Einführende Kapitel der Numerik partieller Differentialgleichungen und Differentialungleichungen in Theorie und Praxis

Mitschrift[1] zur Vorlesung von Prof. Carstensen

Sommersemester 2004

[1] Gefundene Fehler bitte an vigerske@mathematik.hu-berlin.de.

Inhaltsverzeichnis

Kapitel 1

Projektion auf konvexe Mengen

1.1 Skalarprodukt

Definition 1.1 (Skalarprodukt, Prä-Hilbert-Raum, Innenproduktraum). Für einen reellen Vektorraum (VR) X heißt eine Abbildung $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ *Skalarprodukt*, falls

$\forall x, y, z \in X \ \forall \alpha, \beta \in \mathbb{R}$

1. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$,
2. $\langle x, y \rangle = \langle y, x \rangle$,
3. $\langle x, x \rangle \geq 0$ und $(\langle x, x \rangle = 0 \Leftrightarrow x = 0)$.

Dann heißt $(X, \langle \cdot, \cdot \rangle)$ *Prä-Hilbert-Raum* oder *Innenproduktraum*.

Definition 1.2 (Orthogonalität). Zwei Vektoren $x, y \in X$ mit $\langle x, y \rangle = 0$ heißen *orthogonal*. Schreibweise: $x \perp y$

Bemerkung 1.3. Wir betrachten hier nur *reelle* Vektorräume, Banach-Räume und Hilbert-Räume.

Beispiel. Prä-Hilbert-Räume sind $\mathbb{R}^n$,

$$\begin{aligned} \ell^2 &:= \left\{ (x_j) \in \mathbb{R}^n \,\middle|\, \sum_{j=1}^{\infty} x_j^2 < \infty \right\}, \\ L^2(\Omega) &:= \left\{ f : \Omega \to \mathbb{R} \text{ L-meßbar} \,\middle|\, \int_\Omega |f|^2 \, dx < \infty \right\}. \end{aligned}$$

Keine Prä-Hilbert-Räume sind L^p oder ℓ^p für $p \neq 2$.

1.2 Eigenschaften von Innenprodukträumen

In diesem Abschnitt bezeichne $(X, \langle \cdot, \cdot \rangle)$ einen Innenproduktraum und $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$ die induzierte Norm.

1.2.1 Cauchy-Ungleichung

Für alle $x, y \in X$ gilt
$$\langle x, y\rangle \leq \|x\| \, \|y\| \, .$$
Gleichheit (also $\langle x, y\rangle = \|x\| \, \|y\|$) gilt genau dann, wenn $x = 0$ oder $y = \lambda x$ für ein $\lambda \geq 0$.

1.2.2 Parallelogrammgleichung

Für alle $x, y \in X$ gilt
$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \, .$$
Umgekehrt: Wenn $\|\cdot\|$ eine Norm ist, welche die Parallelogrammgleichung erüllt, dann existiert ein Skalarprodukt $\langle\cdot,\cdot\rangle$, welches $\|\cdot\|$ induziert.

1.2.3 Innenprodukträume im $\mathbb{R}^n$

$(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ ist Innenproduktraum genau dann, wenn eine symmetrische positiv definite Matrix $A \in \mathbb{R}^{n\times n}$ existiert, so daß für alle $x, y \in \mathbb{R}^n$ gilt:
$$\langle x, y\rangle = x^T A y.$$

1.3 Normierte lineare Räume

Definition 1.4 (Norm, normierter linearer Raum). Für einen Vektorraum X heißt eine Abbildung $\|\cdot\| : X \to [0, \infty)$ *Norm*, falls

$\forall x, y \in X \ \forall \alpha \in \mathbb{R}$

1. $\|x\| \geq 0$ und $(\|x\| = 0 \Leftrightarrow x = 0)$,
2. $\|\alpha x\| = |\alpha| \, \|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$.

$(X, \|\cdot\|)$ heißt dann *normierter linearer Raum* (NLR).

Definition 1.5 (Cauchy-Folge). Eine Folge $(x_j) \in X^{\mathbb{N}}$ heißt *Cauchy-Folge* (CF) im normierten linearen Raum $(X, \|\cdot\|)$, falls
$$\forall \varepsilon > 0 \, \exists n \in \mathbb{N} \forall j, k \geq n \quad \|x_j - x_k\| \leq \varepsilon.$$

Definition 1.6 (Konvergenz, Limes). Eine Folge $(x_j) \in X^{\mathbb{N}}$ heißt *konvergent*, falls
$$\exists x \in X \, \forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} \forall n \geq n_0 \quad \|x - x_n\| < \varepsilon.$$
Dieses x heißt dann *Limes* der Folge.

Definition 1.7 (Vollständigkeit, Hilbert-Raum, Banach-Raum). Ein normierter linearer Raum heißt *vollständig*, falls darin jede Cauchy-Folge konvergiert.

Ein vollständiger Innenproduktraum heißt *Hilbert-Raum* (HR).

Ein vollständiger normierter linearer Raum heißt *Banach-Raum* (BR).

1.4 Mengen in Banachräumen

In diesem Abschnitt bezeichne $(X, \|\cdot\|)$ einen Banach-Raum.

Definition 1.8 (Abgeschlossenheit). Eine Menge $M \subseteq X$ heißt *abgeschlossen*, falls für jede Folge $(x_j) \in M^{\mathbb{N}}$ mit $(x_j) \to x$ für ein $x \in X$ gilt, daß $x \in M$.

Definition 1.9 (Vollständigkeit). Eine Menge $M \subseteq X$ heißt *vollständig*, falls jede Cauchy-Folge in M konvergiert (mit Limes in M).

In Banach-Räumen ist M genau dann vollständig , wenn M abgeschlossen ist.

Definition 1.10 (Konvexität). Eine Menge $M \subseteq X$ heißt *konvex*, falls

$\forall x, y \in M \ \forall \lambda \in (0, 1)$

$$\lambda x + (1 - \lambda) y \in M.$$

Definition 1.11 (Unterraum). Eine Menge $M \subseteq X$ heißt *Unterraum*, falls $0 \in M$ und

$\forall x, y \in M \ \forall \alpha, \beta \in \mathbb{R}$

$$\alpha x + \beta y \in M.$$

Unterräume sind konvex.

1.5 Bestapproximation in konvexen Mengen

Definition 1.12 (Abstand). Für eine Menge $M \subset X$, $M \neq \emptyset$, in einem normierten linearen Raum $(X, \|.\|)$, und ein $x \in X$ heißt

$$\operatorname{dist}(x, M) = \operatorname{dist}_{\|.\|}(x, M) := \inf_{m \in M} \|x - m\|$$

Abstand von x zu M.

Definition 1.13 (Bestapproximierende). Ein Element $m \in M$ heißt *Bestapproximierende* von x in M, falls

$$\|x - m\| = \operatorname{dist}(x, M).$$

Schreibweise: $\mathcal{P}_M(x)$:=Menge aller Bestapproximierenden von x in M.

Satz 1.14 (Charakterisierungssatz). *In einem Innenproduktraum $(X, \langle \cdot, \cdot \rangle)$ sei $M \subseteq X$ eine konvexe, nichtleere Menge. Für $m \in M$ und $x \in X$ gilt dann*

$$m \in \mathcal{P}_M(x) \quad \Leftrightarrow \quad (\forall k \in M \quad \langle x - m, k - m \rangle \leq 0).$$

Beweis. $(\Rightarrow)$ Für $k \in M$ und $0 < \lambda < 1$ ist $(1 - \lambda) m + \lambda k \in M$, denn M ist konvex. Wegen $m \in \mathcal{P}_M(x)$ gilt

$$\begin{aligned} 0 &\leq \underbrace{\|x - \lambda k - (1 - \lambda) m\|^2}_{\|x - m + \lambda(m - k)\|^2} - \|x - m\|^2 \\ &= \|x - m\|^2 + \lambda^2 \|m - k\|^2 + 2\lambda \langle x - m, m - k \rangle - \|x - m\|^2 \\ &= -2\lambda \langle x - m, k - m \rangle + \lambda^2 \|m - k\|^2. \end{aligned}$$

Also folgt für $\lambda \searrow 0$

$$\langle x - m, k - m \rangle \leq \frac{1}{2} \lambda \|m - k\|^2 \searrow 0.$$

($\Leftarrow$) Für alle $k \in M$ gilt

$$\begin{aligned} 0 \leq \|x-m\|^2 &= \langle x-m, x-k\rangle + \underbrace{\langle x-m, k-m\rangle}_{\leq 0} \\ &\leq \|x-m\|\,\|x-k\| \end{aligned}$$

Also folgt $x = m$ oder $\|x-m\| \leq \|x-k\|$, d.h. $m \in \mathcal{P}_M(x)$. □

Satz 1.15 (Čebyšev-Eigenschaft). *In einem Hilbert-Raum $(X, \langle\cdot,\cdot\rangle)$ sei $M \subset X$ eine konvexe, abgeschlossene, nichtleere Menge. Dann ist für $x \in X$ die Menge $\mathcal{P}_M(x)$ der Bestapproximierenden einelementig und die Abbildung $P_M : X \to X$, $x \mapsto P_M(x)$ mit $\{P_M(x)\} = \mathcal{P}_M(x)$ ist idempotent, nicht-expansiv und monoton.*

Bemerkung 1.16. $\mathcal{P}_M$ heißt *Čebyšev-Operator* oder *Bestapproximations-Operator* zur Menge M.

Beweis. Eindeutigkeit: Für $x \in X$, $m_1, m_2 \in \mathcal{P}_M(x)$ folgt aus dem Charakterisierungssatz 1.14

$$\begin{aligned} \langle x-m_1, m_2-m_1\rangle &\leq 0 \\ \text{und } \langle x-m_2, m_1-m_2\rangle &\leq 0. \end{aligned}$$

Deren Summe zeigt

$$\|m_2-m_1\|^2 = \langle x-m_1+m_2-x, m_2-m_1\rangle \leq 0,$$

also $\|m_2-m_1\| \leq 0$, d.h. $m_1 = m_2$.

Existenz: Für $j \in \mathbb{N}$, $j > 0$, und $d := \operatorname{dist}(x, M) \geq 0$ existiert ein $m_j \in M$ mit

$$d^2 \leq \|x-m_j\|^2 \leq d^2 + 1/j$$

Aus der Parallelogrammgleichung folgt dann für $j, k \in \mathbb{N}$

$$\begin{aligned} \|m_j-m_k\|^2 + \|m_j-x+m_k-x\|^2 &= 2\|m_j-x\|^2 + 2\|m_k-x\|^2 \\ &\leq 4d^2 + 2/j + 2/k \end{aligned}$$

Da $\frac{m_j+m_k}{2} \in M$ ist

$$4d^2 \leq 4\left\|x - \frac{m_j+m_k}{2}\right\|^2 = \|2x-m_j-m_k\|^2$$

und so

$$\|m_j-m_k\|^2 \leq 2/j + 2/k.$$

Also ist (m_j) Cauchy-Folge in M. Mit der Abgeschlossenheit von M im Hilbert-Raum X konvergiert $(m_j) \to m$ in X für ein $m \in M$. Nach Wahl von m_j folgt

$$\|x-m\| \leq \|x-m_j\| + \|m-m_j\| \leq \sqrt{d^2+1/j} + \|m-m_j\| \xrightarrow{j\to\infty} d,$$

also $\|x-m\| = d$.

Idempotenz: $P_M^2 = P_M$ ist klar.

Monotonie: $0 \leq \langle P_M(x) - P_M(y), x-y\rangle$ für $x, y \in X$.

Für $x, y \in X$ gilt nach dem Charakterisierungssatz 1.14

$$\begin{aligned} \langle P_M(x)-x, P_M(x)-P_M(y)\rangle &\leq 0 \\ \text{und } \langle P_M(y)-y, P_M(y)-P_M(x)\rangle &\leq 0. \end{aligned}$$

Die Summe zeigt

$$0 \leq \langle P_M(y)-P_M(x), P_M(y)-P_M(x)\rangle \leq \langle y-x, P_M(y)-P_M(x)\rangle.$$

Nicht-Expansivität: Mit der Cauchy-Ungleichung folgt in der vorhergegangenen Abschätzung weiter

$$\|P_M(y) - P_M(x)\| \leq \|y - x\|.$$

Also ist P_M Lipschitzstetig mit Lipschitz-Konstante ≤ 1, d.h. P_M ist nicht-expansiv. $\square$

1.6 Orthogonale Komplemente

In einem Hilbert-Raum $(X, \langle\cdot,\cdot\rangle)$ sei U ein linearer, abgeschlossener Teilraum von X (also $U \neq \emptyset$, U konvex und abgeschlossen).

Dann ist

$$P := P_U$$

eine lineare Projektion und orthogonal, d.h. für alle $x \in X$ ist

$$x - Px \perp U.$$

Weiter ist

$$Q := \mathrm{id} - P$$

eine orthogonale Projektion auf das orthogonale Komplement

$$U^{\perp} := \{x \in X \,|x \perp U\}$$

von U. Insbesondere ist $U^{\perp}$ abgeschlossen.

Es ist $\mathrm{Bild}(P) = U$ und $\mathrm{Bild}(Q) = U^{\perp}$. Für alle $x \in X$ gilt

$$x = Px + Qx$$

mit $Px \perp Qx$.

1.7 Riesz-Darstellungssatz

Definition 1.17 (Dualraum, Dualnorm). Für einen normierten linearen Raum $(X, \|\cdot\|)$ heißt

$$X^* := \{f : X \to \mathbb{R} \,|\, f \text{ ist linear und stetig}\}$$

Dualraum zu X und

$$\|\cdot\|_{X^*} : X^* \to \mathbb{R}, \qquad f \mapsto \sup_{\|x\| \leq 1} f(x)$$

heißt *Dualnorm*.

Bemerkung 1.18. $(X^*, \|\cdot\|_{X^*})$ ist ein normierter linearer Raum.

Satz 1.19 (Riesz-Darstellung). *Im Hilbert-Raum $(X, \langle\cdot,\cdot\rangle)$ ist*

$$\Lambda : X \to X^*, \qquad x \mapsto \langle x, \cdot\rangle$$

ein Normisomorphismus (d.h. Λ ist eine lineare Bijektion mit $\|\Lambda x\|_{X^*} = \|x\|_X$ für alle $x \in X$).

Beweis der Surjektivität. Für $f \in X^*$ ist ein $x \in X$ mit $\langle x, \cdot \rangle = f$ zu finden.

Für $f = 0$ ist $x = 0$. Sonst ist $U := \ker f$ abgeschlossen in X und $U \neq X$.

