus, um die gewünschten Konvergenzaussagen zu erhalten. Setzt man aber voraus, daß $a(\cdot,\cdot)$ auch auf den Teilräumen $V_h \subset V$ koerzitiv ist und zwar gleichmäßig bzgl. h,

$$\sup_{\psi_h\in V_h} \frac{a(v_h,\psi_h)}{\|\psi_h\|} \geq \gamma_h\|v_h\|, \quad v_h \in V_h, \qquad \gamma_h \geq \gamma_0 > 0,$$

so lassen sich die obigen Aussagen beweisen. Mit beliebigem $\varphi_h \in V_h$ gilt:

$$\begin{aligned}\|u-u_h\| &\le \|u-\varphi_h\| + \|\varphi_h - u_h\| \\ &\le \|u-\varphi_h\| + \frac{1}{\gamma_0}\sup_{\psi_h\in V_h}\frac{a(\varphi_h-u_h,\psi_h)}{\|\psi_h\|} \\ &\le \|u-\varphi_h\| + \frac{1}{\gamma_0}\sup_{\psi_h\in V_h}\frac{a(\varphi_h-u,\psi_h)}{\|\psi_h\|} + \frac{1}{\gamma_0}\sup_{\psi_h\in V_h}\frac{a(u-u_h,\psi_h)}{\|\psi_h\|} \\ &\le \Big(1+\frac{\alpha}{\gamma_0}\Big)\|\varphi_h - u\|.\end{aligned}$$

Der Nachweis der der „diskreten Koerzitivität“ erfordert spezielle Bedingungen an die Bilinearform $a(\cdot,\cdot)$ und den Ansatzraum V_h. Zum Beispiel ist die Bilinearform

$$a(\cdot,\cdot) = (\nabla\cdot,\nabla\cdot) - \mu(\cdot,\cdot),$$

wenn μ kein Eigenwert des Laplace-Operators ist, koerzitiv aber nicht V-elliptisch. Die Gültigkeit der (gleichmäßigen) diskreten Koerzitivität folgt dann über ein Widerspruchsargument mit Hilfe der Kompaktheit der Einbettung $H_0^1(\Omega) \subset L^2(\Omega)$. Angenommen, die Bilinearform $a(\cdot,\cdot)$ ist nicht gleichmäßig koerzitiv auf V_h. Dann existieren eine Folge von Gitterweiten $(h_k)_{k\in\mathbb{N}}$ und Funktionen $v_k \in V_k := V_{h_k}$ mit den Eigenschaften

$$\|\nabla v_k\| = 1, \quad \sup_{\varphi_k\in V_k}\frac{a(v_k,\varphi_k)}{\|\nabla\varphi_k\|} < \frac{1}{k}, \quad k\in\mathbb{N}.$$

Wegen der Kompaktheit der Einbettung $H_0^1(\Omega) \subset L^2(\Omega)$ existieren für die H^1-beschränkte Folge $(v_k)_{k\in\mathbb{N}}$ eine Teilfolge $(v_{k'})_{k'\in\mathbb{N}}$ und ein $v \in L^2(\Omega)$ mit

$$\|v_{k'} - v\| \to 0 \quad (k'\to\infty).$$

Wegen der schwachen Kompaktheit der Einheitskugel in $H_0^1(\Omega)$ kann o.B.d.A. erreicht werden, daß die Folge $(v_{k'})_{k'\in\mathbb{N}}$ auch schwach in $H_0^1(\Omega)$ gegen v konvergiert, d.h.: $v \in H_0^1(\Omega)$ und

$$(\nabla v_{k'},\nabla\varphi) \to (v,\varphi), \quad \varphi\in H_0^1(\Omega).$$

Dies impliziert zunächst für beliebiges $\varphi \in H_0^1(\Omega)$:

$$|(\nabla v,\nabla\varphi) - \mu(v,\varphi)| = \lim_{k'\to\infty}|(\nabla v_{k'},\nabla\varphi) - \mu(v_{k'},\varphi)| \le 0,$$

und folglich $v = 0$, da μ nach Voraussetzung μ kein Eigenwert ist. Dies impliziert dann

$$\frac{(\nabla v_{k'},\nabla v_{k'}) - \mu(v_{k'},v_{k'})}{\|\nabla v_{k'}\|} \le \sup_{\varphi'_k\in V_{k'}}\frac{(\nabla v_{k'},\nabla\varphi_{k'}) - \mu(v_{k'},\varphi_{k'})}{\|\nabla\varphi_{k'}\|} \to 0 \quad (k'\to\infty)$$

und folglich $\|\nabla v_{k'}\| \to 0$ $(k'\to\infty)$ im Widerspruch zur Annahme $\|\nabla v_k\| = 1$.

Lösung A.3.3: a) Exakte Integration ergibt die Werte

$$a_{ij} = (\nabla\varphi_h^i, \nabla\varphi_h^j) = \begin{cases} \frac{8}{3}, & j = i, \\ -\frac{1}{3}, & j \in \{i \pm 1, i \pm m, i \pm m \pm 1\}, \\ 0, & \text{sonst.} \end{cases}$$

b) Verwendung der 2-dimensionalen "Tensorprodukt-Trapezregel" ergibt:

$$a_{ij} = (\nabla\varphi_h^i, \nabla\varphi_h^j) = \begin{cases} 4, & j = i, \\ -1, & j \in \{i \pm 1, i \pm m\}, \\ 0, & \text{sonst.} \end{cases}$$

Der Finite-Elemente-Ansatz mit bilinearen Formfunktionen ergibt bei Verwendung der Trapezregel bis auf den Vorfaktor h^{-2} dieselbe Systemmatrix wie der 5-Punkte-Differenzenoperator.

Lösung A.3.4: Wir betrachten zunächst Dreicke. Dann sind aufgrund des Zusammenhangs

$$\rho_T = h_T \sin(\alpha) \sin(\beta) \sin(\gamma)$$

die Bedingungen (a) und (b) äquivalent.

i) Sei nun a) erfüllt. Dann sind ferner alle Winckel gleichmässig von π wegbeschränkt. Somit gilt für jeden Winkel α

$$\sin(\alpha) \geq c > 0$$

und somit

$$\frac{\sin(\alpha)}{\sin(\beta)} \leq c \sin(\alpha) \leq c$$

Aus dem Sinussatz

$$\frac{\sin(\alpha)}{\sin(\beta)} = \frac{a}{b}$$

folgt also Bedingung (c). Die Umkehrung gilt nicht, wie man sich anhand der Abb. A.1 klar macht.

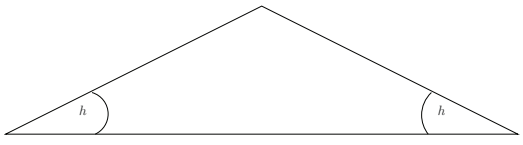

Abbildung A.1: Gegenbeispiel zur Äquivalenz der Formregularität bei Dreiecken

ii) Bei Vierecken sind zusätzlich die Implikationen $(a) \Rightarrow (b), (c)$ sowie $b \Rightarrow (c)$ falsch, wie man sich leicht anhand des ersten bzw. zweiten Rechtecks in Abb. A.2 überlegt.

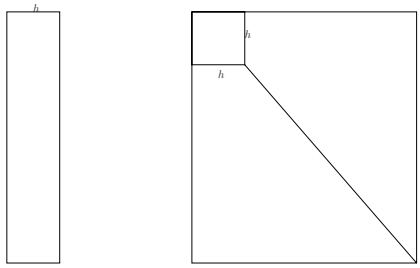

Abbildung A.2: Gegenbeispiele zur Äquivalenz der Formregularität bei Vierecken

Lösung A.3.5: a) Mit beliebigem $\varphi_h \in V_h^{(1)}$ gilt

$$\|u - P_h u\|^2 = (u - P_h u, u - P_h u) = (u - P_h u, u - \varphi_h) \leq \|u - P_h u\| \|u - \varphi_h\|$$

und folglich

$$\|u - P_h u\| = \min_{\varphi_h \in V_h^{(1)}} \|u - \varphi_h\|.$$

Mit Hilfe der Abschätzung für die Knoteninterpolierende $I_h u \in V_h^{(1)}$ für $u \in H^2(\Omega)$,

$$\|u - I_h u\| \leq ch^2 \|\nabla^2 u\|,$$

folgt $\|u - P_h u\| \leq ch^2 \|\nabla^2 u\|$. Im Falle der Minimalregularität $u \in L^2(\Omega)$ kann keine h-Potenz in der Fehlerabschätzung erwartet werden; es gilt aber:

$$\|u - P_h u\| = \min_{\varphi_h \in V_h^{(1)}} \|u - \varphi_h\| \leq \|u\|.$$

Für $u \in H^1(\Omega)$ ist in mehr als einer Dimension die Knoteninterpolierende $I_h u$ gar nicht definiert. In diesem Fall muß mit einer modifizierten „Quasi-Interpolierenden" $\tilde{I}_h : H^1(\Omega) \to V_h^{(1)}$ gearbeitet werden, welche H^1-stabil ist. Dazu ordnen wir jedem Gitterknoten a den umgebenden Zellbereich $S(a) = \cup\{T \in \mathbb{T}_h,\ a \in T\}$ zu und setzen

$$\tilde{I}_h u(a) := |S(a)|^{-1} \int_{S(a)} u(x)\, dx.$$

Für die so definierte Funktion $\tilde{I}_h u \in V_h^{(1)}$ gilt dann die Fehlerabschätzung

$$\|u - \tilde{I}_h u\| \leq ch \|u\|_1.$$

Den recht komplizierten Beweis können wir hier nicht geben.

Alternativ kann man im Falle eines regulären Gebiets (glattberanded oder konvexes Polygon/Polyeder) auch die Ritz-Projektion $P_h u \in S_h^{(1)}$ verwenden. Für diese gilt mit der zugehörigen „dualen Lösung" $z \in H_0^1(\Omega) \cap H^2(\Omega)$ von $-\Delta z = \|u - R_h\|^{-1}(u - R_h u)$:

$$\begin{aligned}\|u - R_h u\| &= (\nabla(u - R_h u), \nabla z) = (\nabla(u - R_h u), \nabla(z - R_h z)) = (\nabla u, \nabla(z - R_h z)) \\ &\le \|\nabla u\| \|\nabla(z - R_h z)\| \le \|\nabla u\| \|\nabla(z - I_h z)\| \\ &\le c_i h \|\nabla u\| \|\nabla^2 z\| \le c_i c_s \|\nabla u\|.\end{aligned}$$

Damit ergibt sich dann

$$\|u - P_h u\| \le \|u - R_h u\| \le c_i c_s h \|\nabla u\|.$$

b) Für die negative Sobolew-Norm erhält man mit Hilfe der obigen Abschätzung für den Interpolationsfehler:

$$\begin{aligned}\|u - P_h u\|_{-1} &= \sup_{\varphi \in H_0^1(\Omega)} \frac{(u - P_h, \varphi)}{\|\nabla \varphi\|} = \sup_{\varphi \in H_0^1(\Omega)} \frac{(u - P_h, \varphi - I_h \varphi)}{\|\nabla \varphi\|} \\ &\le \|u - P_h u\| \sup_{\varphi \in H_0^1(\Omega)} \frac{\|\varphi - I_h \varphi\|}{\|\nabla \varphi\|} \le c_2 c_1 h^3 \|\nabla^2 u\|.\end{aligned}$$

Lösung A.3.6: In einer Dimension, d. h. $\Omega = (0,1)$, gilt mit der Knoteninterpolierenden $I_h u \in V_h^{(1)}$ für beliebiges $\varphi_h \in V_h^{(1)}$:

$$\begin{aligned}(I_h u', \varphi_h') &= \sum_{i=1}^{N+1} \int_{x_{i-1}}^{x_i} I_h u' \varphi_h' \, dx = \sum_{i=1}^{N+1} \Big\{ - \int_{x_{i-1}}^{x_i} I_h u \varphi_h'' \, dx + I_h u \varphi_h' \Big|_{x_{i-1}}^{x_i} \Big\} \\ &= \sum_{i=1}^{N+1} \Big\{ - \int_{x_{i-1}}^{x_i} u \varphi_h'' \, dx + u \varphi_h' \Big|_{x_{i-1}}^{x_i} \Big\} = \sum_{i=1}^{N+1} \int_{x_{i-1}}^{x_i} u' \varphi_h' \, dx = (u', \varphi_h').\end{aligned}$$

In diesem Fall stimmen also Ritz-Projektion $R_h u$ und Knoteninterpolierende $I_h u$ überein. Für die Interpolierende gilt nun i. Allg.:

$$\|u - I_h u\|_{-1} = \sup_{\varphi \in H_0^1(\Omega)} \frac{(u - I_h u, \varphi)}{\|\varphi'\|} \ge c h^2.$$

Lösung A.3.7: a) Mit Hilfe der Hölderschen Ungleichung folgt

$$\Big| \int_\Omega (u - u_h) \omega \, dx \Big| \le \|\omega\| \, \|u - u_h\|,$$

d. h.: Die abzuschätzende Größe ist durch die L^2-Norm des Fehlers beschränkt. Also gilt nach einem Resultat der Vorlesung:

$$\Big| \int_\Omega (u - u_h) \omega \, dx \Big| \le c h^2 \|\nabla^2 u\| \le c h \|f\|.$$

b) Zunächst gilt:

$$\Big| \int_\Gamma (u^2 - u_h^2) \, ds \Big| = \Big| \int_\Gamma (u - u_h)(u + u_h) \, ds \Big| \le \Big(\int_\Gamma |u - u_h|^2 \, ds \Big)^{1/2} \Big(\int_\Gamma |u + u_h|^2 \, ds \Big)^{1/2}.$$

Für den zweiten Faktor folgt mit Hilfe der L^2-Spurabschätzung aus der Vorlesung

$$\Big(\int_\Gamma |u+u_h|^2\,ds\Big)^{1/2} \le c\Big(\int_\Omega |u+u_h|^2\,ds\Big)^{1/2} + \Big(\int_\Omega |\nabla(u+u_h)|^2\,ds\Big)^{1/2} \le c.$$

Für den ersten Faktor erhalten wir mit der L^1-Spurabschätzung angewendet für $(u-u_h)^2$:

$$\begin{aligned}\int_\Gamma |u-u_h|^2\,ds &\le c\int_\Omega |u-u_h|^2\,dx + c\int_\Omega |\nabla(u-u_h)^2|\,dx\\ &\le c\int_\Omega |u-u_h|^2\,dx + c\int_\Omega |u-u_h|\,|\nabla(u-u_h)|\,dx\\ &\le c\int_\Omega |u-u_h|^2\,dx + c\Big(\int_\Omega |u-u_h|^2\,dx\Big)^{1/2}\Big(\int_\Omega |\nabla(u-u_h)|^2\,dx\Big)^{1/2}\end{aligned}$$

Die L^2-Fehlerabschätzungen aus der Vorlesung ergeben also

$$\Big|\int_\Gamma (u^2-u_h^2)\,ds\Big| \le ch^{3/2}\|\nabla^2 u\| \le ch^{3/2}\|f\|.$$

Lösung A.3.8: ai) T kartesisches Einheitsdreieck:

$$P(T)=P_3(T),\quad p(a_i),\ \nabla p(a_i),\ p(z), \qquad P(T)=P_3(T),\quad p(a_i),\ p(b_{ij}),\ p(z).$$

Sei $p\in P(T)$, welches bzgl. aller Knotenfunktionale verschwindet. Dann ist $p_{|\Gamma_i}\equiv 0$. Auf dem Einheitsdreieck ist folglich $p(x,y)=cxy(x-y)$, da jedes entlang von ∂T verschwindende Polynom die drei Faktoren x, y und $x-y$ enthalten muß. Wegen $p(z)=0$ folgt notwendig $c=0$, d. h. $p\equiv 0$.

aii) Ein Polynom $p\in P_5(T)$ mit $p(a_i)=0,\ \nabla p(a_i)=0,\ \nabla^2 p(a_i)=0,\ \partial_n p(m_i)=0$ hat auf dem Einheitsdreieck notwenig die Gestalt $p(x,y)=xy(1-x-y)q(x,y)$ mit einem $q\in P_2(T)$. Wegen $\nabla^2 p(0,0)=0$ gilt:

$$\begin{aligned}0&=\partial_x\partial_y p(0,0)\\ &=\big(\partial_x\partial_y(xy(1-x-y))q+\partial_y(xy(1-x-y))\partial_x q\\ &\quad+\partial_x(xy(1-x-y))\partial_y q+xy(1-x-y)\partial_x\partial_y q\big)(0,0)=q(0,0).\end{aligned}$$

Dies impliziert $q(0,0)=0$. Analog folgt $q(1,0)=q(0,1)=0$. Weiter gilt

$$\begin{aligned}0&=\partial_n p(m_1)=-\partial_y\big(xy(1-x-y)q\big)(\tfrac12,0)\\ &=\big(-x(1-x-y)q+xyq-xy(1-x-y)\partial_y q\big)(\tfrac12,0)=-\tfrac14 q(m_1)\end{aligned}$$

und analog $q(m_2)=q(m_3)=0$. Wegen der Unisolvenz von $P_2(T)$ mit dem Satz $\{q(a_i),q(m_i),i=1,2,3\}$ von Knotenwerten folgt schließlich $q\equiv 0.$ bzw. $p\equiv 0$.

aiii) T kartesisches Einheitsquadrat:

$$P(T)=\tilde Q_1(T):=P_1(T)\oplus\operatorname{span}\{x^2-y^2\},\quad p(m_i),\ i=1,\dots,4.$$

Sei $p \in P(T)$ mit $p(m_i) = 0$. Auf dem Einheitsquadrat ist p linear entlang der schrägen Verbindungslinien zwischen den Mitten benachbarter Kanten und folglich gleich Null entlang dieser Linien. Damit ist auch $\nabla p(m_i) = 0$. Entlang jeder Linie durch eine Seitenmitte m_i verschwindet damit p in mindestens zwei Punkten und seine Richtungsableitung in mindestens einem Punkt. Da p entlang jeder Linie höchstens quadratisch ist, folgt $p \equiv 0$.

$$P(T) = \tilde{Q}_3(T) := P_3(T) \oplus \text{span}\{x^3y, xy^3\}, \quad p(a_i),\ \nabla p(a_i),\ i = 1, \ldots, 4.$$

Sei $p \in P(T)$, welches bzgl. aller Knotenfunktionale verschwindet. Dann ist $p_{|\partial T} \equiv 0$. Folglich müßte p den Faktor $cxy(1-x)(1-y)$ enthalten mit einer Konstante $c \in \mathbb{R}$. Da der Term x^2y^2 aber nicht durch Elemente des Raumes $P(T)$ erzeugt werden kann, muß $c = 0$ sein. Dies bedeutet auch $p \equiv 0$.

b) Bei der Approximation der Laplace-Gleichung auf einem äquidistanten kartesischen Gitter gilt:

$$\dim V_h^{(3)} = \frac{9}{h^2}, \qquad \dim \tilde{V}_h^{(3)} = \frac{5}{h^2}.$$

Die Anzahl der von Null verschiedenen Elemente pro Zeile der Systemmatrizen ist

$$V_h^{(3)}: \ \ N_z = 10,\ N_b = 16,\ N_a = 37, \qquad \tilde{V}_h^{(3)}: \ \ N_z = 10,\ N_a = 27.$$

Lösung A.3.9: Wegen der Gültigkeit der Einbettung $H^2(\Omega) \subset L^\infty(\Omega)$ in 2 und 3 Dimensionen erfüllt das Funktional $F : H^2(\hat{T}) \to \mathbb{R}$,

$$F(v) = \max_{\hat{T}} |v - I_h v|$$

die Voraussetzungen des Bramble-Hilbert-Lemmas. Demnach gilt auf dem Referenzelement $\hat{T}$

$$\max_{\hat{T}} |\hat{v} - I_h \hat{v}| \ \le\ c \|\hat{\nabla}^2 \hat{v}\|_{\hat{T}}.$$

Die Transformationsargumente aus der Vorlesung ergeben

$$\max_{T} |v - I_h v| \ =\ \max_{\hat{T}} |\hat{v} - I_h \hat{v}| \ \le\ c \|\hat{\nabla}^2 \hat{v}\|_{\hat{T}} \ \le\ c h^2 h^{-\frac{d}{2}} \|\nabla^2 v\|_T.$$

Dies impliziert die behauptete Abschätzung in 2 Dimensionen. Für $d = 3$ folgt

$$\max_{\overline{\Omega}} |v - I_h v| \ \le\ c h^{\frac{1}{2}} \|\nabla^2 v\|_\Omega.$$

Lösung A.3.10: Alle 4 Abschätzungen sind richtig:

i) Auf dem Referenzelement $\hat{T}$ sind für den endlich dimensionalen Polynomraum $P(\hat{T})$ alle Normen äquivalent, d. h. es gilt mit einer h-unabhängigen Konstante $c > 0$, so dass

$$\|\hat{\nabla}^2 \hat{v}_h\|_{\hat{T}} \ \le\ \|\hat{v}_h\|_{2,\hat{T}} \ \le\ c \|\hat{v}_h\|_{\hat{T}}.$$

Die üblichen Transformationsargumente liefern nun

$$\|\nabla^2 v_h\|_T \;\le\; ch^{-2}h\|\hat{\nabla}^2\hat{v}_h\|_{\hat{T}} \le ch^{-1}\,\|\hat{v}_h\|_{\hat{T}} \;\le\; ch^{-2}\|v_h\|_T.$$

ii) Auf dem Quotientenraum

$$\frac{P(\hat{T})}{P_0}$$

sind die Normen $\|\nabla\cdot\|_{1,\hat{T}}$ und $\|\nabla\cdot\|_{\hat{T}}$ äquivalent. Mit dem Spurlemma gilt

$$\|\hat{\partial}_n\hat{v}_h\|_{\partial\hat{T}} \;\le\; c\|\hat{\nabla}\hat{v}_h\|_{1,\hat{T}} \;\le\; c\|\hat{\nabla}\hat{v}_h\|_{\hat{T}}.$$

Mit Transformationsargumenten folgt nun

$$\|\partial_n v_h\|_{\partial T} \;\le\; ch^{-1}ch^{\frac{1}{2}}\|\hat{\partial}_n\hat{v}_h\|_{\partial\hat{T}} \;\le\; ch^{-\frac{1}{2}}\|\hat{\nabla}\hat{v}_h\|_{\hat{T}} \;\le\; ch^{-\frac{1}{2}}\|\nabla v_h\|_T.$$

iii) Wir nutzen die Äquivalenz der $W^{1,\infty}$- und der L^2-Norm auf $P(\hat{T})$

$$\|\hat{\nabla}\hat{v}_h\|_{L^\infty(\hat{T})} \;\le\; \|\hat{v}_h\|_{W^{1,\infty}(\hat{T})} \;\le\; c\|\hat{v}_h\|_{\hat{T}}.$$

Nach Transformation folgt

$$\|\nabla v_h\|_{L^\infty(T)} \;\le\; ch^{-1}\|\hat{\nabla}\hat{v}_h\|_{L^\infty(\hat{T})} \;\le\; ch^{-1}\|\hat{v}_h\|_{\hat{T}} \;\le\; ch^{-2}\|v_h\|_T.$$

iv) Ausnutzen der Äquivalenz von L^1- und L^2-Norm auf $P(\hat{T})$ ergibt mit den Transformationsargumenten

$$\|v_h\|_{L^2(T)} \;\le\; ch\|\hat{v}_h\|_{L^2(\hat{T})} \;\le\; ch\|\hat{v}_h\|_{L^1(\hat{T})} \;\le\; ch^{-1}\|\hat{v}_h\|_{L^1(T)}.$$

Lösung A.3.11: i) T kartesisches Einheitsdreieck:

$$P(T) = P_3(T), \quad p(a_i),\ \nabla p(a_i),\ p(z), \qquad P(T) = P_3(T), \quad p(a_i),\ p(b_{ij}),\ p(z).$$

Sei $p \in P(T)$, welches bzgl. aller Knotenfunktionale verschwindet. Dann ist $p_{|\Gamma_i} \equiv 0$. Auf dem Einheitsdreieck ist folglich $p(x,y) = cxy(1-x-y)$, da jedes entlang von ∂T verschwindende Polynom die drei Faktoren x, y und $1-x-y$ enthalten muß. Wegen $p(z) = 0$ folgt notwendig $c = 0$, d. h. $p \equiv 0$.

ii) Ein Polynom $p \in P_5(T)$ mit $p(a_i) = 0,\ \nabla p(a_i) = 0,\ \nabla^2 p(a_i) = 0,\ \partial_n p(m_i) = 0$ hat auf dem Einheitsdreieck notwenig die Gestalt $p(x,y) = xy(1-x-y)q(x,y)$ mit einem $q \in P_2(T)$. Wegen $\nabla^2 p(0,0) = 0$ gilt:

$$\begin{aligned} 0 &= \partial_x\partial_y p(0,0) \\ &= \big(\partial_x\partial_y(xy(1-x-y))q + \partial_y(xy(1-x-y))\partial_x q \\ &\quad + \partial_x(xy(1-x-y))\partial_y q + xy(1-x-y)\partial_x\partial_y q\big)(0,0) = q(0,0). \end{aligned}$$