Also existiert ein $z \in U^\perp \setminus \{0\}$ und das Element

$$x := \frac{f(z)}{\|z\|^2} z \in U^\perp$$

erfüllt für alle $y \in X$

$$\begin{aligned}
(\Lambda x)(y) = \langle x, y \rangle &= \|z\|^{-2} f(z) \langle z, y \rangle \\
&= \|z\|^{-2} \langle z, f(z) y \rangle \qquad \text{und } z \perp U \\
&= \|z\|^{-2} \langle z, f(y) z \rangle \quad \text{da } f(z) y - f(y) z \in U \\
&= f(y).
\end{aligned}$$

Also ist $\Lambda x = f$.

Übrige Beweisdetails als Übungsaufgabe. □

1.8 Sobolev-Räume in einer Dimension

Konstruktion geeigneter Hilberträume

Im ganzen Abschnitt sei $\Omega = (a, b)$ für $-\infty < a < b < \infty$ ein offenes, nichtleeres, beschränktes Intervall und es sei $1 \leq p \leq \infty$.

Sobolev-Funktionen in $W^{1,p}(a, b)$ sind absolut stetige (AC) Funktionen mit punktweise klassischen Ableitungen in $L^p(a, b)$.

Definition 1.20 (absolut stetig, AC). Bezeichne X einen normierten linearen Raum. Eine Abbildung $f : (a, b) \to X$ heißt *absolut stetig* (AC), falls

$$\begin{aligned}
&\forall \varepsilon > 0\, \exists \delta > 0 \quad \forall n \in \mathbb{N} \forall a < a_1 < b_1 < a_2 < b_2 < \ldots < a_n < b_n < b \\
&\text{mit} \quad \textstyle\sum_{j=1}^{n} (b_j - a_j) < \delta \text{ gilt } \displaystyle\sum_{j=1}^{n} \|f(b_j) - f(a_j)\| < \varepsilon.
\end{aligned}$$

Im Rest dieses Abschnittes sei $X = \mathbb{R}$.

Bemerkung 1.21. Es gelten die folgenden und keine weiteren Inklusionen

$$\mathrm{Lip}(a, b) \subsetneqq \mathrm{AC}(a, b) \subsetneqq C^0[a, b] \subsetneqq C^0(a, b)$$

Die Inklusion $\mathrm{AC}(a, b) \subset C^0[a, b]$ gilt, weil für $f \in \mathrm{AC}(a, b)$ ist f gleichmäßig stetig und damit stetig und zu $\hat{f} \in C[a, b]$ fortsetzbar. Insbesondere macht es keinen wesentlichen Unterschied, absolut stetige Funktionen auf offenen oder abgeschlossenen (kompakten) Intervallen zu betrachten.

Satz 1.22 (Hauptsatz der Infinitesimalrechnung). *Für eine Funktion $f : [a, b] \to \mathbb{R}$ sind die folgenden Aussagen (a), (b) und (c) paarweise äquivalent.*

(a) f ist absolut stetig.

(b) Es existiert eine Funktion $g \in L^1(a, b)$, so daß für alle $x \in [a, b]$

$$f(x) - f(a) = \int_a^x g(y)\, dy.$$

(c) *f ist fast überall differenzierbar, $f' \in L^1(a,b)$ und für alle $x \in [a,b]$ ist*

$$f(x) - f(a) = \int_a^x f'(y)\,dy.$$

Bemerkung 1.23 (zum Beweis). $(c) \Rightarrow (b)$ trivial

$(b) \Rightarrow (a)$ folgt mit non-concentration of Lebesgue functions

$(a) \Rightarrow (c)$ Folge von Argumenten:

> f ist absolut stetig $\Rightarrow$ f ist von beschränkter Variation[1] $\Rightarrow f = F_1 - F_2$ für zwei monotone Funktionen F_1 und $F_2 \Rightarrow f'$ existiert fast überall und $|f'| \leq |F_1'| + |F_2'| \Rightarrow f' \in L^1(a,b)$ und schließlich $f(x) - f(a) = \int_a^x f'(y)\,dy$.

Literatur: Real Analysis, H.L. Royden, Prentice Hall, 1988[3], Kapitel 5.

Beispiel. Die Cantor-Funktion ist stetig und monoton wachsend, also von beschränkter Variation. Die Ableitung ist gleich 0 fast überall. Die Funktion ist aber <u>nicht</u> absolut stetig, denn $f(1) - f(0) = 1 - 0 \neq 0 = \int_0^1 f'(x)\,dx$.

Bemerkung 1.24. Ist f absolut stetig mit $f' = 0$ fast überall, so ist f konstant.

Bemerkung 1.25. Ist f von beschränkter Variation und f' existiert <u>für</u> <u>alle</u> x, so ist f absolut stetig. (Literatur: W. Rudin: Real and Complex Analysis, 1987[3] McGraw-Hill, Thm. 7.21 auf S. 149)

Definition 1.26 (Sobolev-Räume und -Normen). Für $1 \leq p \leq \infty$ und für alle $f \in L^p(a,b)$ gilt

$$W^{0,p}(a,b) := L^p(a,b) \text{ und } \|f\|_{W^{0,p}(a,b)} := \|f\|_{0,p} := \|f\|_p := \|f\|_{L^p(a,b)}.$$

Weiter sei

$$W^{1,p}(a,b) := \{f \in AC(a,b) \,|\, f' \in L^p(a,b)\}$$

und für alle $f \in W^{1,p}(a,b)$ gilt

$$\|f\|_{W^{1,p}(a,b)} := \|f\|_{1,p} := \begin{cases} \left(\|f\|_p^p + \|f'\|_p^p\right)^{\frac{1}{p}} & \text{falls } p < \infty, \\ \max\{\|f\|_\infty, \|f'\|_\infty\} & \text{falls } p = \infty. \end{cases}$$

Für $m = 2, 3, 4, \ldots$ definiert man weiter

$$W^{m,p}(a,b) := \left\{f \in AC(a,b) \,\middle|\, f' \in W^{m-1,p}(a,b)\right\}$$

und für alle $f \in W^{m,p}(a,b)$ ist

$$\|f\|_{W^{m,p}} := \|f\|_{m,p} := \begin{cases} \left(\|f\|_p^p + \|f'\|_{m-1,p}^p\right)^{\frac{1}{p}} & \text{falls } p < \infty, \\ \max\left\{\|f\|_\infty, \|f'\|_{m-1,\infty}\right\} & \text{falls } p = \infty. \end{cases}$$

Zusätzlich seien folgende Abkürzungen definiert:

$$W_0^{m,p}(a,b) := \left\{ f \in W^{m,p}(a,b) \,\middle|\, \begin{array}{rcl} f(a) = f'(a) = \ldots = f^{(m-1)}(a) & = & 0 \\ f(b) = f'(b) = \ldots = f^{(m-1)}(b) & = & 0 \end{array} \right\}$$

---

[1] Eine Funktion f ist von *beschränkter Variation* (BV), wenn für alle $a \leq x \leq b$

$$F(x) := \sup\left\{ \sum_{j=1}^n |f(x_j) - f(x_{j-1})| \,\middle|\, n \in \mathbb{N}, x_0 = a < x_1 < x_2 < \ldots < x_n = x \right\} < \infty.$$

sowie

$$H^m := W^{m,2} \quad \text{und} \quad H_0^m := W_0^{m,2}.$$

Das Skalarprodukt in $H^1(a,b)$ lautet

$$\langle f,g\rangle_{H^1(a,b)} := (f,g)_{L^2(a,b)} + (f',g')_{L^2(a,b)} = \int_a^b fg dx + \int_a^b f'g' dx.$$

Satz 1.27. *$W^{m,p}$ und $W_0^{m,p}$ sind Banach-Räume. H^m und H_0^m sind Hilbert-Räume.*

Beweis der Vollständigkeit von $W^{1,p}$. Für eine Cauchy-Folge (f_j) in $W^{1,p}$ sind (f_j) und $\left(f_j'\right)$ Cauchy-Folgen in $L^p(\Omega)$ und daher existieren f und g in $L^p(\Omega)$ mit $(f_j) \to f$ in L^p und $\left(f_j'\right) \to g$ in $L^p(\Omega)$.

Behauptung: $f' = g$.

Beweis: Für eine Teilfolge (f_{j_k}) und für fast alle α, β mit $a \le \alpha \le \beta \le b$ gilt

$$\begin{aligned}\lim_{k\to\infty} f_{j_k}(\alpha) &= f(\alpha),\\ \lim_{k\to\infty} f_{j_k}(\beta) &= f(\beta)\end{aligned}$$

(punktweise fast-überall Konvergenz von einer Teilfolge einer in L^p konvergenten Folge). Dann folgt

$$\begin{aligned} f_{j_k}(\beta) - f_{j_k}(\alpha) &= \int_\alpha^\beta f_{j_k}'(x)\,dx\\ \text{und für } k\to\infty \text{ gilt}\quad f(\beta) - f(\alpha) &= \int_\alpha^\beta g(x)\,dx.\end{aligned}$$

Mit dem Hauptsatz der Infinitesimalrechnung (Satz 1.22) folgt, daß

$$\int_\alpha^\beta g(x)\,dx = f(\beta) - f(\alpha)$$

absolut stetig in β ist und $f'(x) = g(x)$ für fast alle x gilt. □

Satz 1.28 (Produktregel). *Für absolut stetige Funktionen f und g ist fg absolut stetig und*

$$(fg)' = f'g + fg'.$$

Satz 1.29 (Partielle Integration). *Für Funktionen $f, g \in \mathrm{AC}(a,b)$ ist*

$$\int_a^b (f'g + fg')\,dx = [fg]_a^b := (fg)(b) - (fg)(a)$$

Anmerkung (Poincaré-Ungleichung). Für $f \in H^1(a,b)$ sei

$$F := \fint_a^b f(x)\,dx := \frac{1}{b-a}\int_a^b f(x)\,dx.$$

Dann ist

$$\|f - F\|_{L^2(a,b)} \le \frac{b-a}{\pi}\|f'\|_{L^2(a,b)} \tag{1.1}$$

Beweisidee: O.B.d.A. sei $F = 0$. Dann existiert ein $\xi \in (a,b)$, so daß $f(x) = \int_\xi^x f'(y)\,dy$. Mit der Hölder-Ungleichung folgt

$$|f(x)|^2 = \left|\int_\xi^x 1\, f'(y)\,dy\right|^2 \le |x-\xi|\ \|f'\|_{L^2(a,b)}^2.$$

Daraus ergibt sich

$$\|f\|_2^2 \leq \|f'\|_{L^2(a,b)}^2 \underbrace{\max_{a\leq\xi\leq b} \int_a^b |x-\xi|\, dx}_{\leq \frac{1}{2}(b-a)^2}.$$

Anmerkung (Friedrichs-Ungleichung). Für $f \in H_0^1(a,b)$ ist

$$\|f\|_2 \leq \frac{b-a}{\pi} \|f'\|_2 . \tag{1.2}$$

Bemerkung 1.30. Die Konstante $\frac{b-a}{\pi}$ ist scharf: $(b-a)$ folgt aus einem Skalierungsargument. $\frac{1}{\pi}$ ist Eigenwert und wird angenommen für translatiert-skalierte sin- und cos-Funktionen.

Satz 1.31 ($W^{1,p} = H^{1,p}$)**.** *Für $1 \leq p < \infty$ gilt:*

(a) Polynome, (stückweise affine, globalstetige Funktionen) u.v.a. mehr liegen dicht in $W^{1,p}(a,b)$.

(b) $W^{1,p}(a,b)$ ist ein separabler Banach-Raum (hat abzählbar dichte Teilmenge).

(c) $W^{1,p}(a,b)$ ist die Vervollständigung z.B. von $C^1[a,b]$ bzgl. der Sobolev-Norm $\|\cdot\|_{1,p}$,

$$\overline{C^1[a,b]}^{\|\cdot\|_{1,p}} = W^{1,p}(a,b).$$

Bemerkung 1.32. Für $p=\infty$ ist

$$\overline{C^1[a,b]}^{\|\cdot\|_{1,\infty}} = C^1[a,b] \subsetneq W^{1,\infty}(a,b) = \mathrm{Lip}(a,b).$$

Bemerkung 1.33. Analoge Aussagen gelten für $W_0^{1,p}(a,b)$.

Beweis zu (a). Die Menge der Polynome oder die Menge der Treppenfunktionen liegen dicht in $L^p(a,b)$ und daher deren Stammfunktionen dicht in $W^{1,p}(a,b)$: Zu $f \in W^{1,p}(a,b)$ und zu $\varepsilon > 0$ wähle man ein z.B. ein Polynom g mit

$$\|f'-g\|_p < \varepsilon.$$

Dann definiert

$$G(x) := f(a) + \int_a^x g(y)\, dy$$

ein Polynom mit $G' = g$ fast überall und

$$\begin{aligned} |f(x) - G(x)| &= \left| f(a) + \int_a^x f'(y)\, dy - f(a) - \int_a^x g(y)\, dy \right| \\ &\leq \int_a^x |f'(y) - g(y)|\, dy \\ &\leq \underbrace{|x-a|^{\frac{1}{q}}}_{\leq (b-a)^{\frac{1}{q}}} \|f'-g\|_p \qquad \text{für } \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

Also folgt

$$\|f-G\|_p \leq \underbrace{(b-a)^{\frac{1+\frac{1}{q}}{p}}}_{=:c_1} \underbrace{\|f'-g\|_p}_{<\varepsilon} \leq c_1 \varepsilon$$

und weiter

$$\|f-G\|_{1,p} \leq c_2\, \varepsilon.$$

□

Bemerkung 1.34. Zu $f \in W_0^{1,p}$ ist $G \in W_0^{1,p}$ konstruierbar.

1.9 Eindimensionale Randwertprobleme

Der Riesz-Darstellungssatz zeigt: Es existiert genau eine schwache Lösung.

In diesem Abschnitt sei $\Omega = (a,b)$ mit $-\infty < a < b < \infty$.

Definition 1.35 (Starke Form des Randwertproblems (RWP)). Zu $f \in L^2(\Omega)$ und zu $\lambda \in \mathbb{R}$ finde man ein $u \in H_0^1(\Omega) \cap H^2(\Omega)$ mit

$$-u'' + \lambda u = f \qquad \text{f.ü. in } \Omega. \tag{1.3}$$

Ein solches u heißt dann *starke Lösung* des RWP.