Dies impliziert $q(0,0) = 0$. Analog folgt $q(1,0) = \partial_y\partial_y p(1,0) = 0$ und $q(0,1) = \partial_x\partial_x p(0,1) = 0$. Weiter gilt

$$\begin{aligned} 0 = \partial_n p(m_1) &= -\partial_y\big(xy(1-x-y)q\big)(\tfrac{1}{2},0) \\ &= \big(-x(1-x-y)q + xyq - xy(1-x-y)\partial_y q\big)(\tfrac{1}{2},0) = -\tfrac{1}{4}q(m_1) \end{aligned}$$

und analog $q(m_2) = q(m_3) = 0$. Wegen der Unisolvenz von $P_2(T)$ mit dem Satz $\{q(a_i), q(m_i), i = 1,2,3\}$ von Knotenwerten folgt schließlich $q \equiv 0$. bzw. $p \equiv 0$.

iii) T kartesisches Einheitsquadrat:

$$P(T) = \tilde{Q}_1(T) := P_1(T) \oplus \operatorname{span}\{x^2 - y^2\}, \quad p(m_i),\ i = 1,\dots,4.$$

Sei $p \in P(T)$ mit $p(m_i) = 0$. Auf dem Einheitsquadrat ist p linear entlang der schrägen Verbindungslinien zwischen den Mitten benachbarter Kanten und folglich gleich Null entlang dieser Linien. Damit ist auch $\nabla p(m_i) = 0$. Entlang jeder Linie durch eine Seitenmitte m_i verschwindet damit p in mindestens zwei Punkten und seine Richtungsableitung in mindestens einem Punkt. Da p entlang jeder Linie höchstens quadratisch ist, folgt $p \equiv 0$.

$$P(T) = \tilde{Q}_3(T) := P_3(T) \oplus \operatorname{span}\{x^3y, xy^3\}, \quad p(a_i),\ \nabla p(a_i),\ i = 1,\dots,4.$$

Sei $p \in P(T)$, welches bzgl. aller Knotenfunktionale verschwindet. Dann ist $p_{|\partial T} \equiv 0$. Folglich müßte p den Faktor $cxy(1-x)(1-y)$ enthalten mit einer Konstante $c \in \mathbb{R}$. Da der Term x^2y^2 aber nicht durch Elemente des Raumes $P(T)$ erzeugt werden kann, muß $c = 0$ sein. Dies bedeutet auch $p \equiv 0$.

Lösung A.3.12: a) Sei $\mathbb{T}_h$ eine Triangulierung von Ω durch Dreiecke. Dann wählen wir als Ansatz- und Testraum

$$V_h^{(1)} := \left\{\varphi \in C^0 : \varphi|_T \in P^1(T) \quad \forall T \in \mathbb{T}_h\right\}.$$

Die zugehörige Variationsgleichung lautet dann: Finde ein $u_h \in V_h^{(1)}$, so dass

$$(\nabla u_h, \nabla\varphi_h) + (u_h, \varphi_h) = (f, \varphi_h) \quad \forall \varphi_h \in V_h^{(1)}.$$

b) Zunächst gilt für die H^1-Norm und eine beliebige Funktion $\varphi_h \in V_h^{(1)}$ aufgrund der Galerkin-Orthogonalität

$$\begin{aligned} \|u - u_h\|_1^2 &= (\nabla(u-u_h), \nabla(u-u_h)) + (u-u_h, u-u_h) \\ &= (\nabla(u-u_h), \nabla(u-\varphi_h)) + (u-u_h, u-\varphi_h) \\ &\le \|\nabla(u-u_h)\|\|\nabla(u-\varphi_h)\| + \|u-u_h\|\|u-\varphi_h\| \\ &\le \tfrac{1}{2}(\|\nabla(u-u_h)\|^2 + \|\nabla(u-\varphi_h)\|^2 + \|u-u_h\|^2 + \|u-\varphi_h\|^2) \\ &= \tfrac{1}{2}(\|u-u_h\|_1^2 + \|u-\varphi_h\|_1^2) \end{aligned}$$

und somit

$$\|u - u_h\|_1 \le \inf_{\varphi_h \in V_h^{(1)}} \|u - \varphi_h\|_1.$$

Unter Verwendung der Standard-Approximationsabschätzungen folgt also

$$\|u - u_h\|_1 \leq ch\|\nabla^2 u\|$$

Für die L^2-Norm betrachten wir das duale Problem

$$-\Delta z + z = \frac{e}{\|e\|} \quad \text{in } \Omega, \quad \partial_n z = 0 \text{ auf } \partial\Omega,$$

wobei $e = u - u_h$ ist. Es gilt hierfür die a priori Abschätzung $\|\nabla^2 z\| \leq c$ und es folgt

$$\begin{aligned}\|e\| &= (\nabla e, \nabla(z - I_h z)) + (e, z - I_h z) \\ &\leq \|e\|_1 \|z - I_h z\|_1 \leq ch^2 \|\nabla^2 u\| \|\nabla^2 z\| \leq ch^2 \|\nabla^2 u\|.\end{aligned}$$

c) Wir erhalten einen konformen FE-Ansatzraum, indem wir zunächst für unsere Triangulierung $\mathbb{T}_h$ fordern, das alle Rand/Eckpunkte der Triangulierung auf dem Rand von Ω liegen. Dann definieren wir unseren Ansatzraum wie in (a), wobei wir im Falle das $\partial\Omega$ ausserhalb von $\overline{\Omega}_h$ liegt, die Testfunktionen linear bis zum Rand von Ω fortsetzen. Im umgekehrten Fall, ist unsere Testfunktion lediglich auf $\Omega \cap \Omega_h$ zu definieren. Im Falle nichthomogener Neumann-Bedingungen lautet die variationelle Formulierung

$$(\nabla u_h, \nabla \varphi_h) + (u_h, \varphi_h) = (f, \varphi_h) + (g, \varphi_h)_{\partial\Omega} \quad \forall \varphi_h \in V_h^{(1)}.$$

Lösung A.3.13: a) Als Ansatzraum verwenden wir

$$V_h^{(3)} := \{v_h \in H_0^1(\Omega_h) |\ v_{|T} \in P_3(T),\ T \in \mathbb{T}_h\}$$

wobei die Dreiecke entlang des Randes $\partial\Omega$ kubische Randkurven haben. Es sind die folgenden Fehlerabschätzungen zu erwarten:

$$\begin{aligned}\|\nabla(u - u_h)\|_{\Omega_h} &\leq ch^3 \|u\|_4, \\ \|u - u_h\|_{\Omega_h} &\leq ch^4 \|u\|_4,\end{aligned}$$

vorausgesetzt die Lösung ist $u \in H^4(\Omega)$.

b) Für den Fehler im Mittelwert ergibt sich mit Hilfe eines Dualitätsarguments:

$$\Big| \int_\Omega u\, dx - \int_\Omega u_h\, dx \Big| \leq ch^5 \|u\|_4.$$

Lösung A.3.14: a) Für das Referenzelement $\tilde{T}$ definieren wir das Funktional $F : H^2(T) \to P_1(T)$ durch

$$F(v) := \|v - I_h v\|_{\partial\tilde{T}}.$$

Für dieses gilt dann aufgrund des Spursatzes und der Interpolationsabschätzung bzgl. der L^2- und der $H^!$-Norm:

$$
\begin{aligned}
|F(v)| &= \|v - I_h v\|_{\partial\tilde{T}} \le c\|v - I_h v\|_{2;\tilde{T}} \le c\|v\|_{2,\tilde{T}},\\
|F(v+w)| &\le \|v + w - I_h(v+w)\|_{\partial\tilde{T}} \le \|v - I_h v\|_{\partial\tilde{T}} + \|w - I_h w\|_{\partial\tilde{T}} = |F(v)| + |F(w)|,\\
F(q) &= 0, \quad q \in P_1(\tilde{T}).
\end{aligned}
$$

Nach dem Lemma von Bramble/Hilbert folgt die Abschätzung

$$|F(v)| \le c\|\nabla^2 v\|_{\tilde{T}}$$

Das Transformationsargument aus dem Text ergibt dann

$$h_T^{-1/2}\|v - I_h v\|_{\partial T} \le c h_T^2 h_T^{-1}\|\nabla^2 v\|_T,$$

woraus die behauptete Abschätzung folgt. Alternativ kann man auch wie folgt argumentieren: Mit Hilfe des Spursatzes auf der Zelle T mit Durchmesser h_T ergibt sich zunächst die Abschätzung

$$\|v - I_h v\|_{\partial T} \le c h_T^{-1/2}\|v - I_h v\|_T + c h_T^{1/2}\|\nabla(v - I_h v)\|_T$$

und dann mit Hilfe der schon bekannten Fehlerabschätzungen über T:

$$\|v - I_h v\|_{\partial T} \le c h_T^{-1/2} h_T^2\|\nabla^2 v\|_T + c h_T^{1/2} h_T\|\nabla^2 v\|_T = c h_T^{3/2}\|\nabla^2 v\|_T.$$

b) Die Abschätzung

$$\|v - I_h v\|_{\partial T} \le c h_T^{1/2}\|\nabla v\|_T$$

kann nicht gelten, da für Funktionen $v \in H^1(\Omega)$ im Allg. die Knoteninterpolierende $I_h u \in P_2(T)$ gar nicht definiert ist. Sie wäre aber gültig mit der „Quasi-Interpolierenden" $\tilde{I}_h u$ (diskutiert in einer späteren Aufgabe) in der Form

$$\|v - I_h v\|_{\partial T} \le c h_T^{1/2}\|\nabla v\|_{\tilde{T}}, \quad \tilde{T} = \cup\{\tau \in \mathbb{T}_h,\ \tau \cap T \ne \emptyset\}.$$

Die direkte Anwendung des Bramble/Hilbert-Lemmas ist hier aber nicht möglich, da die dort definierte Quasi-Interpolierende i. Allg. Polynome nicht reproduziert. Der Beweis dieser Abschätzung bedarf also noch einiger Arbeit.

Lösung A.3.15: Für die Lagrange-Interpolation in $P(T) := P_2(T)$ gelten die Fehlerabschätzungen

$$
\begin{aligned}
(i) \qquad & \|\nabla^2(v - I_T v)\|_T \le c_i h_T\|\nabla^3 v\|_T;\\
(ii) \qquad & |(v - I_T v)(a)| \le c_i h_T^2\|\nabla^3 v\|_T;\\
(iii) \qquad & \|\partial_n(v - I_T v)\|_{\partial T} \le c_i h^{3/2}\|\nabla^3 v\|_T;\\
(iv) \qquad & \|v - I_T v\|_T \le c_i h_T^2\|\nabla^2 v\|_T.
\end{aligned}
$$

Lösung A.3.16: a) Auf dem konvexen Polygongebiet Ω ist jede Lösung $v \in H_0^1(\Omega)$ der Variationsgleichung

$$(\nabla v, \nabla\varphi) = (f, \varphi) \quad \forall\varphi \in H_0^1(\Omega),$$

auch in $V = H^2(\Omega)$, und es gilt die a priori Abschätzung

$$\|v\|_{H^2} \le c_\Omega \|f\| = c_\Omega \|\Delta v\|.$$

Auf dem Teilraum $H_0^2(\Omega) \subset H_0^1(\Omega)$ gilt also

$$a(v,v) := \|\Delta v\|^2 \ge c_\Omega^{-2} \|v\|_{H^2}^2,$$

d. h.: Die Bilinearform $a(\cdot,\cdot)$ ist V-elliptisch. Folglich hat die Variationsgleichung

$$a(u,\varphi) = (f,\varphi) \quad \forall \varphi \in V,$$

für jedes $f \in L^2(\Omega)$ eine eindeutige Lösung $u \in V$. Diese ist dann die sog. „schwache" Lösung der biharmonischen Gleichung. Auf einem Rechteckgebiet ist $u \in H^4(\Omega)$ und genügt der a priori Abschätzung

$$\|u\|_{H^4} \le c_s \|\Delta^2 u\| = c_s \|f\|.$$

b) Sei $V_h^{(5)} \subset V$ der Finite-Elemente-Raum basierend auf dem quintischen Argyris-Element. Die durch die Galerkin-Gleichungen

$$a(u_h,\varphi_h) = (f,\varphi_h) \quad \forall \varphi_h \in V_h^{(5)},$$

definierte Ritz-Approximation $u_h \in V_h^{(5)}$ besitzt die Bestapproximationseigenschaft

$$\|\Delta(u-u_h)\| = \min_{\varphi_h \in V_h^{(5)}} \|\Delta(u-\varphi_h)\|.$$

Aufgrund der Sobolewschen Einbettung $H^4(\Omega) \subset C^2(\overline{\Omega})$ ist die Knoteninterpolierende $I_h u \in V_h^{(5)}$ von u wohl definiert, und es gilt (Bramble/Hilbert-Lemma):

$$\|\Delta(u - I_h u)\| \le c h^2 \|u\|_{H^4}.$$

Dies impliziert die Energie-Fehlerabschätzung

$$\|u-u_h\|_{H^2} \le c\|\Delta(u-u_h)\| \le c h^2 \|u\|_{H^4}.$$

Zur Abschätzung des Fehlers bzgl. der L^2-Norm verwenden wir ein Dualitätsargument. Sei $z \in V$ die schwache Lösung der Variationsgleichung

$$(\Delta\varphi, \Delta z) = (\varphi, u-u_h)\|u-u_h\|^{-1} \quad \forall \varphi \in V.$$