Umformulierung: Multiplikation mit beliebigem $v \in V := H_0^1(\Omega)$ und Integration über (a,b) gefolgt von partieller Integration zeigt

$$\int_a^b (f - \lambda u)\, v dx = -\int_a^b u'' v dx = \int_a^b u'v'dx$$

(keine Randterme, da $v(a) = v(b) = 0$). Also gilt

$$a(u,v) = (f,v) \tag{1.4}$$

für $(\cdot,\cdot) := (\cdot,\cdot)_{L^2(a,b)}$ und

$$a(u,v) := (u',v') + \lambda(u,v).$$

Definition 1.36 (Schwache Form des RWP). Zu $f \in L^2(\Omega)$ (und zu $\lambda \in \mathbb{R}$) finde man ein $u \in V$ mit (1.4) für alle $v \in V$. Ein solches u heißt dann *schwache Lösung* des RWP.

Satz 1.37 (Äquivalenz der starken und schwachen Form des RWP).

(a) Jedes $u \in V \cap W^{2,1}(a,b)$ mit (1.3) erfüllt (1.4).

(b) Jede schwache Lösung u mit $u \in W^{2,1}(a,b)$ ist auch starke Lösung.

Beweis. $(a) \Rightarrow (b)$: Folgt aus obiger Herleitung.

$(b) \Rightarrow (a)$: Bezeichne u eine schwache Lösung mit $u \in V \cap W^{2,1}(\Omega)$. Die oben ausgeführte partielle Integration zeigt

$$(f,v) = (-u'' + \lambda u, v) \quad \text{für alle } v \in V.$$

Dann folgt (1.3) aus dem Fundamentallemma der Variationsrechnung. $\square$

Lemma 1.38 (Fundamentallemma der Variationsrechnung). *Für $g \in L^1(a,b)$ gelte*

$$\int_a^b g(x)\,\varphi(x)\,dx = 0 \qquad \text{für alle } \varphi \in C_c^\infty(a,b).$$

Dann ist $g = 0$ fast überall in (a,b).

Bemerkung 1.39. Für $g \in L^2(a,b)$ folgt dies aus $\overline{C_c^\infty(a,b)}^{\|\cdot\|_2} = L^2(a,b)$.

Beweis. Für $a < \alpha < \beta < b$ und die charakteristische Funktion $\chi_{(a,b)}$ betrachtet man die Regularisierung

$$\varphi_\varepsilon := \eta_\varepsilon * \chi_{(\alpha,\beta)}$$

für hinreichend kleines $\varepsilon > 0$,

$$\varphi_\varepsilon(x) := \int_\alpha^\beta \eta_\varepsilon(x-y)\,dy \qquad \text{für } x \in (a,b).$$

Die Funktion η_ε heißt der Standard-Mollifier in $C_c^\infty(\mathbb{R})$ mit $\operatorname{supp}\eta_\varepsilon = [-\varepsilon, \varepsilon]$,

$$\eta_\varepsilon(x) := \begin{cases} \varepsilon^{-1} c \exp\left(\dfrac{1}{(x/\varepsilon)^2 - 1}\right) & \text{falls } |x| < \varepsilon, \\ 0 & \text{falls } |x| \geq \varepsilon, \end{cases}$$

wobei die positive Konstante c so gewählt ist, daß $\int_{\mathbb{R}} \eta_\varepsilon(x)\, dx = 1$ gilt.
Man rechnet nach

$$\varphi_\varepsilon(x) \xrightarrow{\varepsilon \searrow 0} \chi_{(\alpha,\beta)}(x) \qquad \text{für fast alle } x \in \Omega$$

und

$$0 = \int_a^b g(x)\, \varphi_\varepsilon(x)\, dx \xrightarrow{\varepsilon \searrow 0} \int_\alpha^\beta g(x)\, dx.$$

Also $\int_\alpha^\beta g(x)\, dx = 0$ für alle $\alpha, \beta \in \Omega$.
Da α und β beliebig gewählt waren, verschwindet g in allen Lebesgue-Punkten. Also ist $g = 0$ fast überall. □

Satz 1.40 (Existenz und Eindeutigkeit schwacher Lösungen). *Sofern*

$$\lambda > -\frac{\pi^2}{(b-a)^2}$$

existiert genau eine schwache Lösung u in V.

Bemerkung 1.41. Diese schwache Lösung u ist dann in der Tat auch eine starke Lösung, d.h. $u \in W^{2,1}(a,b)$, nach dem Regularitätssatz 1.44 auf der nächsten Seite.

Beweis. Da $F := (f, \cdot)_{L^2(\Omega)} \in V^* =: H^{-1}(\Omega)$ genügt es mit dem Riesz-Darstellungssatz (Satz 1.19 auf Seite 9) nachzuweisen, daß

1. a ein Skalarprodukt ist und
2. dessen induzierte Norm $\|\cdot\|_a$ äquivalent zu $\|\cdot\|_{1,2}$ ist.

Nach Definition ist

$$a(v,w) := (v', w') + \lambda (v, w) \qquad \text{für alle } v, w \in V.$$

1. Bilinearität und Symmetrie von a sind direkt einzusehen. Die Positivität und Definitheit folgen mit der Friedrichs-Ungleichung ((1.2) auf Seite 13: $\left(\frac{\pi}{b-a}\right)^2 \|v\|_2^2 \leq \|v'\|_2^2$):

$$\underbrace{\left(\lambda + \left(\frac{\pi}{b-a}\right)^2\right)}_{>0} \|v\|_2^2 \leq \|v'\|_2^2 + \lambda \|v\|_2^2 = a(v,v),$$

also $a(v,v) \geq 0$ und falls $a(v,v) = 0$, so folgt $\|v\|_2 = 0$, also $v = 0$.

2. $a(v,v) \leq \max\{1, \lambda\} \|v\|_{1,2}^2$ und wie oben folgt für jedes $0 < \mu < 1$

$$a(v,v) \geq \underbrace{\left(\lambda + \mu \left(\frac{\pi}{b-a}\right)^2\right)}_{>0 \text{ für } \mu \in (0,1) \text{ groß genug}} \|v\|_2^2 + (1-\mu) \|v'\|_2^2 \gtrsim \|v\|_{1,2}^2$$

("$A \lesssim B$" bezeichnet "$A \leq C \cdot B$" mit einer universellen Konstante $C > 0$). □

Bemerkung 1.42. Der Beweis benutzt $C_F = \frac{b-a}{\pi}$. Jede Schranke von C_F führt auf ein $\lambda_0 < 0$, so daß a für $\lambda > \lambda_0$ ein Skalarprodukt mit $\|\cdot\|_a \approx \|\cdot\|_{1,2}$ definiert. ("$A \approx B$" bezeichnet "$A \lesssim B \lesssim A$").

Bemerkung 1.43. Die Konstante $-\frac{\pi^2}{(b-a)^2}$ im Satz 1.40 ist scharf in dem Sinne, daß für jedes $\mu \in \mathbb{R}$

$$u(x) := \mu \sin\left(\frac{x-a}{b-a}\pi\right)$$

eine starke Lösung von (1.3) mit $\lambda = -\frac{\pi^2}{(b-a)^2}$ und $f \equiv 0$ ist (also die Eindeutigkeit verloren geht).

Jede schwache Lösung ist hier sogar starke Lösung:

Satz 1.44 (Höhere Regularität). *Für $\lambda \in \mathbb{R}$ und $u \in V$ gelte*

$$a(u,v) := (u', v')_{L^2(\Omega)} + \lambda (u,v)_{L^2(\Omega)} \stackrel{!}{=} (f,v) \qquad \text{für alle } v \in V.$$

Dann gilt das "Shift-Theorem" :

$$f \in H^m(\Omega), m \in \mathbb{N}_0 \quad \Rightarrow \quad u \in H^{m+2}(\Omega).$$

Beweis. Mit $g := f - \lambda u \in L^2(\Omega)$ gilt

$$(u', v') = (g, v) \qquad \text{für alle } v \in V.$$

Zu g definiere man $G \in H^1(\Omega)$ vermöge

$$G(x) := \int_a^x g(y)\,dy - \underbrace{\frac{1}{b-a}\int_a^b \int_a^y g(z)\,dzdy}_{\text{Integralmittel von } x \mapsto \int_a^x g(z)dz} \qquad \text{für alle } x \in \Omega.$$

Dann ist $G' = g$ und $\int_a^b G(x)\,dx = 0$. Die Definition

$$v(x) := u(x) + \int_a^x G(y)\,dy \qquad \text{für alle } x \in \Omega$$

ergibt ein $v \in V$ mit $v' = u' + G$ und

$$(u', u' + G) = (u', v') = (g, v) = (G', v).$$

Mit partieller Integration folgt

$$(u', u' + G) = -(G, v') = -(G, G + u')$$

und so insgesamt

$$(u' + G, u' + G)_{L^2(\Omega)} = 0.$$

Also ist $u' = -G \in H^1(\Omega)$.

Dies ist die Behauptung für $m = 0$. Folglich gilt die starke Form $-u'' + \lambda u = f$ und sukzessives Weiterdifferenzieren liefert die Behauptung für $m = 1, 2, 3, \ldots$ □

Bemerkung 1.45. Eigenvektoren sind beliebig glatt!

Bemerkung 1.46. Für "glatte Gebiete" gilt ein Shift-Theorem auch in höheren Raumdimensionen, aber mit anderem Beweis.

1.10 Eindimensionale Finite Elemente Methode

Ziel. Ritz-Projektion auf endlich-dimensionale Teilräume.

Definition 1.47 (P_k-FEM). 1. Zu $\Omega = (0,1)$ und einer *Triangulierung*

$$\mathcal{T} = \{I_j \,|\, j = 1, \ldots, m\}$$

in abgeschlossene Intervalle $I_j = [z_{j-1}, z_j]$ bzgl. der *Knoten* $0 = z_0 < z_1 < \ldots < z_N < z_{N+1} = 1$, $m = N+1$, $N, k \in \mathbb{N}$, definiert man

$$\begin{aligned} P_k(I) &:= \{f : I \to \mathbb{R} \,|\, f \text{ algebraisches Polynom vom Grad} \leq k\}, \\ P_k(\mathcal{T}) &:= \{f \in L^\infty(\Omega) \,|\, \forall I \in \mathcal{T}, f|_I \in P_k(I)\} \text{ und} \\ V_h &:= V \cap P_k(\mathcal{T}) \text{ für } k = 1, 2, 3, \ldots, \end{aligned}$$

wobei $V = H_0^1(\Omega)$ im Randwertproblem (1.3) auf Seite 14.

V_h heißt P_k-FE-Raum (P_k-FER) und hat die Dimension $(k+1)m - (m+1) = km - 1$.

2. Zu der schwachen Form eines Randwertproblems

$$a(u, v) = F(v) \qquad \text{für alle } v \in V \tag{1.5}$$

mit schwacher Lösung $u \in V \cap H^2(\Omega)$ im Hilbert-Raum (V, a) (also $\lambda > -\frac{\pi^2}{(b-a)^2}$, Satz 1.40) und $F \in V^*$ und dem P_k-FER V_h bzgl. der Triangulierung $\mathcal{T}$ heißt ein $u_h \in V_h$ *diskrete Lösung* (oder FE-Lösung), falls

$$a(u_h, v_h) = F(v_h) \qquad \text{für alle } v_h \in V_h. \tag{1.6}$$

3. Der Übergang von (1.5) nach (1.6) heißt hier speziell *Galerkin-Diskretisierung* und für den FER V_h auch *FEM.* (1.5) heißt *kontinuierliches*, (1.6) heißt *diskretes Problem* mit kontinuierlicher bzw. diskreter Lösung.

Bemerkung 1.48. Anstatt der Bezeichnung Triangulierung findet man in der Literatur auch Partition, Netz, Gitter, mesh, grid, triangulation.

Vorsicht: Details können anders sein.

Satz 1.49 (Ritz-Projektion).

(a) (1.5) und (1.6) haben jeweils eindeutige Lösungen.

(b) u_h ist die Bestapproximierende zu u im Hilbert-Raum (V, a).

Definition 1.50 (Ritz-Projektion). Die Abbildung $P_{V_h} : V \to V_h, u \mapsto u_h$ heißt Ritz-Projektion.

Bemerkung 1.51. Die Gleichheit $\|u - u_h\|_a = \min_{v_h \in V_h} \|u - v_h\|_a$ heißt in der FE-Literatur auch Cea-Lemma.

Beweis. (a) V_h ist als endlichdimensionaler Vektorraum ein vollständiger Unterraum von V, also ist $(V_h, a|_{V_h \times V_h})$ ein Hilbert-Raum. Der Riesz-Darstellungssatz 1.19 auf Seite 9 zeigt: Es existiert genau eine Riesz-Darstellung u_h von $F|_{V_h} \in V_h^*$.

(b) Für alle $v_h \in V_h$ ist $a(u_h, v_h) = F(v_h) = a(u, v_h)$, d.h. $u - u_h \perp_a V_h$. Damit folgt 2. aus dem Charakterisierungssatz 1.14 auf Seite 7 (jetzt für Unterräume). $\square$

Beispiel (uniforme P_1-FEM). Es sei $k = 1$, $\dim V_h = n$, $h = (n+1)^{-1}$, $n \in \mathbb{N}$, $n \geq 1$ und

$$z_j := jh \quad \text{für } j = 0, \ldots, n+1.$$

Man definiere $\varphi_j \in V_h$ vermöge

$$\varphi_j(z_k) := \delta_{jk} := \begin{cases} 1 & \text{für } j = k, \\ 0 & \text{sonst,} \end{cases}$$

d.h.

$$\varphi_j(z) = \begin{cases} \frac{z - z_{j-1}}{h} & \text{für } z_{j-1} \leq z < z_j, \\ \frac{z_{j+1} - z}{h} & \text{für } z_j \leq z < z_{j+1}, \\ 0 & \text{sonst,} \end{cases} \qquad \varphi_j'(z) = \begin{cases} \frac{1}{h} & \text{für } z_{j-1} \leq z < z_j, \\ -\frac{1}{h} & \text{für } z_j \leq z < z_{j+1}, \\ 0 & \text{sonst.} \end{cases}$$

Für die diskrete Lösung

$$u_h = \sum_{j=1}^{n} x_j \varphi_j$$

mit Koeffizientenvektor $x = (x_1, \ldots, x_n)^T$ ist (1.6) äquivalent zu

$$\sum_{j=1}^{n} x_j \underbrace{a(\varphi_j, \varphi_i)}_{A_{ji}} = \underbrace{F(\varphi_i)}_{b_i} \qquad i = 1, \ldots, n,$$

also

$$Ax = b$$

für eine positiv definite $n \times n$-Matrix A und mit einem durch F gegebenen Vektor der rechten Seite. Für $\lambda = 0$ ist

$$a(\varphi_j, \varphi_k) = \int_0^1 \varphi_j'(x)\, \varphi_k'(x)\, dx = \begin{cases} 0 & \text{falls } |j-k| > 1, \\ -\frac{1}{h} & \text{falls } |j-k| = 1, \\ \frac{2}{h} & \text{falls } j = k, \end{cases}$$

$$A = (a(\varphi_j, \varphi_k))_{j,k} = \frac{1}{h} \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

und

$$b_j = F(\varphi_j) = \int_\Omega f(x)\, \varphi_j(x)\, dx \approx f(z_j)\, h,$$

also

$$\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} x = hb \approx h^2 \begin{pmatrix} f(z_1) \\ \vdots \\ \vdots \\ f(z_n) \end{pmatrix}.$$

Definition 1.52 (Nodale Interpolation, nodale Basis). Für ein Triangulierung $\mathcal{T}$ von $(0,1) = \Omega$ in abgeschlossene Intervalle und die P_1-FEM heißt der Operator

$$I : C[0,1] \to V_h, \quad f \mapsto \sum_{j=1}^{N} f(x_j)\, \varphi_j$$

nodale Interpolation. Dabei sind $z_0 = 0 < z_1 < \ldots < z_{N+1} = 1$ die *Knoten* der Triangulierung,

$$\mathcal{N} := \{z_0, \ldots, z_{N+1}\}.$$

$(\varphi_1, \ldots, \varphi_N)$ heißt *nodale Basis* von $V_h = P_1(\mathcal{T}) \cap V$ definiert vermöge $\varphi_j(z_k) = \delta_{jk}$ für $j, k = 1, \ldots, N$.