Dann ist $z \in H^4(\Omega)$, und es gilt $\|z\|_{H^4} \le c$. Damit folgt mit Hilfe der Galerkin-Orthogonalität und der Interpolationsabschätzungen

$$\begin{aligned}\|u-u_h\| &= a(u-u_h, z) = a(u-u_h, z-I_h z) \le \|\Delta(u-u_h)\|\,\|\Delta(z-I_h z)\| \\ &\le c h^2 \|u\|_{H^4}\, h^2 \|z\|_{H^4} \le c h^4 \|u\|_{H^4}.\end{aligned}$$

c) Die Spektral-Konditionen der zugehörigen Steifigkeitsmatrix $A_h = \big(a(\varphi_h^j, \varphi_h^i)\big)_{i,j}$ und Massematrix $M_h = \big((\varphi_h^j, \varphi_h^i)\big)_{ij}$ sind gegeben durch

$$\kappa(A_h) = \frac{\lambda_{\max}(A_h)}{\lambda_{\min}(A_h)}, \qquad \kappa(M_h) = \frac{\lambda_{\max}(M_h)}{\lambda_{\min}(M_h)} \le c.$$

Mit Hilfe der inversen Beziehung folgt

$$\begin{aligned}\lambda_{\max}(A_h) &= \max_{x\in\mathbb{R}^N} \frac{(A_h x, x)}{|x|^2} \le \max_{x\in\mathbb{R}^N} \frac{(A_h x, x)}{(M_h x, x)} \max_{x\in\mathbb{R}^N} \frac{(M_h x, x)}{|x|^2} \\ &= \max_{v_h\in V_h^{(5)}} \frac{a(v_h, v_h)}{\|v_h\|^2}\, \lambda_{\max}(M_h) \le c h^{-4} \lambda_{\max}(M_h).\end{aligned}$$

Ferner

$$\begin{aligned}\lambda_{\min}(A_h) &= \min_{x\in\mathbb{R}^N} \frac{(A_h x, x)}{|x|^2} \ge \min_{x\in\mathbb{R}^N} \frac{(A_h x, x)}{(M_h x, x)} \min_{x\in\mathbb{R}^N} \frac{(M_h x, x)}{|x|^2} \\ &= \min_{v_h\in V_h^{(5)}} \frac{a(v_h, v_h)}{\|v_h\|^2}\, \lambda_{\min}(M_h) \ge \min_{v\in V} \frac{a(v, v)}{\|v\|^2}\, \lambda_{\min}(M_h) = \lambda_{\min}(\Delta^2)\, \lambda_{\min}(M_h).\end{aligned}$$

Dies impliziert

$$\kappa(A_h) \le c h^{-4} \lambda_{\min}(\Delta^2)^{-1} \frac{\lambda_{\max}(M_h)}{\lambda_{\min}(M_h)} \le c h^{-4} \lambda_{\min}(\Delta^2)^{-1} \kappa(M_h).$$

Da die Kondition der Massematrix $\kappa(M_h) = \mathcal{O}(1)$ ist, folgt $\kappa(A_h) = \mathcal{O}(h^{-4})$.

Lösung A.3.17: a) Konvergenz in der Energie-Norm ist garantiert, für Quadraturordnung $r \ge m-2$, mit der Ordnung m des FE-Ansatzes (Polynomgrad $m-1$) ist. Die optimale Konvergenzordnung wird erreicht für $r \ge 2m-3$. Im Fall „kubische" Elemente, d. h. $m = 4$, bedeutet dies $r \ge 2$ für Konvergenz und $r \ge 5$ für *optimale* Konvergenz. Geeignete, möglichst ökonomische Quadraturformeln wären für Konvergenz z. B. die Mittelpunktsregel und für optimale Konvergenz eine „Quasi"-Gauß-Formel 5-ter Ordnung.

b) Die 1. RWA des Laplace-Operators sei auf dem Einheitsquadrat mit bilinearen finiten Elementen auf einem äquidistanten, kartesischen Gitter diskretisiert. Die Elemente $(\nabla\varphi_h^j, \nabla\varphi_h^i)$ der Systemmatrix werden mit der Mittelpunktregel berechnet:
Die Koordinaten (a_i, b_j) beschreiben einen Gitterpunkt auf einem äquidistanten kartesischen Gitter mit der Gitterweite h. Die quadratischen Zellen um (a_i, b_j) sollen der Einfachheit halber mit lo, ro, lu und ru für *l*inks *o*ben etc. bezeichnet werden. Dann ergibt sich für die Basisfunktionen die Darstellung

$$\varphi_{i,j} := \begin{cases} -\frac{1}{h^2}(x-(a_i-h))(y-(b_j+h)) & (x,y)\in lo \\ \frac{1}{h^2}(x-(a_i+h))(y-(b_j+h)) & (x,y)\in ro \\ \frac{1}{h^2}(x-(a_i-h))(y-(b_j-h)) & (x,y)\in lu \\ -\frac{1}{h^2}(x-(a_i+h))(y-(b_j-h)) & (x,y)\in ru \end{cases}$$

mit dem Gradienten

$$\nabla\varphi_{i,j} = \begin{cases} \frac{1}{h^2}\begin{pmatrix} -y+b_j+h \\ -x+a_i-h \end{pmatrix} & (x,y) \in lo \\ \frac{1}{h^2}\begin{pmatrix} y-b_j-h \\ x-a_i-h \end{pmatrix} & (x,y) \in ro \\ \frac{1}{h^2}\begin{pmatrix} y-b_j+h \\ x-a_i+h \end{pmatrix} & (x,y) \in lu \\ \frac{1}{h^2}\begin{pmatrix} -y+b_j-h \\ -x+a_i+h \end{pmatrix} & (x,y) \in ru \end{cases}$$

Damit berechnen sich die Skalarprodukte mit Hilfe der Mittelpunktsregel zu

$$\begin{aligned} (\nabla\varphi_{i,j}, \nabla\varphi_{i,j}) &= \frac{4}{h^4}\int_{a_i-h}^{a_i}\int_{b_j}^{b_j+h}\begin{pmatrix} -y+b_j+h \\ -x+a_i-h \end{pmatrix}^2 dxy \\ &\approx \frac{4}{h^2}\left(\frac{h^2}{4}+\frac{h^2}{4}\right) = 2 \\ (\nabla\varphi_{i,j}, \nabla\varphi_{i+1,j}) &= \frac{2}{h^4}\int_{a_i}^{a_i+h}\int_{b_j}^{b_j+h}\begin{pmatrix} y-b_j-h \\ x-a_i-h \end{pmatrix}\begin{pmatrix} -y+b_{j+1}+h \\ -x+a_{i+1}-h \end{pmatrix} dxy \\ &\approx \frac{2}{h^2}\left(-\frac{h^2}{4}+\frac{h^2}{4}\right) = 0 \\ (\nabla\varphi_{i,j}, \nabla\varphi_{i+1,j+1}) &= \frac{1}{h^4}\int_{a_i}^{a_i+h}\int_{b_j}^{b_j+h}\begin{pmatrix} y-b_j-h \\ x-a_i-h \end{pmatrix}\begin{pmatrix} y-b_{j+1}+h \\ x-a_{i+1}+h \end{pmatrix} dxy \\ &\approx \frac{1}{h^2}\left(-\frac{h^2}{2}-\frac{h^2}{2}\right) = -\frac{1}{2} \end{aligned}$$

Die Steifigkeitsmatrix hat damit die Blockform

$$\begin{pmatrix} D & N & 0 & 0 & \cdots \\ N & D & N & 0 & \cdots \\ 0 & N & D & N & \cdots \\ \vdots & & \ddots & \ddots & \ddots \end{pmatrix},$$

mit den Untermatrizen

$$D = \begin{pmatrix} 2 & 0 & 0 & \cdots \\ 0 & 2 & 0 & \cdots \\ 0 & 0 & 2 & \cdots \\ \vdots & & & \ddots \end{pmatrix}, \quad N = \begin{pmatrix} 0 & -1/2 & 0 & 0 & \cdots \\ -1/2 & 0 & -1/2 & 0 & \cdots \\ 0 & -1/2 & 0 & -1/2 & \cdots \\ \vdots & & \ddots & & \ddots \end{pmatrix}$$

Dieses Verfahren entspricht dem 5-Punkte-Differenzenoperator, wobei hier die Richtungen zur Bestimmung des Differenzenquotienten diagonal zu den kartesischen Achsen verlaufen. Es handelt sich aber trotzdem um eine konsistente (Ordnung 2) Approximation des Laplace-Operators, da dieser invariant gegenüber Rotation des Koordinatensystems ist.

Lösung A.3.18: i) Eine Triangulierung $\mathbb{T}_h$ heißt „quasi-gleichförmig“ falls sie formregulär,

$$\sup_{h>0}\left(\max_{T\in\mathcal{T}_h}\frac{h_T}{\rho_T}\right)\leq c,$$

und „größenregulär“,

$$\sup_{h>0}\left(\frac{\max h_T}{\min h_T}\right)\leq c,$$

ist.

ii) Die negative h-Potenz bei der Konditionsabschätzung stammt von der Benutzung einer inversen Beziehung zur Abschätzung von $a(v_h,v_h)$, diese und alle weiteren Schritte sind unabhängig von der konkreten Diskretisierung $V_h\subset V$, wie man sich beim Durchgehen des Beweises überzeugt:

$$\begin{aligned}\lambda_{\min}(A)&\geq\min_{\xi\in\mathbb{R}}\frac{\langle A\xi,\xi\rangle}{\langle M\xi,\xi\rangle}\min_{\xi\in\mathbb{R}}\frac{\langle M\xi,\xi\rangle}{\langle\xi,\xi\rangle}=\min_{v_h\in V_h}\frac{a(v_h,v_h)}{\|v_h\|^2}\lambda_{\min}(M)\\&\geq\min_{v\in V}\frac{a(v,v)}{\|v\|^2}\lambda_{\min}(M)=\lambda_{\min}(-\Delta)\,\lambda_{\min}(M)\end{aligned}$$

und

$$\lambda_{\max}(A)\leq\max_{\xi\in\mathbb{R}}\frac{\langle A\xi,\xi\rangle}{\langle M\xi,\xi\rangle}\max_{\xi\in\mathbb{R}}\frac{\langle M\xi,\xi\rangle}{\langle\xi,\xi\rangle}=\max_{v_h\in V_h}\frac{a(v_h,v_h)}{\|v_h\|^2}\lambda_{\max}(M).$$

Mit der inversen Beziehung folgt:

$$a(v_h,v_h)=\sum_T\|\nabla v_h\|_T^2\leq c\sum_T\rho_T^{-2}\|v_h\|_T^2\leq c\big(\max_T\rho_T^{-2}\big)\|v_h\|^2.$$

Es gilt weiterhin $\operatorname{cond}_2(M)=O(1)$ und somit folgt zusammen mit der Formregularität:

$$\operatorname{cond}_2(A)\leq c\max_T\rho_T^{-2}\operatorname{cond}_2(M)=O\left(\max_T\rho_T^{-2}\right)=O\left(h^{-2}\right).$$

iii) Für den Fall, dass die Anzahl der an einer Ecke zusammenstoßenden Zellen beschränkt bleibt, gilt auch ohne Formregularität weiterhin $\operatorname{cond}_2(M)=O(1)$. Die Argumentation in (ii) benötigte nur im letzten Schritt Formregularität, so dass folgt:

$$\operatorname{cond}_2(A_h)=O\left(\rho^{-2}\right),$$

wobei ρ der minimale Innenkreisdurchmesser ist.