Definition 1.53 (Netzweite, Gewichtsfunktion). Zu einer Triangulierung $\mathcal{T}$ definiere

$$h_j := |I_j| := \text{Länge von } I_j = z_j - z_{j-1}$$

die *Netzweite.* Die *Gewichtsfunktion* $h \in P_0(\mathcal{T})$ ist definiert vermöge

$$h(x) := h_j \quad \text{für } j \text{ so, daß } x \in \operatorname{int} I_j = (z_{j-1}, z_j).$$

Satz 1.54 (Approximationssatz). *Für $u \in H^2(\Omega)$ gilt*

$$\begin{aligned} \|u - Iu\|_{L^2(\Omega)} &\leq \frac{1}{\pi^2} \left\|h^2 u''\right\|_{L^2(\Omega)}, \\ \left\|(u - Iu)'\right\|_{L^2(\Omega)} &\leq \frac{1}{\pi} \|hu''\|_{L^2(\Omega)}. \end{aligned}$$

Beweis. Die Anwendung der Friedrichs-Ungleichung (1.2) auf I_j für $u - Iu$ zeigt

$$\|u - Iu\|_{L^2(I_j)} \leq \frac{h_j}{\pi} \left\|(u - Iu)'\right\|_{L^2(I_j)}.$$

$(Iu)'\big|_{I_j}$ ist das Integralmittel von u' auf I_j, denn

$$\int_{z_{j-1}}^{z_j} (u - Iu)'\, dx = [u - Iu]_{z_{j-1}}^{z_j} = 0.$$

Dann zeigt die Poincaré-Ungleichung (1.1) auf I_j für $(u - Iu)'$:

$$\left\|(u - Iu)'\right\|_{L^2(I_j)} \leq \frac{h_j}{\pi} \left\|(u - Iu)''\right\|_{L^2(I_j)} = \frac{h_j}{\pi} \|u''\|_{L^2(I_j)}.$$ □

<u>Aufgabe</u>: Für welche $\lambda \in \mathbb{R}$ gilt $u_h = Iu$ in 1D?

Obwohl die exakte Lösung u unbekannt ist, liefert die FEM eine diskrete Lösung u_h als Bestapproximierende zu u vermöge der Lösung eines linearen Gleichungssystems mit einer SPD-Tridiagonalmatrix (in 1D) und es gilt

$$\begin{aligned} \|u - u_h\|_a &= \sqrt{\|u' - u_h'\|_2^2 + \lambda \|u - u_h\|_2^2} \\ &\leq \sqrt{\underbrace{\left\|u' - (Iu)'\right\|_2^2}_{\leq \frac{1}{\pi^2}\|hu''\|_2^2} + \lambda \underbrace{\|u - Iu\|_2^2}_{\leq \|h\|_\infty^2 \|hu''\|_2^2 \frac{1}{\pi^4}}} \\ &\leq \underbrace{\sqrt{\frac{1}{\pi^2} + |\lambda| \frac{\|h\|_\infty^2}{\pi^4}}}_{\text{const}} \|hu''\|_2. \end{aligned}$$

Das bedeutet lineare Konvergenz in h, denn für $f \in L^2(\Omega)$ ist $u \in H^2(\Omega)$.

Eine effektive Netzweitensteuerung sollte $\|hu''\|_2$ effizient klein machen, also h dort klein bzw. groß machen wo u'' groß bzw. klein ist.

1.11 Eindimensionale Variationsungleichungen

Ziel. Hindernisproblem mit exakter und diskreter Lösung kennenlernen.

In diesem Abschnitt sei $\Omega = (a, b)$ und $V = H_0^1(\Omega)$ wie in §1.9. Weiter sei $f \in L^2(\Omega)$ und $\chi \in H^1(\Omega)$ mit

$$K = \{v \in V \,|\, v \leq \chi \text{ f.ü. in } \Omega\} \neq \emptyset.$$

Letzteres bedeutet $0 \leq \chi(a)$ und $0 \leq \chi(b)$. (Anmerkung: $\min\{\chi, 0\} \in K$)

Notation. $a(v, w) := \int_\Omega v'w'dx$ (also $\lambda = 0$ in (1.4), Seite 14), $\|\cdot\|_a := a(\cdot,\cdot)^{\frac{1}{2}}$.

Lemma 1.55. *Die Menge K ist abgeschlossen und konvex im Hilbert-Raum (V, a).*

Beweis. Konvexität: Für $v, w \in K$ und $0 < \lambda < 1$ ist

$$\lambda v + (1-\lambda) w \leq \lambda\chi + (1-\lambda)\chi = \chi \quad \text{f.ü.},$$

also $\lambda v + (1-\lambda) w \in K$.

Vollständigkeit: Für eine Cauchy-Folge (v_j) in $K^{\mathbb{N}}$ ist (v_j) eine Cauchy-Folge im Hilbert-Raum (V, a), also existiert genau ein $v \in V$, so daß $(v_j) \to v$. Dann existiert eine Teilfolge (v_{j_k}), die punktweise fast überall gegen v konvergiert,

$$\lim_{k\to\infty} \underbrace{v_{j_k}(x)}_{\leq\chi(x)} = v(x).$$

Daher gilt $v(x) \leq \chi(x)$ für fast alle $x \in \Omega$, also $v \in K$. □

Notation. Mit $u_f \in V$ sei der Riesz-Darsteller von $F := (f, \cdot)_{L^2(\Omega)} \in V^*$ bezeichnet, d.h.

$$a(u_f, \cdot) = (f, \cdot)_{L^2(\Omega)}.$$

Es ist $u_f \in V \cap H^2(\Omega)$. Warnung: Im allgemeinen ist $u_f \notin K$.

Satz 1.56. *(i) Die folgenden Aussagen (a)-(c) sind für jedes $u \in K$ paarweise äquivalent:*

(a) u ist Bestapproximierende zu u_f in K, d.h.

$$\|u_f - u\|_a \leq \|u_f - v\|_a \qquad \textit{für alle } v \in K. \tag{1.7a}$$

(b) u löst die Variationsungleichung

$$F(v-u) \leq a(u, v-u) \qquad \textit{für alle } v \in K. \tag{1.7b}$$

(c) u minimiert das Funktional

$$E(v) := \frac{1}{2}a(v,v) - F(v) \qquad \textit{unter allen } v \in K. \tag{1.7c}$$

(ii) Es existiert genau ein $u \in K$ mit (1.7a), (1.7b) oder (1.7c) [und dann mit (a), (b) und (c)].

Beweis. (i) (a) $\Leftrightarrow$ (b): Folgt aus dem Charakterisierungssatz 1.14 auf Seite 7: Für alle $v \in K = M$ gilt

$$0 \geq a(u_f - u, v-u) = F(v-u) - a(u, v-u).$$

(ii): Folgt aus der Čebyšev-Eigenschaft 1.15 auf Seite 8 für Bestapproximierende.

(i) (a) $\Leftrightarrow$ (c): Es ist

$$\begin{aligned} E(v) - E(u) &= \frac{1}{2} a(v,v) - a(u_f, v) - \frac{1}{2} a(u,u) + a(u_f, u) \\ &= \frac{1}{2} \|u_f - v\|_a^2 - \frac{1}{2} \|u_f - u\|_a^2 . \end{aligned}$$

□

Notation. Für die Lösung u definiere man

$$\begin{aligned} \rho &:= \underbrace{F - a(u,\cdot)}_{=a(u_f - u,\cdot)} \in V^* =: H^{-1}(\Omega) = \left(H_0^1(\Omega)\right)^*, \\ \{u = \chi\} &:= \{x \in \Omega \,|\, u(x) = \chi(x)\} =: C \end{aligned}$$

den Kontaktbereich (relativ abgeschlossen, möglicherweise leer) und

$$\{u < \chi\} := \Omega \setminus C = \{x \in \Omega \,|\, u(x) < \chi(x)\} = \bigcup \mathcal{I}$$

den (offenen) Nicht-Kontaktbereich für eine Familie offener nichtleerer Intervalle $\mathcal{I}$, den Zusammenhangskomponenten von $\Omega \setminus C$.

Satz 1.57. *Der Operator ρ ist nicht-negativ (in Zeichen $\rho \geq 0$) in dem Sinne, daß für alle $\varphi \in V$ mit $\varphi \geq 0$ fast überall gilt $\rho(\varphi) \geq 0$.*

Für fast alle $x \in \Omega$ gilt

$$(u - \chi)(x) \cdot (f + u'')(x) = 0,$$

wobei $u \in H^2(\mathcal{I}) := \left\{v \in L^2(\Omega) \,\middle|\, \forall I \in \mathcal{I}, V|_I \in H^2(I)\right\}$.

Beweis. Für $0 \leq \varphi \in V$ ist $v := u - \varphi \leq \chi$, also $v \in K$. Dann ist

$$\rho(-\varphi) = a(u_f - u, v - u) \leq 0, \quad \text{also } \rho \geq 0.$$

Für ein Intervall $I = (\alpha, \beta) \in \mathcal{I}$ und $v \in C_c^1(I) := \left\{w \in C^1[\alpha,\beta] \,|\, \operatorname{supp} w \Subset I\right\}$ (dicht in V; "$\Subset$": kompakt enthalten) gilt für ein $\delta > 0$ auf einer kompakten Teilmenge $S := \operatorname{supp} v = \overline{\{v \neq 0\}} \Subset (\alpha, \beta)$, daß

$$u - \chi \leq -\delta < 0.$$

Also existiert ein $\varepsilon > 0$ mit

$$u - \chi \leq -\varepsilon |v| \leq \pm \varepsilon v \qquad (v \equiv 0 \text{ auf } \mathbb{R} \setminus S)$$

Dann ist $u \pm \varepsilon v \in K$ und

$$\begin{aligned} F(\pm\varepsilon v) &\leq a(u, \pm\varepsilon v) \\ \Rightarrow F(v) &= a(u,v) \quad \text{für alle } v \in C_c^1(I) \\ \Rightarrow F &= a(u,\cdot) \quad \text{in } H^{-1}(I) := \left(H_0^1(I)\right)^*. \end{aligned}$$

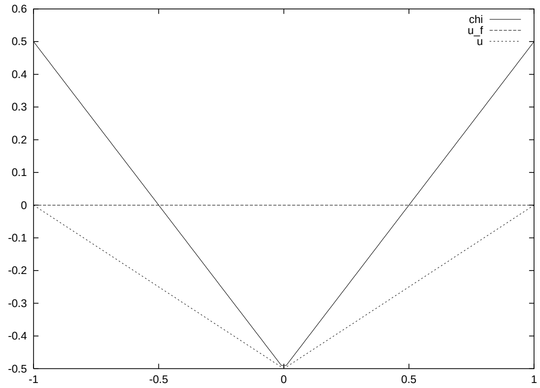
Nach dem Regularitätssatz 1.44 auf Seite 16 folgt $u \in H^2(I)$. Zudem ist $-u''|_I = f|_I$. □

Bemerkung 1.58. Für glattes χ ist tatsächlich $u \in H^2(\Omega)$. Dann bedeutet $\rho \geq 0$ punktweise, daß $f + u'' \geq 0$ fast überall, und es gelten die punktweisen Kuhn-Tucker-Bedingungen (Komplementaritätsbedingungen)

$$0 \leq \chi - u, \quad 0 \leq f + u'', \quad (\chi - u)(f + u'') \leq 0.$$

Beispiel. Deutung: χ beschreibt einen Stempel und u die heruntergedrückte Membran.

1. Für $\Omega = (-1,1)$, $V = H_0^1(\Omega)$, $f \equiv 0$, $u_f \equiv 0 \notin K$ und



$$\chi(x) = |x| - \frac{1}{2}$$

definiert

$$u(x) = \frac{1}{2}(|x| - 1)$$

eine Funktion mit

$$\begin{aligned} \{\chi = u\} &= \{0\}, \\ u'(x) &= \frac{1}{2}\operatorname{sign}(x). \end{aligned}$$

Für dieses $u \in K$ und jedes $v \in K$ gilt

$$\begin{aligned} a(u, v-u) &= -\frac{1}{2}\int_{-1}^{0}\left(v' + \frac{1}{2}\right)dx + \frac{1}{2}\int_{0}^{1}\left(v' - \frac{1}{2}\right)dx \\ &= -\frac{1}{2} - \frac{1}{2}v(0) + \frac{1}{2}\underbrace{v(-1)}_{=0} + \frac{1}{2}\underbrace{v(1)}_{=0} - \frac{1}{2}v(0) \\ &= -\underbrace{v(0)}_{\le -\frac{1}{2}} - \frac{1}{2} \ge 0 = F(v-u). \end{aligned}$$

Also ist u <u>die</u> Lösung. □

2. Für $\Omega = (0,1)$ und $\chi(x) = x^2 - lx + \frac{l^2}{8}$, $f \equiv 0$ bestimme man die Lösung u (und beweise, daß es eine ist).

1.11.1 P_1-Diskretisierung

Für den FER V_h wie in §1.10 und $\chi_h \in P_1(\mathcal{T}) \cap C^0[\Omega]$, z.B. $\chi_h(z) := \chi(z)$ für alle Knoten $z \in \mathcal{N}$, sei

$$K_h := \{v_h \in V_h \,|\, \forall z \in \mathcal{N}, v_h(z) \le \chi_h(z)\}.$$

Sprechweise: *Konforme FEM*, falls $K_h \subseteq K$, und sonst *nicht-konforme FEM*.

Das *diskrete Problem* ergibt sich durch Projektion auf K_h, also analog zum kontinuierlichen Fall. Sei weiterhin $K_h \neq \emptyset$ vorausgesetzt ($K_h = \emptyset$ ist möglich, sei hier aber ausgeschlossen).

Es existiert genau eine Lösung $u_h \in K_h$ mit

$$\begin{aligned} \|u_f - u_h\|_a &\le \|u_f - v_h\|_a \quad \text{für alle } v_h \in K_h; && (1.8a) \\ F(v_h - u_h) &\le a(u_h, v_h - u_h) \quad \text{für alle } v_h \in K_h; && (1.8b) \\ E(u_h) &= \min E(K_h). && (1.8c) \end{aligned}$$

Mit der Matrix $A_{jk} := a(\varphi_j, \varphi_k)$ und dem Vektor $b_j := F(\varphi_j)$ aus §1.10 führt (1.8b) zu

$$b \cdot (y - x) \le (Ax) \cdot (y - x)$$

für Koeffizienten $y_j := v_h(z_j) \le \chi(z_j) =: \chi_j$, $x_j := u_h(z_j)$. Das Problem (1.8c) heißt in der Optimierung QP oder quadratisches Problem:

$$\min_{\substack{y_j \le \chi(z_j) \\ j=1,\dots,n}} \frac{1}{2} y \cdot Ay - b \cdot y$$

Zur Lösung gibt es Standardsoftware in MATLAB, z.B. Verfahren von Wolfe oder Einzelschrittverfahren.

Satz 1.59. *Sofern $Iu \in K_h$ und $u_h \in K$ (also im wesentlichen im konformen Fall) folgt*

$$\|u - u_h\|_a \le a\left(u_f - Iu_f, u - Iu\right)^{\frac{1}{2}} \le \left\|\frac{h}{\pi}f\right\|_{L^2(\Omega)} = O\left(\|h\|_{L^\infty}\right)$$

Beweis. Für alle $v \in K$ und alle $v_h \in K_h$ ist

$$\begin{aligned} a(u, u-v) &\le F(u-v) = a(u_f, u-v), \\ a(u_h, u_h - v_h) &\le F(u_h - v_h) = a(u_f, u_h - v_h). \end{aligned}$$

und so

$$\begin{aligned} \|u-u_h\|_a^2 &= a(u_h, u_h - v_h) + a(u, u-v) + a(u_h, v_h - u) + a(u, v - u_h) \\ &\le a(u_f, u_h - v_h) + a(u_f, u - v) + a(u_h, v_h - u) + a(u, v - u_h) \\ &= F(u_h) - F(v_h) + F(u) - F(v) + a(u_h, v_h - u) + a(u, v - u_h) \\ &= a(u, v - u_h) - F(v - u_h) + a(u_h, v_h - u) - F(v_h - u). \end{aligned}$$

Also ist

$$\begin{aligned} \|u-u_h\|_a^2 &\le \underbrace{\min_{v \in K}\left(a(u, v-u_h) - F(v-u_h)\right)}_{\le 0, \text{ da } v = u_h \in K \text{ möglich}} \\ &\quad + \underbrace{\min_{v_h \in K_h}\left(a(u_h, v_h - u) - F(v_h - u)\right)}_{\le a(u_h, Iu-u) - F(Iu-u), \text{ denn } Iu \in K_h} \\ &\le 0 + a(u_f - u_h, u - Iu). \end{aligned}$$

Für die 1D FEM hier ist $u - Iu \perp_a V_h$ (Beweis als Aufgabe!). Dann ist $u_h - Iu_f \perp_a u - Iu$ und damit

$$a(u_f - u_h, u - Iu) = a(u_f - Iu_f, u - Iu).$$

Da nun auch $Iu_f \perp_a u - Iu$ gilt, erhalten wir

$$a(u_f - Iu_f, u - Iu) = a(u_f, u - Iu) = (f, u - Iu).$$

Die Anwendung einer Friedrichs-Ungleichung liefert

$$\begin{aligned} (f, u - Iu) &\le \left\|\frac{h}{\pi}f\right\|_{L^2(\Omega)} \left\|(u - Iu)'\right\|_{L^2(\Omega)} \\ &= \left\|\frac{h}{\pi}f\right\|_{L^2(\Omega)} \|u - Iu\|_a \\ &\le \left\|\frac{h}{\pi}f\right\|_{L^2(\Omega)} \|u - u_h\|_a, \end{aligned}$$

da $\|u - Iu\|_a = \min_{v_h \in V_h} \|u - v_h\|$. Benutzen wir nun die Young-Ungleichung[2] (mit $p = q = 2$), so erhalten wir weiter

$$\begin{aligned} (f, u - Iu) &\le \frac{1}{2}\left\|\frac{h}{\pi}f\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\|u - u_h\|_a^2 \\ &\le \frac{1}{2}\left\|\frac{h}{\pi}f\right\|_{L^2(\Omega)}^2 + \frac{1}{2}a(u_f - Iu_f, u - Iu). \end{aligned}$$

[2] Für $1 < p, q < \infty$, $1/p + 1/q = 1$, $a, b \in \mathbb{R}$ ist $ab \le a^p/p + b^q/q$.

Daraus folgt insgesamt

$$a\left(u_f - Iu_f, u - Iu\right) = (f, u - Iu) \leq \frac{1}{2}\left\|\frac{h}{\pi}f\right\|^2_{L^2(\Omega)} + \frac{1}{2}a\left(u_f - Iu_f, u - Iu\right),$$

d.h.

$$a\left(u_f - Iu_f, u - Iu\right) \leq \left\|\frac{h}{\pi}f\right\|^2_{L^2(\Omega)}. \qquad \square$$

Kapitel 2

Finite Elemente Methode

Kennenlernen von Algorithmen zur zweidimensionalen P_1-FEM.

2.1 Reguläre Triangulierung

Notation. Eine Menge von affin-unabhängigen Punkten $P_1, \ldots, P_{n+1} \in \mathbb{R}^n$ definieren ein nicht-entartetes *Simplex*

$$T := \operatorname{conv}\{P_1, \ldots, P_{n+1}\}$$

mit *Baryzentrischen Koordinaten* (Konvexkoeffizienten $\lambda_1(x), \ldots, \lambda_{n+1}(x)$), d.h. jeder Punkt $x \in \mathbb{R}^n$ besitzt <u>die</u> (eindeutige) Darstellung

$$\begin{aligned} x &= \lambda_1(x) P_1 + \ldots + \lambda_{n+1}(x) P_{n+1} \\ \text{mit } 1 &= \lambda_1(x) + \ldots + \lambda_{n-1}(x). \end{aligned}$$

Das heißt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \lambda_1(x) \begin{pmatrix} (P_1)_1 \\ \vdots \\ (P_1)_n \\ 1 \end{pmatrix} + \ldots + \lambda_{n+1}(x) \begin{pmatrix} (P_{n+1})_1 \\ \vdots \\ (P_{n+1})_n \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}.$$

Unter der Voraussetzung der linearen Unabhängigkeit von $P_2 - P_1, \ldots, P_{n+1} - P_1$ existiert eine Bijektion

$$\begin{aligned} \Lambda : \mathbb{R}^n &\to \operatorname{Bild}(\Lambda) = \left\{ \lambda \in \mathbb{R}^{n+1} \,\middle|\, \sum_{i=1}^{n+1} \lambda_i = 1 \right\} \subset \mathbb{R}^{n+1}, \\ x &\mapsto (\lambda_1(x), \ldots, \lambda_{n+1}(x)), \end{aligned}$$

nämlich die Lösung $\Lambda \in \mathbb{R}^{n+1}$ von

$$\begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix} \Lambda = \begin{bmatrix} x \\ 1 \end{bmatrix},$$

also

$$\lambda_j(x) = \Lambda_j(x) = \frac{\det \begin{bmatrix} x & P_1 & \cdots & P_{j-1} & P_{j+1} & \cdots & P_{n+1} \\ 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{bmatrix}}{\det \begin{bmatrix} P_j & P_1 & \cdots & P_{j-1} & P_{j+1} & \cdots & P_{n+1} \\ 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{bmatrix}}.$$

Die Koeffizientenmatrix $\begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix}$ ist regulär, denn $P_2 - P_1, \ldots, P_{n+1} - P_1$ sind linear unabhängig.

Konvention. Die Ecken $P_1, \ldots, P_{n+1}$ des Simplex

$$T := \{\lambda_1 P_1 + \ldots + \lambda_{n+1} P_{n+1} \,|\, \lambda_1, \ldots, \lambda_{n+1} \geq 0 \text{ mit } \lambda_1 + \ldots + \lambda_{n+1} = 1\}$$

seien so durchnummeriert, dass

$$\det\left(\begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix}\right) > 0$$

gilt. (Damit ist auch garantiert, daß keine Entartung auftreten kann.)

Im Fall $n = 2$ sind die Ecken dann gegen den Uhrzeigersinn zu nummerieren, um der Konvention zu genügen.

Bemerkung 2.1. Das n-dimensionale Volumen von T ist

$$|T| = \frac{1}{n!} \det\left(\begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix}\right) > 0.$$

Für nichtentartete Punkte $P_1, \ldots, P_{n+1} \in \mathbb{R}^n$ gilt

$$(\lambda_1(x), \ldots, \lambda_{n+1}(x)) = \Lambda(x) = \begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix}^{-1} \begin{pmatrix} x \\ 1 \end{pmatrix}$$

und für $j = 1, \ldots, n+1$ und $k = 1, \ldots, n$ ist

$$\lambda_{j,k}(x) := \frac{\partial \Lambda_j}{\partial x_k}(x) = \left(\begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix}^{-1}\right)_{j,k}.$$

Bemerkung 2.2. $\Lambda_j(x)$ ist die affine Funktion mit $\Lambda_j(P_k) = \delta_{j,k}$ für $j, k = 1, \ldots, n+1$.

Bemerkung 2.3. Für $n = 2$ und Zahlen $p, q, r \in \mathbb{N}_0$ ist

$$\int_T \lambda_1^p(x)\, \lambda_2^q(x)\, \lambda_3^r(x)\, dx = \frac{p!q!r!}{(p+q+r+2)!} 2\,|T|$$

(Gilt $\int_T \Lambda^\alpha(x)\, dx = \frac{\alpha! n!}{(|\alpha|+n)!} |T|$ für allgemeine Multiindizes $\alpha \in \mathbb{N}_0^{n+1}$?)

Notation. Ein *polygonal (polyhydral) berandetes Gebiet* Ω ist eine beschränkte, offene, zusammenhängende, nichtleere Teilmenge des $\mathbb{R}^n$, so daß sich $\bar{\Omega}$ als endliche Vereinigung von (nichtentarteten) Simplizes schreiben läßt.

Definition 2.4 (Triangulierung, Knoten, Kanten, Seitenflächen). Eine *Triangulierung* $\mathcal{T}$ eines polygonal berandeten Gebietes Ω ist eine endliche Menge

$$\mathcal{T} = \{T_1, \ldots, T_m\}$$

von nicht-entarteten Simplices

$$T_j = \operatorname{conv}\left\{P_1^j, \ldots, P_{n+1}^j\right\}$$

mit Ecken $P_1^j, \ldots, P_{n+1}^j \in \mathbb{R}^n$, so daß gilt:

$$\begin{aligned} \bigcup \mathcal{T} := T_1 \cup T_2 \cup \ldots \cup T_m &= \bar{\Omega} \\ \text{und } |T_j \cap T_k| &= 0 \qquad \text{für } j, k = 1, \ldots, m, j \neq k. \end{aligned}$$

Zu $\mathcal{T}$ bezeichne

$$\mathcal{N} := \{z \in \mathbb{R}^n \,| z \text{ Ecke eines } T \in \mathcal{T}\}$$

die Menge aller *Knoten* und

$$\mathcal{E} := \{E \subseteq \mathbb{R}^n \,| E \text{ Kante eines } T \in \mathcal{T}\}$$

die Menge aller *Kanten* (E ist Kante von $T = \operatorname{conv}\{P_1, \ldots, P_{n+1}\}$, wenn $E = \operatorname{conv}\{P_j, P_k\}$ für $j, k \in \{1, \ldots, n+1\}$ mit $j \neq k$.) und für $n \geq 3$ bezeichne

$$\mathcal{F} := \{F \subseteq \mathbb{R}^n \,| F \text{ Seitenfläche eines } T \in \mathcal{T}\}$$

die Menge aller *Seitenflächen* (F ist Seitenfläche von $T = \operatorname{conv}\{P_1, \ldots, P_{n+1}\}$, wenn $F = \operatorname{conv}\{P_j, P_k, P_l\}$ für $j, k, l \in \{1, \ldots, n+1\}$ mit j, k, l paarweise verschieden.) etc.

Definition 2.5 (reguläre Triangulierung). Eine Triangulierung $\mathcal{T}$ im $\mathbb{R}^n$ mit $n \leq 3$ heißt *regulär*, sofern für alle $T, K \in \mathcal{T}$ gilt: $T \cap K$ ist gemeinsamer Simplex, oder gemeinsame Fläche ($n \geq 3$) oder gemeinsame Kante ($n \geq 2$) oder gemeinsamer Knoten ($n \geq 1$) oder leer.