Lösung A.3.19: Es ist

$$\begin{aligned}\lambda_{\min}(A) &\geq \min_{\xi\in\mathbb{R}} \frac{\langle A\xi,\xi\rangle}{\langle M\xi,\xi\rangle} \min_{\xi\in\mathbb{R}} \frac{\langle M\xi,\xi\rangle}{\langle \xi,\xi\rangle} = \min_{v_h\in V_h} \frac{a(v_h,v_h)}{\|v_h\|^2}\lambda_{\min}(M)\\ &\geq \min_{v\in V} \frac{a(v,v)}{\|v\|^2}\lambda_{\min}(M) = \lambda_{\min}(-\Delta)\,\lambda_{\min}(M)\end{aligned}$$

und

$$\lambda_{\max}(A) \leq \max_{\xi\in\mathbb{R}} \frac{\langle A\xi,\xi\rangle}{\langle M\xi,\xi\rangle} \max_{\xi\in\mathbb{R}} \frac{\langle M\xi,\xi\rangle}{\langle \xi,\xi\rangle} = \max_{v_h\in V_h} \frac{a(v_h,v_h)}{\|v_h\|^2}\lambda_{\max}(M).$$

Mit der inversen Beziehung folgt:

$$a(v_h,v_h) = \sum_T \|\nabla v_h\|_T^2 \leq c\sum_T \rho_T^{-2}\|v_h\|_T^2 \leq c\max_T \rho_T^{-1}\|v_h\|^2.$$

Aufgrund der Formregulatität ist $\operatorname{cond}_2(M) = O(1)$ und es folgt insgesamt

$$\operatorname{cond}_2(A) \leq c\max_T \rho_T^{-2}\operatorname{cond}_2(M) = O\left(\max_T \rho_T^{-2}\right).$$

Lösung A.3.20: a) Die variationelle Formulierung lautet:

$$\int_\Omega \alpha\nabla u\cdot\nabla\varphi\,dx + \int_\Omega \gamma u\varphi\,dx = \int_\Omega f\varphi\,dx + \int_{\partial\Omega} g\varphi\,do_x$$

b) Analog zum Vorgehen in der Vorlesung formuliert man das duale Problem: Finde $z\in H^1$, so dass

$$\int_\Omega \alpha\nabla\varphi\cdot\nabla z\,dx + \int_\Omega \gamma\varphi z\,dx = \frac{(\nabla e_h,\nabla\varphi)}{\|e_h\|_E}.$$

Setzen von $\varphi = e_h$ liefert die Fehleridentität

$$\|e_h\|_E = \int_\Omega f(z+\psi_h)\,dx + \int_{\partial\Omega} g(z+\psi_h)\,dx - (\alpha\nabla U,\nabla z+\nabla\psi_h) - (\gamma U, z+\psi_h)$$

für beliebiges $\psi_h\in V_h$. Zellweises partielles Integrieren führt auf die Ungleichung

$$\|e_h\|_E \leq \sum_{T\in\mathbb{T}_h}\left\{\|f+\Delta U-\gamma U\|_T\,\|z+\psi_h\|_T + \tfrac{1}{2}\|[\partial_n U]\|_{\partial T}\,\|z+\psi_h\|_{\partial T}\right\},$$

hierbei sei $[\partial_n U]\big|_{\partial T\cap\partial\Omega} = 2(g-\partial_n U)$, ansonsten der übliche Sprung über eine innere Kante. Man wählt nun wieder $\psi_h = I_h z$ und folgert mit der Clément-Interpolationsabschätzung,

$$\|z-I_h z\|_T + h_T^{1/2}\|z-I_h z\|_{\partial T} \leq ch_T\|\nabla z\|_{\tilde{T}}$$

die Abschätzung:

$$\begin{aligned}\|e_h\|_E &\le \sum_{T\in\mathbb{T}_h} ch_T \Big\{\|f+\Delta U-\gamma U\|_T + \tfrac{1}{2}h_T^{-1/2}\|[\partial_n U]\|_{\partial T}\Big\}\,\|\nabla z\|_{\tilde T}\\ &\le \Big(\sum_{T\in\mathbb{T}_h} ch_T^2\,\big\{\|f+\Delta U-\gamma U\|_T^2 + \tfrac{1}{4}h_T^{-1}\|[\partial_n U]\|_{\partial T}^2\big\}\Big)^{1/2}\Big(\sum_{T\in\mathbb{T}_h}\|\nabla z\|_{\tilde z}^2\Big)^{1/2}\\ &\le \Big(\sum_{T\in\mathbb{T}_h} ch_T^2\,\big\{\|f+\Delta U-\gamma U\|_T^2 + \tfrac{1}{4}h_T^{-1}\|[\partial_n U]\|_{\partial T}^2\big\}\Big)^{1/2}.\end{aligned}$$

since $\|\nabla z\|^2 \le c$.

c) Die Argumentation verläuft analog zu (b) mit dem modifizierten dualen Problem

$$\int_\Omega \alpha\nabla\varphi\cdot\nabla z\,dx + \int_\Omega \gamma\varphi z\,dx = \frac{(e_h,\varphi)}{\|e_h\|_E}.$$

Im Gegensatz zum dualen Problem in (b) ist hierbei aber die rechte Seite aus L^2, so dass $z\in H^2(\Omega)$ erwartet werden kann. Dies ermöglicht daher die Nutzung der herkömmlichen Knoteninterpolierenden für ψ_h mit der Interpolationsfehlerabschätzung

$$\|z-I_h z\|_T + h_T^{1/2}\|z-I_h z\|_{\partial T} \le ch_T^2\|\nabla^2 z\|_T.$$

mit $\|\nabla^2 z\|\le c$. Dies liefert

$$\|e_h\| \le \Big(\sum_{T\in\mathbb{T}_h} ch_T^4\,\big\{\|f+\Delta U-\gamma U\|_T^2 + \tfrac{1}{4}h_T^{-1}\|[\partial_n U]\|_{\partial T}^2\big\}\Big)^{1/2}.$$

Lösung A.3.21: Wir rekapitulieren die Spurabschätzung

$$\|\partial_n v\|_{\partial T} \le c\big\{h_T^{1/2}\|\Delta v\|_T + h_T^{-3/2}\|v\|_T\big\}.$$

Wegen $[\partial_n u]=0$ ergibt sich

$$\begin{aligned}\|[\partial_n u_h]\|_{\partial T\setminus\partial\Omega}^2 = \|[\partial_n e_h]\|_{\partial T\setminus\partial\Omega}^2 &\le c\sum_{T\subset\tilde T}\|\partial_n e_h\|_{\partial T\setminus\partial\Omega}^2\\ &\le c\sum_{T\subset\tilde T}\big\{h_T\|\Delta e_h\|_T^2 + h_T^{-3}\|e_h||_T^2\big\}\\ &\le c\sum_{T\subset\tilde T}\big\{h_T\|f+\Delta u_h\|_T^2 + h_T^{-3}\|e_h||_T^2\big\}\end{aligned}$$

und damit

$$\begin{aligned}\eta_{L^2}(u_h) &= \Big(\sum_{T\in\mathbb{T}_h} h_T^4\big\{\|f+\Delta u_h\|_T^2 + \tfrac{1}{2}h_T^{-1}\|[\partial_n u_h]\|_{\partial T\setminus\partial\Omega}^2\big\}\Big)^{1/2}\\ &\le c\|e_h\|_2 + \Big(\sum_{T\in\mathbb{T}_h} h_T^4\|f+\Delta u_h\|_T^2\Big)^{1/2}.\end{aligned}$$

Lösung A.3.22: Für die Nebenbedingung $\eta(h) \leq \mathrm{TOL}$ wird im Optimum Gleichheit angenommen. Wir suchen also stationäre Punkte des Lagrangefunktionals

$$\mathcal{L}(h, \lambda) = N(h) + \lambda(\eta(h) - \mathrm{TOL}).$$

dies sind gerade solche Punkte in denen

$$\frac{d}{dt}\mathcal{L}(h + t\varphi, \lambda + t\mu)|_{t=0} = 0$$

für alle $\varphi \in C\left(\overline{\Omega}\right)$ und alle $\mu \in \mathbb{R}$. Indem wir zunächst $\mu = 0$ und dann $\varphi = 0$ wählen erhalten wir die beiden Variationsgleichungen

$$2\int_\Omega \varphi\left(-h^{-3} + h\lambda A\right) dx = 0 \quad \forall \varphi \in C(\overline{\Omega})$$

und

$$\mu\left(\int_\Omega h^2 A\, dx - \mathrm{TOL}\right) = 0 \quad \forall \mu \in \mathbb{R}.$$

Aus der ersten folgt

$$h = (\lambda A)^{-1/4}$$

durch dieser obiger Beziehung in die zweite Gleichung folgt

$$\lambda^{1/2}\mathrm{TOL} = \int_\Omega A^{1/2}\, dx =: W.$$

Somit ist

$$\lambda = \left(\frac{W}{\mathrm{TOL}}\right)^2$$

und es folgt:

$$h(x) = \left(\frac{\mathrm{TOL}}{W}\right)^{1/2} A(x)^{-1/4}.$$

Lösung A.3.23: Mit $u \in W^{2,\infty}(\Omega)$ liegt f in $L^\infty(\Omega)$. Da u_h stückweise linear ist, folgt:

$$\|f + \Delta u_h\|_T = \|f\|_T \leq \|f\|_\infty h_T.$$

Wir betrachten nun den zweiten Summanden:

$$\begin{aligned}
\|[\partial_n u_h]\|_{\partial T} = \|[\partial_n e]\|_{\partial T} &\leq \|\partial_n e\|_{\partial T} + \|\partial_n \tilde{e}\|_{\partial T} \\
&\leq \|\nabla e\|_{\partial T} + \|\nabla \tilde{e}\|_{\partial \tilde{T}} \\
&\leq c h^{1/2}\left(\|\nabla e\|_{\infty,T} + \left\|\nabla^2 \tilde{e}\right\|_{\infty,\tilde{T}}\right) \\
&\leq c h^{3/2}\|\nabla^2 e\|_\infty = c h^{3/2}\|\nabla^2 u\|_\infty
\end{aligned}$$

wobei $\tilde{e}$ der Fehler auf den an ∂T angrenzenden Zellen ist. Insgesamt folgt somit die Behauptung.

Lösung A.3.24: Für die Lagrange-Interpolation in $P(T) := P_2(T)$ gelten die Fehlerabschätzungen

$$\begin{aligned} (i) \qquad & \|\nabla^2(v - I_T v)\|_T \le c_i h_T \|\nabla^3 v\|_T; \\ (ii) \qquad & |(v - I_T v)(a)| \le c_i h_T^2 \|\nabla^3 v\|_T; \\ (iii) \qquad & \|\partial_n(v - I_T v)\|_{\partial T} \le c_i h^{3/2} \|\nabla^3 v\|_T; \\ (iv) \qquad & \|v - I_T v\|_T \le c_i h_T^2 \|\nabla^2 v\|_T. \end{aligned}$$

Lösung A.3.25: a) Variationelle Formulierung im Raum $V = H^1(\Omega)$: Finde $u \in V$ mit

$$a(u, \varphi) := (\alpha \nabla u, \nabla \varphi)_\Omega + (\gamma u, \varphi)_\Omega = (f, \varphi)_\Omega + (g, \varphi)_{\partial\Omega}, \quad \forall \varphi \in V.$$

Die Bilinearform $a(\cdot, \cdot)$ ist offenbar symmetrisch, V-elliptisch und beschränkt auf V. Nach dem Darstellunsgsatz von Riesz (oder dem Lemma von Lax-Milgram) existiert also eine eindeutige Lösung $u \in V$. Ist diese hinreichend glatt, folgt durch partielle Integration

$$(-\nabla \cdot (\alpha \nabla u) + \gamma u - f, \varphi)_\Omega + (n \cdot (\alpha \nabla u) - g, \varphi)_{\partial\Omega} = 0, \quad \varphi \in V.$$

Nach dem Fundamentalsatz der Variationsrechnung impliziert dies, dass u klassische Lösung der RWA ist:

$$-\nabla \cdot (\alpha \nabla u) + \gamma u = f \quad \text{in } \Omega, \quad n \cdot (\alpha \nabla u)_{|\partial\Omega} = g.$$

b) Zur Herleitung der a posteriori Energienorm-Fehlerabschätzung schreiben wir mit Hilfe der Galerkin-Orthogonalität

$$\Sigma_h := (\alpha \nabla e_h, \nabla e_h)_\Omega + (\gamma e_h, e_h)_\Omega = (\alpha \nabla e_h, \nabla(e_h - i_h e_h))_\Omega + (\gamma e_h, e_h - i_h e_h)_\Omega.$$

Zellweise partielle Integration ergibt weiter

$$\begin{aligned} \Sigma_h &= \sum_{T \in \mathbb{T}_h} \Big\{ (-\nabla \cdot (\alpha \nabla e_h), e_h - i_h e_h)_T + (\gamma e_h, e_h - i_h e_h)_T + (n \cdot (\alpha \nabla e_h), e_h - i_h e_h)_{\partial T} \Big\} \\ &= \sum_{T \in \mathbb{T}_h} \Big\{ (R(u_h), e_h - i_h e_h)_T + (r(u_h), e_h - i_h e_h)_{\partial T} \Big\} \end{aligned}$$

mit den „Zell- und Kanten-Residuen" (bzw. „Gleichungs- und Sprung-Residuen")

$$R(u_h)_{|T} := f + \nabla \cdot (\alpha \nabla u_h) - \gamma u_h, \qquad r_{h|\Gamma} := \begin{cases} \frac{1}{2}[n \cdot (\alpha \nabla u_h)], & \text{für } \Gamma \subset \partial T \backslash \partial\Omega, \\ n \cdot (\alpha \nabla u_h) - g, & \text{für } \Gamma \subset \partial\Omega. \end{cases}$$

Mit Hilfe der Hölderschen Ungleichung folgt hieraus

$$|\Sigma_h| \le \sum_{T \in \mathbb{T}_h} \big\{ \|R(u_h)\|_T \|e_h - i_h e_h\|_T + \|r(u_h)\|_{\partial T} \|e_h - i_h e_h\|_{\partial T} \big\}$$

Wir wählen die Approximation $i_h e_h$ als verallgemeinerte Knoteninterpolierende mit der

Eigenschaft

$$\|e_h - i_h e_h\|_T + h_T^{1/2}\|e_h - i_h e_h\|_{\partial T} \le \tilde{c}_i h_T \|\nabla e_h\|_{\tilde{T}},$$

wobei $\tilde{T} := \cup\{T' \in \mathbb{T}_h | \, T' \text{ hat gemeinsame Kante mit } T\}$. Damit ergibt sich weiter

$$\begin{aligned}|\Sigma_h| &\le \tilde{c}_i \sum_{T\in\mathbb{T}_h} \big\{ h_T \|R(u_h)\|_T \|\nabla e_h\|_{\tilde{T}} + h_T^{1/2}\|r(u_h)\|_{\partial T}\|\nabla e_h\|_{\tilde{T}}\big\} \\ &\le \tilde{c}_i \Big(\sum_{T\in\mathbb{T}_h} h_T^2 \big\{ \|R(u_h)\|_T^2 + h_T^{-1}\|r(u_h)\|_{\partial T}^2\big\}\Big)^{1/2} \Big(\sum_{T\in\mathbb{T}_h} \|\nabla e_h\|_{\tilde{T}}^2\Big)^{1/2}.\end{aligned}$$

Wegen

$$\sum_{T\in\mathbb{T}_h} \|\nabla e_h\|_{\tilde{T}}^2 \le c\|\nabla e_h\|_\Omega^2 \le c(\alpha\nabla e_h, \nabla e_h)_\Omega + (\gamma e_h, e_h)_\Omega = c\Sigma_h.$$

ergibt sich die gewünschte a posteriori Fehlerabschätzung:

$$\|e_h\|_E \le c\tilde{c}_i \Big(\sum_{T\in\mathbb{T}_h} h_T^2 \big\{ \|R(u_h)\|_T^2 + h_T^{-1}\|r(u_h)\|_{\partial T}^2\big\}\Big)^{1/2}.$$

c) Zur Herleitung der L^2-Norm-Fehlerabschätzung verwenden wir wieder ein Dualitätsargument mit dem Funktional

$$J(\varphi) := \|e_h\|_\Omega^{-1}(\varphi, e_h)_\Omega, \qquad J(e_h) = \|e_h\|_\Omega.$$

Sei $z \in V \cap H^2(\Omega)$ die Lösung des dualen Problems

$$a(\varphi, z) = J(\varphi) \quad \forall \varphi \in V,$$

bzw.

$$-\nabla\cdot(\alpha\nabla z) + \gamma z = \|e_h\|_\Omega^{-1} e_h \;\text{ in }\; \Omega, \quad n\cdot(\alpha\nabla u)_{|\partial\Omega} = 0 * .$$

Für diese gilt auf dem konvexen Polygongebiet die a priori Abschätzung $\|\nabla^2 z\|_\Omega \le c_s$. Damit erschließen wir wie zuvor:

$$\begin{aligned}\|e_h\| &= J(e_h) = a(e_h, z) = a(e_h, z - i_h z) \\ &= \sum_{T\in\mathbb{T}_h} \big\{ R(u_h), z - i_h z)_T + (r(u_h), z - i_h z)_{\partial T}\big\} \\ &\le c_i \sum_{T\in\mathbb{T}_h} \big\{ \|R(u_h)\|_T h_T^2 \|\nabla^2 z\|_T + \|r(u_h)\|_{\partial T} h_T^{3/2}\|\nabla^2 z\|_T\big\} \\ &\le \Big(\sum_{T\in\mathbb{T}_h} h_T^4 \big\{ \|R(u_h)\|_T + h_T^{-1}\|r(u_h)\|_{\partial T}^2\big\}\Big)^{1/2} \Big(\sum_{T\in\mathbb{T}_h} \|\nabla^2 z\|_T^2\Big)^{1/2}\end{aligned}$$

Dies ergibt die gewünschte a posteriori Fehlerabschätzung

$$\|e_h\| \le \Big(\sum_{T\in\mathbb{T}_h} h_T^4 \big\{ \|R(u_h)\|_T + h_T^{-1}\|r(u_h)\|_{\partial T}^2\big\}\Big)^{1/2}.$$

Lösung A.3.26: Ein Glättungsschritt des Richardson-Verfahren benötigt im Wesentlichen eine Matrix-Vektor-Multiplikation. Die Matrix-Vektor-Multiplikation $A_h x_h$ kann mit $9N_l$ a. Op. ausgeführt werden, da A_h maximal 9 Nicht-Nulleinträge pro Zeile hat. Ein Glättungsschritt braucht damit $12N_l$ a. Op.
Die Berechnung des Defekts $d_l = f_l - A_l x^l$ kostet $10N_l$ a. Op. Für die L^2-Projektion auf das nächstfeinere Gitter müssen wir nun

$$\tilde{d}^{l-1} := r_l^{l-1} d_l.$$

berechnen. Innerhalb das Mehrgitter-Algorithmus kann die L^2-Projektion sehr effizient berechnet werden. Wir bezeichnen die Basisfunktionen auf Gitterlevel l mit φ_i^l. Nach Definition der L^2-Projektion ist die i-te Komponente von $\tilde{d}^{l-1}$ durch

$$\tilde{d}_i^{l-1} = (r_l^{l-1} d^l, \varphi_i^{l-1}) = (d^l, \varphi_i^{l-1}).$$

gegeben. Da $V_{l-1} \subset V_l$, können wir φ_i^{l-1} darstellen als

$$\varphi_i^{l-1} = \sum_{j=1}^{N_l} \mu_{ij} \varphi_j^l,$$

wobei für einen Index i maximal 9 Werte μ_{ij} nichttrivial sind. Daher reduziert sich die Berechnung der L^2-Projektion auf

$$\tilde{d}_i^{l-1} = \sum_{j=1}^{N_l} \mu_{ij} (d^l, \varphi_i^l) = \sum_{j=1}^{N_l} \mu_{ij} d_i^l$$

und benötigt $9N_l$ Operationen.
Die Prolongation kostet 2 a.op. für die Interpolation in jedem neuen Knoten (bei weniger als N_l solcher Knoten). Zusätzlich kostet das Aufaddieren der Korrektur N_l Operationen. Zusammen sind das

$$(2 \cdot 12 + 10 + 9 + 2 + 1) N_l = 46 N_l$$

Operationen auf Gitterlevel l. Die Dimension der diskreten Räume verhält sich wie

$$N_{l-k} \approx 2^{-2k} N_l$$

Innerhalb eines V-Zyklus, müssen wir die genannten Operationen genau einmal pro Gitterlevel ausführen. Für genügend großes l können wir die Kosten zum Lösen auf dem gröbsten Gitter vernachlässigen. Daher kostet ein V-Zyklus insgesamt

$$\sum_{k=0}^{l} 46 N_{l-k} = \sum_{k=0}^{l} \frac{46}{2^{2k}} N_l = \frac{4}{3} 46 N_l \left(1 - 2^{-(2k+2)}\right) \leq \frac{4}{3} 46 N_l.$$

Innerhalb eines W-Zyklus, führen wir auf Gitterlevel $l - k$ 2^k Schritte mit oben gezählten Operationen durch. Zusammen kostet das:

$$\sum_{k=0}^{l} 2^k 46 N_{l-k} = \sum_{k=0}^{l} \frac{46}{2^k} N_l = 2 \cdot 46 N_l \left(1 - 2^{-k-1}\right) \leq 2 \cdot 46 N_l$$

arithmetische Operationen.

Lösung A.3.27:

a) Man erhält den 5-Punkte-Differenzenoperator.

b) „Unisolvenz“ bedeutet, dass für ein Polynom $p \in P(T)$ aus $\chi(p) = 0$ $(r = 1, \ldots, R)$ notwendig $p \equiv 0$ folgt.

c) Die „Minimalwinkelbedingung“ besagt, dass alle Winkel der Dreiecke gleichmäßig von Null weg beschränkt sind; dagegen besagt die „Maximalwinkelbedingung“, dass alle diese Winkel gleichmäßig von π wegbeschränkt sind. Die „Minimalwinkelbedingung“ ist äquivalent zur „Formregularität“ (Quotient aus Umkreis- und Inkreisradius gleichmäßig beschränkt), während die „Größenregulärität“ damit gar nichts zu tun hat.

d) Es ist $\dim(P_2) = 6,\ \dim(P_5) = 21,\ \dim(Q_2) = 9\,.$

e) Die Spektralkondition der FE-Matrix wird durch die Ordnung des Differentialoperators bestimmt; im gegebenen Fall gilt $\mathcal{O}(h^{-2})\,.$

A.4 Kapitel 4

Lösung A.4.1: Wir betrachten O.B.d.A. nur den ersten N-Zyklus. Das Gauss-Seidel Verfahren hat gerade die Form

$$\hat{x}_j^{(1)} = \frac{1}{a_{jj}} \Big(b_j - \sum_{k<j} a_{jk} \hat{x}_k^{(1)} - \sum_{k>j} a_{jk} \hat{x}_k^{(0)} \Big).$$

Aufgrund der Wahl der Abstiegsrichtung $d^{(t)} = e_{t+1}$ gilt für die Iterierten des Koordinatenrelaxationsverfahrens $x_j^{(t+1)} = x_j^{(t)}$ für $j \neq t+1$. Es genügt also zu zeigen, dass im Schritt $t \to t+1$ die $t+1$-te Komponente von $x^{(t+1)}$ auf den Richtigen Wert gesetzt wird, der Rest folgt per Induktion.

Hierzu beachten wir $r^{(t)} = b - Ax^{(t)}$ und somit ist

$$\alpha_{t+1} = \frac{r_{t+1}^{(t)}}{a_{t+1,t+1}} = \frac{1}{a_{t+1,t+1}} \left(b_{t+1} - \sum_k a_{t+1,k} x_k^{(t)} \right).$$

Durch Einsetzen in die Verfahrensvorschrift folgt:

$$\begin{aligned} x_{t+1}^{(t+1)} &= x_{t+1}^{(t)} + \frac{b_{t+1}}{a_{t+1,t+1}} - \frac{1}{a_{t+1,t+1}} \sum_k a_{t+1,k} x_k^{(t)} \\ &= \frac{1}{a_{t+1,t+1}} \left(b_{t+1} - \sum_{j<t+1} a_{t+1,k} x_k^{(t)} - \sum_{j>t+1} a_{t+1,k} x_k^{(t)} \right). \end{aligned}$$

Verwendet man nun die Induktionsannahme $x_k^{(t)} = \hat{x}_k^{(1)}$ für $k < t+1$ und $x_k^{(t)} = \hat{x}_k^{(0)}$ für $k > t+1$, so folgt die behauptete Äquivalenz.