Notation. Zum Rand

$$\Gamma := \partial\Omega = \bar{\Omega} \setminus \Omega$$

mit

$$\Gamma = \Gamma_D \dot{\cup} \Gamma_N$$

($\dot{\cup}$ bezeichne die disjunkte Vereinigung) für Γ_D abgeschlossen und $\Gamma_N := \Gamma \setminus \Gamma_D$ relativ offen, bezeichne Γ_D den *Dirichlet-Rand* und Γ_N den *Neumann-Rand* von Ω. Für eine gegebene Triangulierung von Ω mit Kanten $\mathcal{E}$ ($n = 2$) bzw. Seitenflächen $\mathcal{F}$ ($n = 3$) ist dann

$$\begin{aligned}
\Gamma_D &= \begin{cases} \bigcup \mathcal{F}_D & \text{für } n = 3, \\ \bigcup \mathcal{E}_D & \text{für } n = 2, \end{cases} \\
\text{und } \overline{\Gamma_N} &= \begin{cases} \bigcup \mathcal{F}_N & \text{für } n = 3, \\ \bigcup \mathcal{E}_N & \text{für } n = 2, \end{cases}
\end{aligned}$$

für

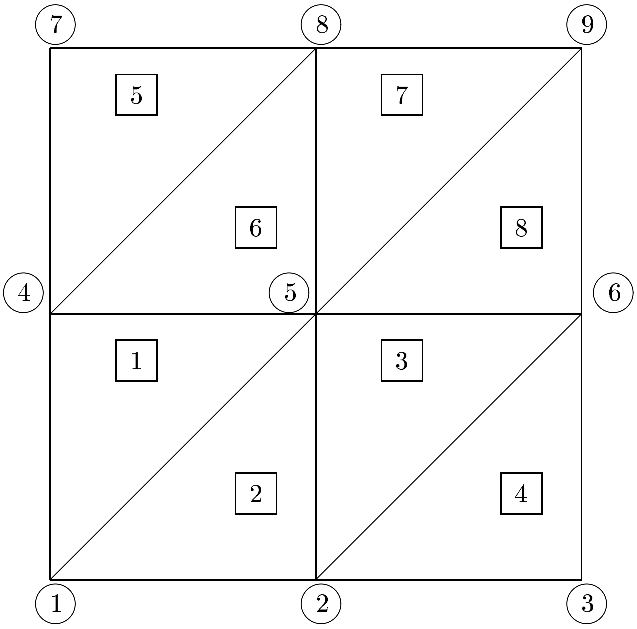
$$\begin{aligned}
\mathcal{E}_D &:= \{E \in \mathcal{E} \,| E \subseteq \Gamma_D\}, \\
\mathcal{E}_N &:= \{E \in \mathcal{E} \,\big| E \subseteq \overline{\Gamma_N}\}, \\
\mathcal{F}_D &:= \{F \in \mathcal{F} \,| F \subseteq \Gamma_D\}, \\
\mathcal{F}_N &:= \{F \in \mathcal{F} \,\big| F \subseteq \overline{\Gamma_N}\}.
\end{aligned}$$

($n = 1$: keine Extrabedingung)

Implementation. Die Knoten $\mathcal{N} = \{P_1, \ldots, P_\ell\}$ werden in einer Datei mit Namen `coordinates.dat` oder `coordinates4nodes.dat` gespeichert. In Zeile j stehen die n Komponenten von P_j, d.h. `coordinates(j,k)`$= (P_j)_k$, die k-te Komponente von P_j, $j = 1, \ldots, \ell$, $k = 1, \ldots, n$.

Die Elemente $\mathcal{T} = \{T_1, \ldots, T_m\}$ werden in einer Datei mit Namen `elements.dat` gespeichert. In Zeile $j \in \{1, \ldots, m\}$ und Spalte $k \in \{1, \ldots, n+1\}$ steht `elements(j,k)`, die Nummer des Knotens $P_k \in \mathcal{N}$ in einer fixierten Aufzählung $(P_1, \ldots, P_{n+1})$ der Ecken von T_j, welche der Konvention $\det\left(\begin{bmatrix} P_1 & \cdots & P_{n+1} \\ 1 & \cdots & 1 \end{bmatrix}\right) > 0$ genügt. Im Falle $n = 2$ bezeichnen die Knoten P_1, P_2 die *Referenzkante*.

Beispiel. Für $n = 2$ betrachte eine Triangulierung von $\Omega = (0,1)^2$ in $m = 8$ Elemente:

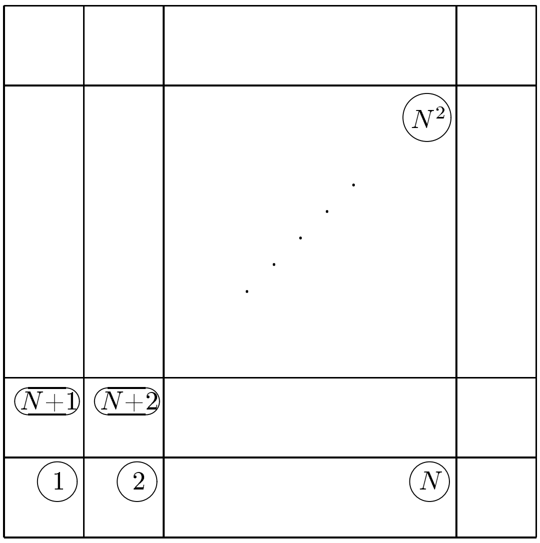


`coordinates.dat`:

1	0	0
2	.5	0
⋮		
6	1	.5
⋮		
9	1	1

`elements.dat`:

1	1	5	4
2	5	1	2
⋮			
6	8	4	5
⋮			
8	9	5	6



2.2 Finite Elemente Räume

Für eine reguläre Triangulierung $\mathcal{T}$ von Ω und $k \in \mathbb{N}_0$ bezeichne

$$\begin{aligned} P_k(\mathcal{T}) &:= \{f \in L^\infty(\Omega) \,|\, \forall T \in \mathcal{T}, f|_T \in P_k(T)\} \\ P_k(T) &:= \{f \in C^\infty(T) \,|\, f \text{ algebraisches Polynom vom totalen Grad} \leq k\} \end{aligned}$$

Bemerkung 2.6. Für $n \in \mathbb{N}$ und $k \in \mathbb{N}_0$ sei

$$P_k := \left\{ \sum_{\substack{\alpha \in \{0,\dots,k\}^n \\ |\alpha| \leq k}} a_\alpha x^\alpha \,\middle|\, a_\alpha \in \mathbb{R} \forall \alpha \right\}$$

der Vektorraum der Polynome von *totalem Grad* höchstens k und

$$Q_k := \left\{ \sum_{\alpha \in \{0,\ldots,k\}^n} a_\alpha x^\alpha \middle| a_\alpha \in \mathbb{R} \forall \alpha \right\}$$

der Vektorraum der Polynome von *partiellem Grad* höchstens k.

Beispiel. Für $n = 1$ ist

$$P_k = Q_k = \left\{ a_0 + a_1 x + \ldots + a_k x^k \,\middle|\, a_0, \ldots, a_k \in \mathbb{R} \right\}.$$

Für $n = 2$ ist

$$\begin{aligned} P_1 &= \operatorname{span}\{1, x, y\}, \\ Q_1 &= \operatorname{span}\{1, x, y, xy\} \\ \text{und } P_2 &= \operatorname{span}\left\{1, x, y, x^2, y^2, xy\right\}, \\ Q_2 &= \operatorname{span}\left\{1, x, y, xy, x^2, x^2 y, x^2 y^2, xy^2, y^2\right\}. \end{aligned}$$

Wir definieren nun den P_k-FE-Raum durch

$$V_h := P_k(\mathcal{T}) \cap C_0(\Omega), \qquad C_0(\Omega) := \left\{ f : \bar{\Omega} \to \mathbb{R} \text{ stetig} \,|\, f = 0 \text{ auf } \partial\Omega \right\}.$$

Der P_1-FER ist V_h für $k = 1$.

Für jeden *freien Knoten* z, d.h.

$$z \in \mathcal{K} := \mathcal{N} \cap \Omega,$$

bezeichne φ_z die Funktion $\varphi_z \in P_1(\mathcal{T})$ mit

$$\varphi_z(x) = \begin{cases} 1 & x = z \\ 0 & \text{sonst} \end{cases} \qquad \text{für alle } x \in \mathcal{N},$$

d.h. für alle $T \in \mathcal{T}$ ist $\varphi_z|_T(z) = 1$, falls $z \in T$, und $\varphi_z|_T(x) = 0$ für $x \in (\mathcal{N} \cap T) \setminus \{z\}$. Da $\mathcal{T}$ regulär ist, ist φ_z an allen Knoten stetig und damit global stetig auf $\bar{\Omega}$.

Lemma 2.7 (nodale Basis). *Für eine reguläre Triangulierung $\mathcal{T}$ ist $(\varphi_z)_{z \in \mathcal{K}}$ Basis von V_h für $k = 1$, $n = 2$ und heißt* nodale Basis.

Beweis. Die Eigenschaft $\varphi_z \in V_h$ folgt nach Konstruktion von V_h und der Stetigkeit von φ_z für innere Knoten $z \in \mathcal{K}$. Es bleibt zu zeigen, daß jede Funktion $f \in V_h$ als Linearkombination der $(\varphi_z)_{z \in \mathcal{K}}$ darstellbar ist.

Für $x \in \bar{\Omega}$ betrachte Dreiecke $T \in \mathcal{T}$ mit $x \in T$. Sei $x = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3$ mit den Ecken P_1, P_2, P_3 von T und den Baryzentrischen Koordinaten λ_j, $j = 1, 2, 3$. Wegen

$$\text{"}\{\varphi_z > 0\}\text{"} := \operatorname{Int}\left(\bigcup \{ T \in \mathcal{T} \,|\, z \in T \} \right) = \operatorname{Int}(\operatorname{supp} \varphi_z)$$

ist $\lambda_j(x) = \varphi_{P_j}(x)$ und

$$\begin{aligned} f(x) &= \lambda_1(x) f(P_1) + \lambda_2(x) f(P_2) + \lambda_3(x) f(P_3) \\ &= \varphi_{P_1}(x) f(P_1) + \varphi_{P_2}(x) f(P_2) + \varphi_{P_3}(x) f(P_3), \end{aligned}$$

also

$$f(x) = \sum_{z \in \mathcal{K}} f(z) \varphi_z(x). \qquad \square$$

Bemerkung 2.8. Allgemeines Prinzip der FEM: Reguläre Triangulierungen werden so definiert, daß sie nodale Basisfunktionen erlauben.

Für $E = \operatorname{conv}\{A, B\} \in \mathcal{E}$ definiere mittels

$$\varphi_E := 4\varphi_A\varphi_B$$

eine "*edge-bubble*". Dann ist eine (hierarchische) Basis vom P_2-FER gegeben durch

$$P_2(\mathcal{T}) \cap C(\bar{\Omega}) = \operatorname{span}\{\varphi_z \,|z \in \mathcal{N}\} \oplus \operatorname{span}\{\varphi_E \,|E \in \mathcal{E}\}$$

Weiterhin sei für $T = \operatorname{conv}\{A, B, C\} \in \mathcal{T}$

$$\varphi_T := 27\varphi_A\varphi_B\varphi_C$$

eine "*element-bubble*". Dann ist

$$P_3(\mathcal{T}) \cap C(\bar{\Omega}) \supsetneq \operatorname{span}\{\varphi_z \,|z \in \mathcal{N}\} \oplus \operatorname{span}\{\varphi_E \,|E \in \mathcal{E}\} \oplus \operatorname{span}\{\varphi_T \,|T \in \mathcal{T}\}$$

2.3 Diskretes Problem

Zum Modellproblem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{auf } \partial\Omega \end{aligned}$$

und einer gegebenen Triangulierung von Ω in lauter Dreiecke lautet das diskrete Problem: Man finde ein

$$u_h = \sum_{z \in \mathcal{K}} x_z\varphi_z \in V_h := P_1(\mathcal{T}) \cap C_0(\bar{\Omega})$$

($\mathcal{K} := \mathcal{N} \cap \Omega$ die freien Knoten, φ_z nodale Basisfunktionen) mit

$$a(u_h, v_h) = F(v_h) \qquad \text{für alle } v_h \in V_h,$$

wobei

$$\begin{aligned} a(u_h, v_h) &:= \int_\Omega \nabla u_h \cdot \nabla v_h dx, \\ F(v_h) &:= \int_\Omega f(x)\, v_h(x)\, dx. \end{aligned}$$

Dies ist äquivalent zu

$$a(u_h, \varphi_z) = \sum_{y \in \mathcal{K}} x_y a(\varphi_y, \varphi_z) = F(\varphi_z) \qquad \text{für alle } z \in \mathcal{K}.$$

Also $Ax = b$ mit

$$\begin{aligned} A &:= (a(\varphi_y, \varphi_z))_{y,z \in \mathcal{K}} \in \mathbb{R}^{N \times N}, \\ b &:= (b_z)_{z \in \mathcal{K}} := (F(\varphi_z))_{z \in \mathcal{K}} \in \mathbb{R}^N, \\ x &:= (x_z)_{z \in \mathcal{K}} \in \mathbb{R}^N \end{aligned}$$

für $N = \operatorname{card}(\mathcal{K})$.

Die Berechnung von A und b erfolgt elementweise: Für $\alpha, \beta \in \mathcal{N}$ ist

$$\mathrm{STEMA}_{\alpha,\beta} := a(\varphi_\alpha, \varphi_\beta) = \sum_{T \in \mathcal{T}} \int_T \nabla\varphi_\alpha \cdot \nabla\varphi_\beta dx.$$

Sofern $\alpha \notin T$ oder $\beta \notin T$ ist $\int_T \nabla\varphi_\alpha \cdot \nabla\varphi_\beta = 0$.

Für $T = \operatorname{conv}\{P_1, P_2, P_3\}$ mit $P_j = \alpha$ und $P_k = \beta$ ist

$$\int_T \nabla\varphi_\alpha \cdot \nabla\varphi_\beta dx = \left(\mathrm{STEMA}(T)_{j,k}\right)_{j,k \in \{1,2,3\}},$$

wobei

$$\mathrm{STEMA}(T)_{j,k} := \int_T \nabla\lambda_j \cdot \nabla\lambda_k dx$$

mit den Baryzentrischen Koordinaten $\lambda_1, \lambda_2, \lambda_3$ zu den Ecken P_1, P_2, P_3 von T.

α und β sind globale Bezeichner, j und k sind lokale Bezeichner.