Lösung A.4.2: a) Wir rekapitulieren aus einer früheren Aufgabe: Die Eigenwerte der Iterationsmatrix $B_\theta = I - \theta A$ sind $\mu = 1 - \theta\lambda$ und folglich $\mu_{\max} = 1 - \theta\lambda_{\min}$ sowie $\mu_{\min} = 1 - \theta\lambda_{\max}$ bzw.

$$\rho(B_\theta) = \max|\mu| = \max_{i=1,\ldots,N} |1 - \theta\lambda_i|.$$

Im Falle $0 < \lambda_{\min} \leq \lambda_{\max}$ gilt

$$0 < \theta < \frac{2}{\lambda_{\max}} \quad \Leftrightarrow \quad 0 < \theta\lambda_{\min} \leq \theta\lambda_{\max} < 2.$$

Dies wiederum ist äquivalent mit $\max\{|1 - \theta\lambda_{\min}|, |1 - \theta\lambda_{\max}|\} < 1$. Ein einfaches geometrische Argument ergibt den Wert

$$\theta_{\text{opt}} = \frac{2}{\lambda_{\min} + \lambda_{\max}}$$

für minimales $\rho(B_\theta)$, d. h. für schnellste Konvergenz.

b) Die beste Glättungseigenschaft ist aber charakterisiert durch

$$|1 - \theta\lambda_{\max}| = \min_{i=1,\ldots,N} |1 - \theta\lambda_i|,$$

da dann die hochfrequenten Fehleranteile in der Darstellung

$$|e_h^{(t)}|^2 = \sum_{i=1}^{N_h} \varepsilon_i^2 (1 - \theta\lambda_i)^{2t}$$

am schnellsten gedämpft werden. Dies ist der Fall für

$$\theta_{\text{opt}} = \frac{1}{\lambda_{\max}}.$$

Dass in diesem Fall die niederfrequenten Fehleranteile nur sehr langsam gedämpft werden, spielt keine Rolle, da man ja nur an dem Glättungseffekt interessiert ist.

Lösung A.4.3: Es gibt Probleme beim Beweis der Approximationseigenschaft. Im Beweis aus dem Text ist

$$v_{L-1} = A_{L-1}^{-1} r_L^{L-1} f_L, \quad v_L = A_L^{-1} f_L.$$

Aufgrund der Definition der Operatoren A_l folgt

$$a(v_L, \varphi_L) = (f_L, \varphi_L) \quad \forall \varphi_L \in V_L$$

sowie

$$a\left(v_{L-1},\varphi_{L-1}\right)=\left(r_L^{L-1}f_L,\varphi_{L-1}\right)\quad\forall\varphi_{L-1}\in V_{L-1}.$$

Da r_L^{L-1} nicht die L^2-Projektion ist, ist i. Allg.

$$\left(r_L^{L-1}f_L,\varphi_{L-1}\right)\neq\left(f_L,\varphi_{L-1}\right)$$

somit sind v_L und v_{L-1} nicht die Ritz-Projektion der selben kontinuierlichen Funktion. Was bleibt ist, dass v_L die Ritz-Projektion von

$$a\left(v,\varphi\right)=\left(f_L,\varphi\right)\quad\forall\varphi\in V$$

sowie v_{L-1} die Ritz-Projektion von

$$a\left(\tilde{v},\varphi\right)=\left(r_L^{L-1}f_L,\varphi\right)\quad\forall\varphi\in V$$

ist. Damit ist

$$\left\|v_L-v\right\|\leq ch_L^2\left\|\nabla^2 v\right\|\leq ch_L^2\left\|f_L\right\|$$

sowie

$$\left\|v_{L-1}-\tilde{v}\right\|\leq ch_{L-1}^2\left\|\nabla^2\tilde{v}\right\|\leq ch_L^2\left\|r_L^{L-1}f_L\right\|.$$

Damit bleibt noch zu zeigen, dass

$$\left\|r_L^{L-1}f_L\right\|\leq c\left\|f_L\right\|$$

sowie

$$\left\|v-\tilde{v}\right\|\leq ch_L^2\left\|f_L\right\|.$$

Während man die erste Ungleichung unter Ausnutzung der endlichen Dimension des Raumes V_L zeigen kann, ist die 2. Ungleichung i. Allg. nicht gültig.

Lösung A.4.4: Hier tritt das Problem bei der Glättungseigenschaft auf. Mithilfe einer inversen Ungleichung können wir zwar zeigen

$$\left\|A_L\right\|\leq ch^{-2}.$$

Es bleibt zu zeigen, dass

$$\left\|S_L\right\|\leq c<1$$

für $S_L=I_L-\theta A_L$ mit einer L-unabhängigen Konstante c. Da A_L nicht symmetrisch ist, können wir die Spektralargumente aus dem Text hier nicht direkt übertragen. A_L ist jedoch positiv definit. Um das zu sehen, wenden wir den Gaußschen Integralsatz an:

$$\left(\partial_1 u,u\right)=\frac{1}{2}\int_\Omega\partial_1(u^2)\,dx=\frac{1}{2}\int_{\partial\Omega}n_1u^2\,ds=0$$

für $u\in H_0^1(\Omega)$. Es folgt

$$a\left(u,u\right)=\left\|\nabla u\right\|^2.$$

Damit sind die Realteile der (komplexen) Eigenwerte von A_L positiv

$$\Re(\lambda_i) > 0 \quad i = 1...N_L.$$

Die Eigenwerte von $S_L = I_L - \theta A_L$ sind $1 - \theta\lambda_i, i = 1...N_L$. Es gilt

$$\begin{aligned}|1 - \theta\lambda| = |1 - \theta\Re(\lambda) - \theta\Im(\lambda)| &= \left\{(1 - \theta\Re(\lambda))^2 + \theta^2(\Im(\lambda))^2\right\}^{\frac{1}{2}} \\ &= \left\{1 - 2\theta\Re(\lambda) + \theta^2(\Im(\lambda)^2 + \Re(\lambda)^2)\right\}^{\frac{1}{2}}\end{aligned}$$

Wählen wir

$$\theta < \max_{i=1...N_L} \frac{2\Re(\lambda_i)}{|\lambda_i|^2},$$

so ist

$$\mathrm{spr}(S_L) = \max_{i=1...N_L} |1 - \theta\lambda_i| < c < 1.$$

gleichmäßig in L. Somit gibt es zu jedem $\epsilon > 0$ eine Norm $\|\cdot\|_*$ mit

$$\|S_L\|_* \leq c + \epsilon.$$

Es bleibt noch die Frage offen, inwieweit die L-unabhängige Konvergenzrate in der Norm $\|\cdot\|$ erhalten bleibt.

Lösung A.4.5: Ein Glättungsschritt des Richardson-Verfahren benötigt im Wesentlichen eine Matrix-Vektor-Multiplikation. Die Matrix-Vektor-Multiplikation $A_h x_h$ kann mit $9N_l$ a. Op. ausgeführt werden, da A_h maximal 9 Nicht-Nulleinträge pro Zeile hat. Ein Glättungsschritt braucht damit $12N_l$ a. Op.

Die Berechnung des Defekts $d_l = f_l - A_l x^l$ kostet $10N_l$ a. Op. Für die L^2-Projektion auf das nächstfeinere Gitter müssen wir nun

$$\tilde{d}^{l-1} := r_l^{l-1} d_l$$

berechnen. Innerhalb das Mehrgitter-Algorithmus kann die L^2-Projektion sehr effizient berechnet werden. Wir bezeichnen die Basisfunktionen auf Gitterlevel l mit φ_i^l. Nach Definition der L^2-Projektion ist die i-te Komponente von $\tilde{d}^{l-1}$ durch

$$\tilde{d}_i^{l-1} = (r_l^{l-1} d^l, \varphi_i^{l-1}) = (d^l, \varphi_i^{l-1}).$$

gegeben. Da $V_{l-1} \subset V_l$, können wir φ_i^{l-1} darstellen als

$$\varphi_i^{l-1} = \sum_{j=1}^{N_l} \mu_{ij} \varphi_j^l,$$

wobei für einen Index iv maximal 9 Werte μ_{ij} nichttrivial sind. Daher reduziert sich die Berechnung der L^2-Projektion auf

$$\tilde{d}_i^{l-1} = \sum_{j=1}^{N_l} \mu_{ij}(d^l, \varphi_i^l) = \sum_{j=1}^{N_l} \mu_{ij} d_i^l$$

und benötigt $9N_l$ Operationen.

Die Prolongation kostet 2 a. Op. für die Interpolation in jedem neuen Knoten (bei wenigwe als N_l solcher Knoten). Zusätzlich kostet das Aufaddieren der Korrektur N_l Operationen. Zusammen sind das

$$(2 \cdot 12 + 10 + 9 + 2 + 1)N_l = 46N_l$$

Operationen auf Gitterlevel l. Die Dimension der diskreten Räume verhält sich wie

$$N_{l-k} \approx 2^{-2k} N_l$$

Innerhalb eines V-Zyklus, müssen wir die genannten Operationen genau einmal pro Gitterlevel ausführen. Für genügend großes l können wir die Kosten zum Lösen auf dem gröbsten Gitter vernachlässigen. Daher kostet ein V-Zyklus insgesamt

$$\sum_{k=0}^{l} 46N_{l-k} = \sum_{k=0}^{l} \frac{46}{2^{2k}} N_l = \frac{4}{3} 46N_l \left(1 - 2^{-(2k+2)}\right) \leq \frac{4}{3} 46N_l.$$

Innerhalb eines W-Zyklus, führen wir auf Gitterlevel $l-k$ 2^k Schritte mit oben gezählten Operationen durch. Zusammen kostet das die folgenden a. Op.:

$$\sum_{k=0}^{l} 2^k 46N_{l-k} = \sum_{k=0}^{l} \frac{46}{2^k} N_l = 2 \cdot 46N_l \left(1 - 2^{-k-1}\right) \leq 2 \cdot 46N_l.$$

Lösung A.4.6: a) Die Eigenwerte der Iterationsmatrix $B_\theta = I - \theta A$ sind $\mu = 1 - \theta\lambda$ und folglich $\mu_{\max} = 1 - \theta\lambda_{\min}$ sowie $\mu_{\min} = 1 - \theta\lambda_{\max}$ bzw.

$$\rho(B_\theta) = \max\{|\mu|\} = \max\{|1 - \theta\lambda_{\max}|, |1 - \theta\lambda_{\min}|\}.$$

b) Im Falle $0 < \lambda_{\min} \leq \lambda_{\max}$ gilt

$$0 < \theta < \frac{2}{\lambda_{\max}} \quad \Leftrightarrow \quad 0 < \theta\lambda_{\min} \leq \theta\lambda_{\max} < 2.$$

Dies wiederum ist äquivalent mit $\max\{|1 - \theta\lambda_{\min}|, |1 - \theta\lambda_{\max}|\} < 1$.

c) Ein einfaches geometrische Argument ergibt

$$\theta_{\text{opt}} = \frac{2}{\lambda_{\min} + \lambda_{\max}}.$$

Lösung A.4.7: a) Die schwache Formulierung lautet

Finde $u \in H_0^1(\Omega)$ *, so dass*

$$(a\nabla u, \nabla\varphi) + (bu, \varphi) = (f, \varphi) \quad \forall\varphi \in H_0^1(\Omega).$$

Wir definieren die Bilinearform

$$a(u, \varphi) = (a\nabla u, \nabla\varphi) + (bu, \varphi)$$

Um Existenz und Eindeutigkeit einer Lösung zu gewährleisten (mit dem Lax-Milgram-Lemma oder dem Rieszschen Darstellungssatz) brauchen wir die Abschätzung

$$a(u, u) \geq \alpha\|u\|^2_{H^1(\Omega)}$$

für eine Konstante $\alpha > 0$. Sei

$$a_0 := \min_{x\in\Omega} a(x), \quad b_0 := \min_{x\in\Omega} b(x).$$

Es gilt

$$a(u, u) \geq a_0\|\nabla u\|^2 + b_0\|u\|^2.$$

Daher sind die Bedingungen $a_0 > 0$ und $b_0 \geq 0$ hinreichend, denn es folgt mit der Poincaréschen Ungleichung

$$a(u, u) \geq a_0\|\nabla u\|^2 \geq \frac{a_0}{c_P + 1}\|u\|^2_{H^1(\Omega)}.$$

b) Wir definieren eine reguläre Triangulierung $\mathbb{T}_h$ des Polyeders Ω in abgeschlossene Tetraeder T sowie einen Satz von linearen Funktionalen $\mathcal{X}$ auf $P(T)$, so dass ein Polynom $p \in P_2(T)$ eindeutig durch die Werte von $\mathcal{X}(p)$ festgelegt ist. Für den Fall quadratischer Polynome ist das einzig übliche konforme Element das Lagrange-Element, welches Punktwerte in Knotenpunkten und Seitenmittelpunkten verwendet. Der dazugehörige Ansatzraum ist

$$V_h^{(2)} := \left\{v_h \in C(\overline{\Omega}) \,\middle|\, v_h \in P_2(T)\ \forall T \in \mathbb{T}_h, v_h = 0 \text{ auf } \partial\Omega\right\}.$$