Lemma 2.9. *Für*

$$\begin{aligned} Q &:= \begin{bmatrix} P_1 & P_2 & P_3 \\ 1 & 1 & 1 \end{bmatrix}, \\ R &:= Q^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

ist

$$\mathrm{STEMA}(T) = \frac{1}{2} \det(Q)\, RR^T.$$

Beweis.

$$\mathrm{STEMA}(T)_{j,k} = \int_T (\lambda_{j,1}\lambda_{k,1} + \lambda_{j,2}\lambda_{k,2})\, dx$$

mit

$$\lambda_{j,\ell} = \frac{\partial\lambda_j}{\partial x_\ell} = \left(\begin{bmatrix} P_1 & P_2 & P_3 \\ 1 & 1 & 1 \end{bmatrix}^{-1}\right)_{j,\ell}, \quad \ell = 1, 2.$$

Dann also $\left(Q^{-T} := \left(Q^{-1}\right)^T\right)$

$$\begin{aligned} \mathrm{STEMA}(T)_{j,k} &= |T| \left(\left(Q^{-1}\right)_{j,1} \left(Q^{-T}\right)_{1,k} + \left(Q^{-1}\right)_{j,2} \left(Q^{-T}\right)_{2,k}\right) \\ &= \frac{1}{2} \det(Q) \left(RR^T\right)_{j,k}. \end{aligned}$$

□

Implementation. In Matlab: Für eine 3×2 Matrix `vertices`:

`function stema=stema(vertices)`

$P = \begin{bmatrix} 1 & \cdots & 1 \\ & \text{vertices'} & \end{bmatrix}$ %3×3

$Q = P \backslash \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

`stema`$= \det(P) * Q * Q'/2$

Assemblieren von STEMA (ZUSBAU):

$(\text{STEMA}_{\alpha,\beta})_{\alpha,\beta\in\mathcal{N}}$ ist Anfangs Null und wird elementweise durch STEMA (T) additiv verändert

$$\text{STEMA} = \sum_{T\in\mathcal{T}} \sum_{j,k=1}^{3} \text{STEMA}\,(T)_{j,k} \underbrace{e_{\alpha(j,T)} \otimes e_{\alpha(k,T)}}_{\text{Boolsche Matrizen}}$$

wobei $e_z : \mathcal{N} \to \{0,1\}$ definiert vermöge

$$e_z\,(x) = \begin{cases} 1 & \text{falls } x = z, \\ 0 & \text{sonst} \end{cases}$$

und $\alpha\,(j,T)$ ist der (globale) Knoten der in T die lokale Nummer j hat, d.h. $\alpha\,(j,T) = P_j$.

Implementation. In Matlab:

```
STEMA(elements(j,:),elements(j,:))=...
STEMA(elements(j,:),elements(j,:))+STEMA(T)
```

mit T=`vertices=coordinates(elements(j,:),:)` bzw.
T=`coordinates4node(nodes4elements(j,:),:)`.

b wird elementweise mittels $b_j = b_j + \int_T \varphi_j dx$, wobei $\int_T \varphi_j dx = |T|\,\frac{1}{3!}$, berechnet.

Einbau der geometrischen Randbedingungen durch Reduzierung des Gleichungssystems

$$\text{STEMA} \cdot x = b$$

auf die freien Knoten $\mathcal{K}$:

$$x\,(\mathcal{K}) = \text{STEMA}\,(\mathcal{K},\mathcal{K})\setminus b\,(\mathcal{K})\,.$$

2.4 Zweidimensionale FEM für allgemeine elliptische partielle Differentialgleichungen 2. Ordnung

Starke Form: Gegeben seien Funktionen $A, A^{-1} \in L^\infty\left(\Omega;\mathbb{R}^{d\times d}\right)$, $b \in L^2\left(\Omega;\mathbb{R}^d\right)$, $\gamma \in L^\infty\,(\Omega;\mathbb{R})$ auf einem polygonal berandetem Gebiet $\Omega \Subset \mathbb{R}^d$, $d = 2$, mit Rand $\partial\Omega = \Gamma = \Gamma_D \dot\cup \Gamma_N$ und (äußerem) Einheitsnormalenvektor ν. Γ_D sei abgeschlossen und habe positives Oberflächenmaß. $\Gamma_N := \Gamma \setminus \Gamma_D$ sei relativ offen.

Zu den rechten Seiten $f \in L^2\,(\Omega)$, $u_D \in C^1\left(\bar\Omega\right)$ (oder größer) und $g \in L^2\,(\Gamma_N)$ finde man ein $u : \Omega \to \mathbb{R}$ mit hinreichenden Glattheitseigenschaften (z.B. $u \in C^2\,(\Omega) \cap C^1\left(\bar\Omega\right)$) und

$$\begin{aligned} -\operatorname{div}\,(ADu) + b\cdot Du + \gamma u &= f && \text{f.ü. in } \Omega, \\ u &= u_D && \text{f.ü. auf } \Gamma_D, \\ (ADu)\cdot\nu &= g && \text{f.ü. auf } \Gamma_N. \end{aligned}$$

Formale partielle Integration führt auf die schwache Form mit Bilinearform

$$\begin{aligned} a\,(u,v) &:= \int_\Omega (ADu)\cdot Dv dx + \int_\Omega (b\cdot Du + \gamma u)\, v dx. \\ F\,(v) &:= \int_\Omega f v dx + \int_{\Gamma_N} g v ds. \end{aligned}$$

Aus den Randtermen bei der partiellen Integration entsteht

$$-\int_{\partial\Omega} v\,\underbrace{(ADu)\cdot\nu}_{=g \text{ auf } \Gamma_N}\,ds = -\int_{\Gamma_N} gv ds - \int_{\Gamma_D} v\,(ADu)\cdot\nu ds$$

und für $v:\bar{\Omega}\to\mathbb{R}$ mit $Dv\in L^2\left(\Omega;\mathbb{R}^d\right)$ und $v=0$ auf Γ_D gilt

$$a(u,v)=F(v).$$

<u>Schwache Form</u>: Zu F finde man ein

$$u\in H^1(\Omega):=\left\{v\in L^2(\Omega)\,\middle|\,Dv \text{ ex. im schwachen Sinne und } Dv\in L^2\left(\Omega;\mathbb{R}^d\right)\right\},$$

mit $u=u_D$ auf Γ_D und für alle

$$v\in V:=H^1_D(\Omega):=\left\{w\in H^1(\Omega)\,|w=0 \text{ auf } \Gamma_D\right\}$$

gilt

$$a(u,v)=F(v).$$

<u>Definitheit von a</u>: Obwohl a i.a. ($b\not\equiv 0$, $A\not\equiv A^T$) keine symmetrische Bilinearform ist, kann man unter gewissen Bedingungen an A, b, γ zeigen, daß $a\geq 0$ gilt im Sinne von $0\leq a(v,v)$ für alle $v\in V=H^1_D(\Omega)$ ("Testfunktionen"). Dann definiert man $|||\cdot|||:=a(\cdot,\cdot)^{\frac{1}{2}}$ obwohl dies i.a. <u>keine</u> Norm ist und spricht von der "*Energie-Norm*".

Bemerkung 2.10. Die Forderung $u\in H^1(\Omega)$ mit $u=u_D$ auf Γ_D ist äquivalent als $u\in u_D+V$ notiert. Also ist V Vektorraum und u_D+V affiner Raum (Bisher $u_D\equiv 0$, also beide Räume gleich.). V heißt *Test(funktionen)raum* und u_D+V heißt *Ansatz(funktionen)raum*.

Entsprechend definiere

$$V_h:=\left\{v_h\in P_1(\mathcal{T})\cap C\left(\bar{\Omega}\right)|v_h=0 \text{ auf } \Gamma_D\right\}\subseteq V,$$

und $u_{Dh}\in P_1(\mathcal{T})\cap C\left(\bar{\Omega}\right)$ approximiert u_D (z.B. mittels nodaler Interpolation).

2.5 Netzverfeinerungen

Definition 2.11 (Referenzkante). In einer regulären Triangulierung $\mathcal{T}$ mit Kantenmenge $\mathcal{E}$ wird jedem Dreieck $T\in\mathcal{T}$ genau eine Kante $E(T)\in\mathcal{E}(T):=\{E\in\mathcal{E}\,|E\subset T\}$ zugeordnet. Diese Kante $E(T)$ heißt *Referenzkante* und definiert eine Abbildung $E:\mathcal{T}\to\mathcal{E}$.

Konvention. Im Datensatz `elemente` zu $\mathcal{T}$ bezeichnen die Einträge der ersten beiden Spalten die Endpunkte der jeweiligen Referenzkante.

Beispiel. Zu Beginn wird eine längste Kante als Referenzkante gewählt.

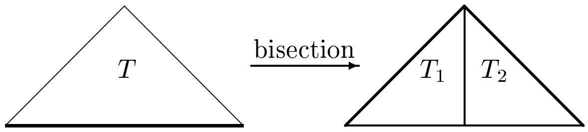
Basierend auf der Markierung durch Fixierung der Referenzkante betrachte man die folgenden Verfeinerungen von $\mathcal{T}$:

2.5.1 Bisections-Netzverfeinerung

Zu $\mathcal{M} \subseteq \mathcal{T}$ und Referenzkantenabbildung $E : \mathcal{T} \to \mathcal{E}$ bezeichne

$$\begin{aligned} \tilde{\mathcal{T}} &:= (\mathcal{T} \setminus \mathcal{M}) \cup \operatorname{bisec}(\mathcal{M}) \\ \text{mit } \operatorname{bisec}(\mathcal{M}) &:= \{K \,|\, K \in \operatorname{bisec}(T), T \in \mathcal{M}\} \end{aligned}$$

die (i.a. nicht reguläre) Triangulierung, die als Verfeinerung von $\mathcal{T}$ durch Bisection der Elemente in $\mathcal{M}$ entsteht:



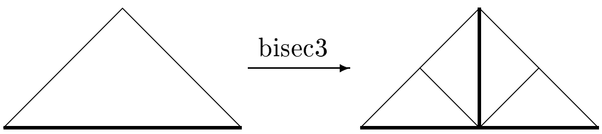
$\{T_1, T_2\} := \operatorname{bisec}(T)$. Dicke Linien deuten Referenzkanten an.

Bemerkung 2.12. Als hängende Knoten kommen hier nur Mittelpunkte von Kanten der Ausgangstriangulierung in Frage, d.h. eine Ecke von 2 Dreiecken ist Mittelpunkt einer Kante eines 3. Dreiecks.

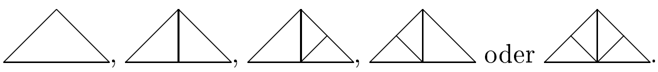
Algorithmus 2.13 (Bisection).
Input: Eine reguläre Triangulierung $\mathcal{T}$ mit Kantenmenge $\mathcal{E}$.
Eine Referenzkantenabbildung $E : \mathcal{T} \to \mathcal{E}$.
Eine Menge $\mathcal{M} \subseteq \mathcal{T}$ von Elementen.
Output: Eine Triangulierung $\hat{\mathcal{T}}$ mit Kantenmenge $\hat{\mathcal{E}}$ und Referenzkantenabbildung $E : \hat{\mathcal{T}} \to \hat{\mathcal{E}}$.
Compute:
$\ell := 0$
$\mathcal{T}_0 := (\mathcal{T} \setminus \mathcal{M}) \cup \operatorname{bisec}(\mathcal{M})$ mit Referenzkantenabbildung $E_0 : \mathcal{T}_0 \to \mathcal{E}_0$.
repeat
Berechne hängende Knoten $\mathcal{H}_\ell$ in $\mathcal{T}_\ell$.
Berechne $\mathcal{M}_\ell := \{T \in T_\ell \,|\, \exists z \in \mathcal{H}_\ell \, \exists E \in \mathcal{E}(T), \, z = \operatorname{mid}(E)\}$
$\mathcal{T}_{\ell+1} := (\mathcal{T}_\ell \setminus \mathcal{M}_\ell) \cup \operatorname{bisec}(\mathcal{M}_\ell)$
$\ell := \ell + 1$
until $\mathcal{H}_\ell = \emptyset$
$\hat{\mathcal{T}} := \mathcal{T}_\ell$

Satz 2.14. *Der Algorithmus bricht nach endlich vielen Schritten ab und berechnet eine reguläre Triangulierung $\hat{\mathcal{T}}$.*

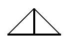
Beweis. In jedem Schritt des Algorithmus ist $\mathcal{T}_\ell$ eine Vergröberung der regulären Triangulierung $\operatorname{bisec3}(\mathcal{T})$, bei der jedes Dreieck T in $\mathcal{T}$ in, wie nachfolgend gezeigt, 4 Teildreiecke zerlegt wird:

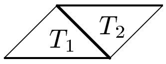


D.h. für alle $T \in \mathcal{T}$ besteht $\mathcal{T}_\ell|_T := \{K \in \mathcal{T}_\ell \,|\, K \subseteq T\}$ aus

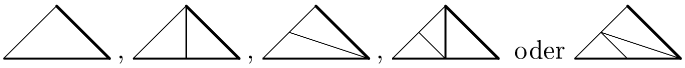


Beweis dieser Aussage durch vollständige Induktion über ℓ:

$\ell = 0$: Nach Konstruktion besteht $\mathcal{T}_0$ nur aus Dreiecken  und .

$\ell \to \ell + 1$: Betrachte die Situation 

mit der Referenzkante $F = T_1 \cap T_2$, für die $\operatorname{mid}(F) \in \mathcal{H}_\ell$ gilt. Also ist $T_1, T_2 \in \mathcal{M}_\ell$ und Bisection von T_1 ergibt



d.h. $\mathcal{T}_{\ell+1}$ ist auch eine Vergröberung von bisec3 $(\mathcal{T})$.

Also gilt

$$\operatorname{card}(\mathcal{T}) \leq \operatorname{card}(\mathcal{T}_\ell) \leq \operatorname{card}(\operatorname{bisec3}(\mathcal{T})) = 4 \operatorname{card}(\mathcal{T})$$

und $\operatorname{card}(\mathcal{T}_\ell) < \operatorname{card}(\mathcal{T}_{\ell+1})$ oder $\operatorname{card}(\mathcal{T}_\ell) = \operatorname{card}(\mathcal{T}_{\ell+1})$ mit Abbruch.