Eine Basis von $V_h^{(2)}v$ wird durch

$$\varphi_i(a_j) = \delta_{ij}, \quad i, j = 1...N,$$

definiert, wobei die Menge $\{a_i, i = 1...N\}$ aus Knoten- und Seitenmittelpunkten bestehe. Die Steifigkeitsmatrix $A_h = (a_{ij})_{i,j=1}^N$ und der Lastvektor $b_h = (b_i)_{i=1}^N$ berechnen sich zu

$$a_{ji} = a(\varphi_i, \varphi_j), \quad b_i = (f, \varphi_i).$$

Das diskrete System

$$a(u_h, \varphi_i) = (f, \varphi_i) \quad \forall\varphi_i \in V_h^{(2)}$$

ist nun äquivalent zum linearen Gleichungssystem

$$A_h x_h = b_h,$$

wobei x_i die Komponenten von u_h bezüglich der oben definierten Basis bezeichne:

$$u_h = \sum_{i=1}^{N} x_i \varphi_i.$$

c) Unter der Annahme $u \in H^3(\Omega)$, können wir mit Standardargumenten (Galerkin-Orthogonalität, Bestapproximationseigenschaft, Interpolationsabschätzungen) zeigen, dass

$$\|u - u_h\|_E := a(u,u)^{\frac{1}{2}} \leq ch^2 \|\nabla^3 u\|.$$

Mithilfe eines Dualitätsarguments bekommen wir in der L^2-Norm

$$\|u - u_h\| \leq ch^3 \|\nabla^3 u\|.$$

Die Spektralkondition hängt von der Ordnung des Differentialoperators ab. Für den gegeben elliptischen Differentialoperator zweiter Ordnung gilt

$$\kappa_2(A_h) = \mathcal{O}(h^{-2}).$$

di) Von einem Startwert $x^{(0)}$ definiert das Gauß- Seidel-Verfahren Iterierte durch

$$x_i^{(k+1)} = b_i - \sum_{j<k} a_{ij} x_j^{(k)} - \sum_{j>k} a_{ij} x_j^{(k+1)}, \quad i = 1...N.$$

Ein Gauß-Seidel-Schritt reduziert den Gesamtfehler um den Faktor

$$\rho = 1 - \mathcal{O}(h^2).$$

Die Anzahl an Iterationen T, die man benötigt, um den Anfangsfehler um den Faktor $\epsilon = 10^{-3}$ zu reduzieren, ist daher

$$\rho^T = \epsilon \qquad \Leftrightarrow \quad T = \frac{\ln(\epsilon)}{\ln(\rho)} = -3\,\frac{\ln(10)}{\ln(1-ch^2)} \simeq ch^{-2},$$

da $\ln(1 - ch^r) \simeq ch^r + \mathcal{O}(h^{2r})$. Eine Iteration besteht aus $2N$ Subtraktionen und $N-1$ Matrix-Vektor-Multiplikationen. Da die Matrix A_h dünn besetzt ist, kostet jede Iteration $\mathcal{O}(N)$ a. Op.. Da $N \approx h^{-3}$, brauchen wir insgesamt $\mathcal{O}(N^{\frac{5}{3}})$ a. Op..

dii) Das Gradientenverfahren ist ausgehend von einem Startwert $x^{(0)}$ definiert durch

$$x^{(t+1)} = x^{(t)} + \alpha_t(r^{(t)}), \quad \alpha_t = \frac{\|r^{(t)}\|^2}{(A_h r^{(t)}, r^{(t)})},$$

wobei $r^{(t)} := b_h - A_h x^{(t)}$ das Residuum in Schritt t bezeichnet. Ein Schritt des Gradien-

tenverfahrens reduziert den Fehler um den Faktor

$$\rho = 1 - \frac{1 - \kappa(A_h)^{-1}}{1 + \kappa(A_h)^{-1}} \approx 1 - 2\kappa(A_h)^{-1} = 1 - \mathcal{O}(h^2).$$

Hier haben wir benutzt, dass die Spektralkondition $\kappa(A_h)$ sich wie $\mathcal{O}(h^{-2})$ verhält. Eine Iteration besteht im Wesentlichen wieder aus einer konstanten Zahl an Matrix-Vektor-Multiplikationen und kostet daher $\mathcal{O}(N)$ a. Op. Die Iterationszahl ist wieder $\mathcal{O}(N^{\frac{5}{3}})$. Ein besseres Ergebnis liefert das CG-Verfahren. Ein Schritt des CG-Verfahrens reduziert den Fehler um einen Faktor

$$\rho = 1 - \frac{1 - \kappa(A_h)^{-\frac{1}{2}}}{1 + \kappa(A_h)^{-\frac{1}{2}}} \approx 1 - 2\kappa(A_h)^{-\frac{1}{2}} = 1 - \mathcal{O}(h).$$

Die Anzahl Iterationen T, um den Anfangsfehler um den Faktor ϵ zu reduzieren, kann daher mit

$$T \leq \frac{1}{2}\sqrt{\kappa(A_h)}\ln(\frac{2}{\epsilon}) + 1.$$

abgeschätzt werden, d. h.

$$T \in \mathcal{O}(N^{\frac{1}{3}}).$$

Die Anzahl a. Op. ist damit $\mathcal{O}(N^{\frac{4}{3}})$.

A.5 Kapitel 5

Lösung A.5.1: a) Es genügt zu zeigen, dass bei numerischer Integration über dem Einheitsdreieck $\hat{T}$ mit Hilfe der Dreiecks-Trapezregel für die drei Knotenbasisfunktionen gilt:

$$Q_T(\hat{\varphi}_i\varphi_j) = \delta_{ij}.$$

Paarweises Einsetzen der Basisfunktionen

$$\hat{\varphi}_1 = x, \quad \hat{\varphi}_2 = y, \quad \hat{\varphi}_3 = 1 - x - y,$$

in die Quadraturformel liefert aber sofort:

$$\hat{Q}_T(\hat{\varphi}_i\varphi_j) = \frac{1}{6}\left\{\hat{\varphi}_i(0,0)\hat{\varphi}_j(0,0) + \hat{\varphi}_i(1,0)\hat{\varphi}_j(1,0) + \hat{\varphi}_i(0,1)\hat{\varphi}_j(0,1)\right\} = \frac{1}{6}\delta_{ij}.$$

Durch Rücktransformation auf die Gitterzellen und „Assemblierung“ erhält man:

$$(\tilde{M}_h)_{ij} = \delta_{ij}\sum_{\overline{T}\ni x_{ii}}\frac{|T|}{3}.$$

Durch Masse-Lumping erhält man also eine positive Diagonalmatrix.

b) Nach dem diskreten Maximumsprinzip für finite Elemente ist im Fall, wenn alle Innenwinkel der Triangulierung kleiner oder gleich $\pi/2$ sind, die Steifigkeitsmatrix A_h eine

M-Matrix Es muss also nur gezeigt werden, dass beim Skalieren mit k und Addieren der ‘„gelumpten“ Massematrix $\tilde{M}_h$ die M-Matrixeigenschaft nicht verloren geht:

1. Die Diagonaldominanz,
$$\sum_{j \neq i} a_{ij} \leq a_{ii},$$
bleibt unter Skalierung und Addition positiver Beiträge auf der Diagonalen erhalten.

2. Die Eigenschaft ,,von nicht-negativen Typ”, $a_{ii} > 0$, $a_{ij} \leq 0$, $\forall i, j \neq i$, ebenso.

3. Da durch die Addition positiver Diagonalelemente keine Einträge der Steifigkeitsmatrix ausgelöscht werden können, ist die resultierende Systemmatrix wieder irreduzibel.

c) Nicht ausgeführt.

Lösung A.5.2: Nicht ausgeführt.

Lösung A.5.3: Nicht ausgeführt.

A.6 Kapitel 6

Lösung A.6.1:

1. Das Ritzsche Projektionsverfahren ist für variationelle Probleme mit *symmetrischer* Bilinearform $a(.,.)$ definiert. Ausgangspunkt ist die Formulierung über das Optimierungsproblem
$$\min_{u_h \in V_h} E(u_h); \quad E(\varphi) = \frac{1}{2} a(\varphi, \varphi) - (f, \varphi).$$
Dem gegenüber ist das Galerkinverfahren auch für nicht-symmetrische Bilinearformen definiert. Es bedient sich direkt der variationellen Formulierung: Finde $u_h \in V_h$, so dass:
$$a(u_h, \varphi_h) = (f, \varphi_h) \quad \forall \varphi_h \in V_h.$$
Die Petrov-Galerkin-Verfahren verwenden unterschiedliche Ansatz- und Testräume für die variationelle Formulierung:
$$u_h \in V_h^{\text{Ansatz}}, \quad \varphi_h \in V_h^{\text{Test}}.$$

2. Der Glätter auf den feineren Gitterlevel.

3. Für einen quadratischen Ansatz erhält man in der Energienorm $\|\nabla e_h\| \leq h^2\|\nabla^3 u\|$ und der L^2-Norm $\|e_h\| \leq h^3\|\nabla^3 u\|$. Es ist aber sogar möglich die Norm noch eine Stufe auf eine sog. „negative Sobolevnor“ abzuschwächen und damit 4te Ordnung 4 zu erhalten: Zu beliebigen $\psi \in H^1$ sei z Lösung des dualen Problems

$$(\nabla\varphi, \nabla z) = \frac{(\varphi, \psi)}{\|\psi\|_1}.$$

Es gilt die a priori-Abschätzung $\|z\|_3 \leq \left\|\psi/\|\psi\|_1\right\|_1 \leq c$ unäbhängig von ψ. Einsetzen von e_h in die Gleichung:

$$\frac{(e_h, \psi)}{\|\psi\|_1} \leq \|\nabla e_h\|\|\nabla z\| \leq ch^2\|\nabla^3 u\|\, ch^2\|\nabla^3 z\| \leq ch^4\|\nabla^3 u\|.$$

4. Die Maximalwinkelbedingung fordert, dass alle in einer Gitterhierarchie auftretenden Innenwinkel gleichmäßig nach oben von $180°$ weg beschränkt sind; analog fordert die Minimalwinkelbedingung eine Abschätzung nach unten von $0°$ weg.

5. Unter der Glättungseigenschaft versteht man die Gültigkeit einer Abschätzung der Form

$$\|U_h^M - u_h^m\| \leq c\frac{k^r}{t_m^r}\|u_h^0\|,$$

wobei r die Ordnung des Verfahrens, und U^M die FE-Näherung an die kontinuierliche Lösung u_h^m zum Zeitpunkt t_m ist.
Das Crank-Nicolson-Schema ist zwar A-stabil aber nicht *stark* A-stabil. Eine Glättungseigenschaft ist deshalb nicht zu erwarten und tatsächlich zeigen numerische Ergebnisse, dass das Crank-Nicolson-Schema keine Glättungseigenschaft besitzt.

6. Bei einem isoparametrischen Ansatz ist die Transformation einer lokalen Zelle (Dreieck, Rechteck) auf die Referenzzelle vom selben polynomialen Ansatzraum wie der FE-Ansatz.

7. Die Ordnung r der Quadraturformel sollte so gewählt werden, dass er zum einen zulässig ist und zum anderen $r \geq 2m - 3$.

8. Es ergibt sich eine Schrittweitenbedingung der Form

$$k \leq c\frac{1}{h^2}.$$

9. Der biharmonische Operator ist von 4ter Ordnung, die Kondition ist also

$$cond_2(A_h) = \mathcal{O}\left(h^{-4}\right).$$

10. Eine M-Matrix ist eine quadratische Matrix mit starker Diagonaldominanz:

$$\sum_{j\neq i} a_{ij} \leq a_{ii}\ \forall i, \text{ und } \exists k \text{ s.d. } \sum_{j\neq k} a_{kj} < a_{kk},$$

und von nicht-negativen Typs:

$$a_{ii} > 0,\ a_{ij} \leq 0,\ \forall i, j \neq i.$$

Eine M-Matrix ist regulär und ihre Inverse ist elementweise nicht negativ, $A_h^{-1} \geq 0$.