Also bricht der Algorithmus nach maximal $4 \operatorname{card}(\mathcal{T})$ Schritten ab. Da dann keine hängenden Knoten vorliegen, ist $\hat{\mathcal{T}}$ regulär. □

Satz 2.15. *Für eine reguläre Triangulierung $\mathcal{T}$ mit Referenzkantenabbildung $E : \mathcal{T} \to \mathcal{E}$, eine Menge $\mathcal{M} \subseteq \mathcal{T}$ von Elementen und eine reguläre Triangulierung $\mathcal{T}'$, welche Verfeinerung von $(\mathcal{T} \setminus \mathcal{M}) \cup \operatorname{bisec}(\mathcal{M})$ ist, gilt* $\operatorname{card}\left(\hat{\mathcal{T}}\right) \leq \operatorname{card}(\mathcal{T}')$ *für den Output $\hat{\mathcal{T}}$ aus dem Algorithmus.*

Beweis. Übungsaufgabe. □

Satz 2.16 (Binev-Dahmen-DeVore (2004)). *Zu einer Triangulierung $\mathcal{T}$ mit Referenzkantenabbildung $E : \mathcal{T} \to \mathcal{E}$ und Mengen $\mathcal{M}_0, \mathcal{M}_1, \ldots, \mathcal{M}_{L-1}$ mit $\mathcal{M}_0 \subseteq \mathcal{T}_0 := \mathcal{T}$, $\mathcal{M}_1 \subseteq \mathcal{T}_1$, $\ldots$, $\mathcal{M}_{L-1} \subseteq \mathcal{T}_{L-1}$ sei $\mathcal{T}_{\ell+1}$ das Ergebnis des Bisectionsalgorithmus mit Input $\mathcal{T}_\ell$, $E_\ell : \mathcal{T}_\ell \to \mathcal{E}_\ell$ und $\mathcal{M}_\ell$, $\ell = 0, 1, \ldots, L-1$.*

Unter gewissen Bedingungen an $E : \mathcal{T}_0 \to \mathcal{E}_0$ gilt

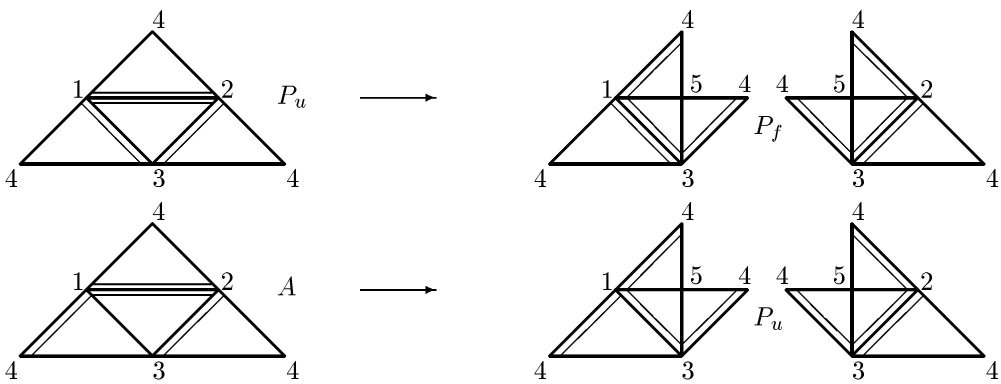
$$\operatorname{card}(\mathcal{T}_L \setminus \mathcal{T}_0) \leq C \sum_{\ell=0}^{L-1} \operatorname{card}(\mathcal{M}_\ell)$$

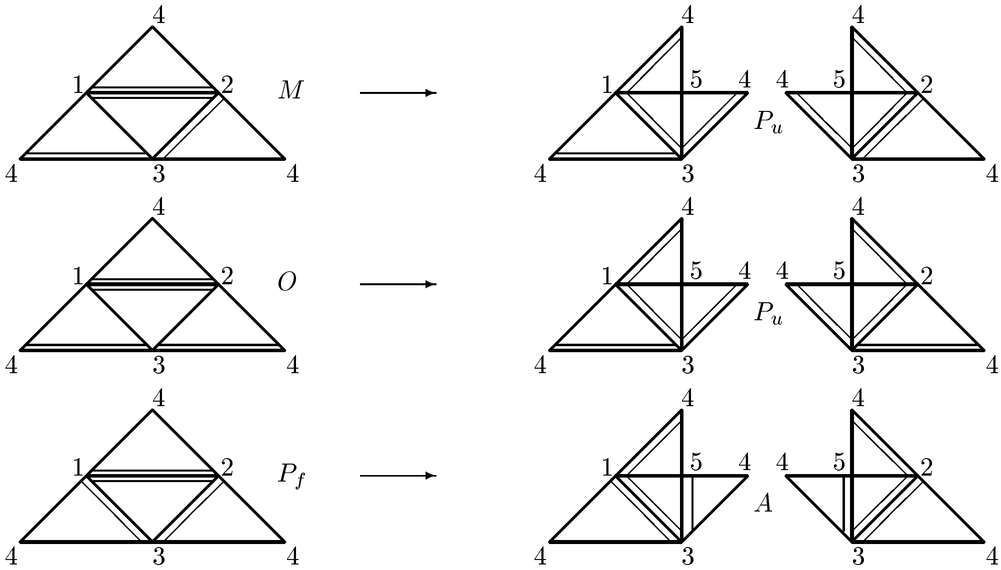
mit einer Konstanten C, die nur von $\mathcal{T}_0$ und $E_0 : \mathcal{T}_0 \to \mathcal{E}_0$ abhängt und insbesondere unabhängig von L ist.

Beweis. Später für allgemeinere Verfeinerungen und Details an die Anfangsbedingungen. □

Bemerkung 2.17 (Bisektion von Tetraedern). Auch dreidimensionale Bisection ist möglich, vgl. Literatur, z.B. Bänsch: E. Local Mesh Refinement in 2 and 3 Dimensions. *IMPACT of Computing in Science and Engineering* 1991; 3: 181–191.

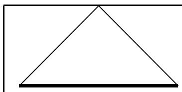
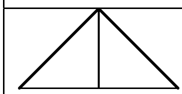
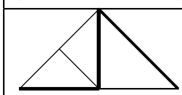
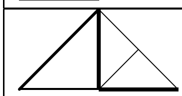
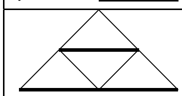
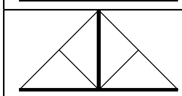
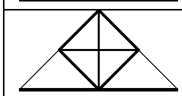
Sie beinhaltet 5 Typen von Tetraedern (bezeichnet als P_u, A, M, O und P_f auf der linken Seite) und ihre Bisektionen die mit $M \to 2\,P_u$ und $O \to 2\,P_u$ beginnen und dann von einer sukzessiven Schleife der Form $A \to 2\,P_u \to 4\,P_f \to 8\,A$ etc. fortgesetzt werden. Beachte, dass die Formen von P_u und P_f identisch sind, aber eine unterschiedliche Rolle in der Verfeinerungsschleife spielen.





2.5.2 Erweiterte Rot-Grün-Blau-Netzverfeinerung

Dieser Abschnitt behandelt die nachfolgend skizzierten Verfeinerungsregeln, die auch die neuen Referenzkanten (durch dicke Linien) beschreiben.

	keine
	bisec2=bisec="grün"
	bisec3ℓ="blau$_\ell$"
	bisec3r="blau$_r$"
	rot
	bisec3
	bisec5 (mit neuem Innenknoten)

Alle diese Verfeinerungen entstehen durch Markierung von Kanten und deren Bisection und sind Vergröberungen von bisec5 (T).

Übergang von Element- zu Kanten-Markierungen

In einer Triangulierung $\mathcal{T}$ werden gewisse Kanten $\mathcal{M} \subseteq \mathcal{E}$ zur Bisection markiert.

Definition 2.18 ((R)). Eine Menge $\mathcal{M} \subseteq \mathcal{E}$ markierter Kanten in einer regulären Triangulierung $\mathcal{T}$ (in lauter Dreiecke in 2D) und eine Referenzkantenabbildung $E : \mathcal{T} \to \mathcal{E}$ erfüllt (R), gdw.

$$\forall T \in \mathcal{T} \qquad (\mathcal{M} \cap \mathcal{E}(T) = \emptyset \text{ oder } E(T) \in \mathcal{M}), \tag{R}$$

$\mathcal{E}(T) := \{E \in \mathcal{E} \,|\, E \subset \partial T\}$.

Bemerkung 2.19. Sofern (R) gilt, kann jedes $T \in \mathcal{T}$ mit den obigen Regeln so verfeinert werden, daß alle Kanten in $\mathcal{M} \cap \mathcal{E}(T)$ halbiert werden. Jede dieser so entstandenen Verfeinerungen ist eine *reguläre* Triangulierung!

(Beweis: Es werden *keine* hängenden Knoten erzeugt.)

Idee: (R) herstellen durch sukzessive Vergrößerung von $\mathcal{M}$:

Algorithmus 2.20 (Rot-Grün-Blau-Abschluß).
Input: Eine reguläre Triangulierung $\mathcal{T}$ mit Kantenmenge $\mathcal{E}$.
Eine Referenzkantenabbildung $E : \mathcal{T} \to \mathcal{E}$.
Eine Menge $\mathcal{M} \subseteq \mathcal{E}$ markierter Kanten.
Compute:
$k := 0$, $\mathcal{M}^{(0)} := \mathcal{M}$, $\mathcal{N}^{(0)} := \mathcal{M}$.
while $\mathcal{N}^{(k)} \neq \emptyset$ **do**
Wähle $E \in \mathcal{N}^{(k)}$ mit Nachbarelementen $T_\pm \in \mathcal{T}$, d.h. $E \subset \partial T_\pm$ (bzw. nur ein Nachbarelement T_+, falls E Außenkante ist).
$\mathcal{M}^{(k+1)} := \mathcal{M}^{(k)} \cup \{E_+, E_-\}$ (bzw. $\mathcal{M}^{(k+1)} := \mathcal{M}^{(k)} \cup \{E_+\}$, falls E Außenkante ist) für die Referenzkante $E_\pm := E(T_\pm)$ von $T_\pm$.

$$\mathcal{N}^{(k+1)} := \begin{cases} \mathcal{N}^{(k)} \setminus \{E\} & \text{falls } \mathcal{M}^{(k)} = \mathcal{M}^{(k+1)}, \\ \left(\mathcal{N}^{(k)} \cup \{E_+, E_-\}\right) \setminus \{E\} & \text{falls } \mathcal{M}^{(k)} \neq \mathcal{M}^{(k+1)}. \end{cases}$$

$k := k + 1$.
end while
Output: $\hat{\mathcal{M}} := \mathcal{M}^{(k)}$.

Satz 2.21. *(a) Der Algorithmus bricht nach endlich vielen Schritten ab.*

(b) $\hat{\mathcal{M}}$ *erfüllt* (R).

(c) $\hat{\mathcal{M}}$ *ist minimal mit* (R) *und* $\mathcal{M} \subseteq \hat{\mathcal{M}}$, *d.h. für jedes* $\tilde{\mathcal{M}}$ *mit* (R) *und* $\mathcal{M} \subseteq \tilde{\mathcal{M}}$ *gilt* $\hat{\mathcal{M}} \subseteq \tilde{\mathcal{M}}$.

Beweis. (a): Da $m_j := \operatorname{card}(\mathcal{M}_j)$ monoton wächst, aber $m_j \leq \operatorname{card}(\mathcal{E})$, gilt

$$\mathcal{M}_J = \mathcal{M}_{J+1} = \mathcal{M}_{J+2} = \dots$$

ab einem Index J, oder der Algorithmus bricht ab.

Im ersten Fall erfüllt $n_J := \operatorname{card}\left(\mathcal{N}^{(J)}\right)$ aber

$$n_J > n_{J+1} > n_{J+2} > \dots > 0.$$

Dieser Widerspruch zeigt $n_k = 0$ für ein $k < \infty$.

(b): 1. Behauptung:

$$\hat{\mathcal{M}} \subseteq \hat{\mathcal{N}} := \bigcup_{m=0}^{k} \mathcal{N}^{(m)}.$$

Beweis: Für jedes $E \in \hat{\mathcal{M}}$ gilt $E \in \mathcal{M}^{(\ell)} \setminus \mathcal{M}^{(\ell-1)}$ für ein $\ell \in \{0, 1, \dots, k\}$, wobei $\mathcal{M}^{(-1)} := \emptyset$.
Für $\ell = 0$ ist $E \in \mathcal{N}^{(0)} = \mathcal{M} \subseteq \hat{\mathcal{N}}$.

Für $\ell \geq 1$ gilt $E = E(K_+)$ für ein $K_+ \in \mathcal{T}$ mit $K_+ \cap K_- = F \in \mathcal{N}^{(\ell-1)} \setminus \mathcal{N}^{(\ell)}$ (und Modifikation für $F \subseteq \partial\Omega$).

Da $E \notin \mathcal{M}^{(\ell-1)}$ ist $\mathcal{M}^{(\ell)} \neq \mathcal{M}^{(\ell-1)}$ und daher $\mathcal{N}^{(\ell)} = \left(\mathcal{N}^{(\ell-1)} \cup \{E(K_+), E(K_-)\}\right) \setminus \{F\}$.

Falls $E = F$, so ist $E \in \hat{\mathcal{N}}$, da $F \in \mathcal{N}^{(\ell-1)} \subseteq \hat{\mathcal{N}}$. Falls $E \neq F$, so ist $E \in \mathcal{N}^{(\ell)} \subseteq \hat{\mathcal{N}}$. Also stets $E \in \hat{\mathcal{N}}$.

Nebenbemerkung:

$$\hat{\mathcal{N}} \subseteq \hat{\mathcal{M}}.$$

Beweis: Zu $E \in \hat{\mathcal{N}}$ existiert ein j mit $E \in \mathcal{N}^{(j)} \setminus \mathcal{N}^{(j+1)}$. $E = T_+ \cap T_-$, $E(T_\pm) \in \hat{\mathcal{M}}$ (und Modifikation für $E \subseteq \partial\Omega$).

Wenn $E \in \{E(T_+), E(T_-)\}$, so ist $E \in \hat{\mathcal{M}}$. Andernfalls ist E keine Referenzkante. Da in $\hat{\mathcal{N}} \setminus \mathcal{N}^{(0)}$ nur Referenzkanten liegen, ist $E \notin \hat{\mathcal{N}} \setminus \mathcal{N}^{(0)}$. Da $E \in \hat{\mathcal{N}}$, folgt $E \in \mathcal{N}^{(0)} \subseteq \hat{\mathcal{M}}$.

2. Behauptung:

$$\left\{ E(K) \,\middle|\, \exists K \in \mathcal{T} \quad \hat{\mathcal{N}} \cap \mathcal{E}(K) \neq \emptyset \right\} \subseteq \hat{\mathcal{M}}.$$

Beweis: Betrachte ein $K \in \mathcal{T}$ mit $\hat{\mathcal{N}} \cap \mathcal{E}(K) \neq 0$ und ein $E \in \hat{\mathcal{N}} \cap \mathcal{E}(K)$. Dann existiert ein Index $j \in \{0, 1, \ldots, k\}$, so daß $E \in \mathcal{N}^{(j)} \setminus \mathcal{N}^{(j+1)}$, also $E = T_+ \cap T_-$ und $E(T_+), E(T_-) \in \hat{\mathcal{M}}$, $T_+, T_- \in \mathcal{T}$ (und Modifikation für $E \in \partial\Omega$).

Da $E \in \mathcal{E}(K)$, ist $K = T_+$ oder $K = T_-$ und somit $E(K) \in \hat{\mathcal{M}}$.

3. Behauptung: $\hat{\mathcal{M}}$ erfüllt (R).

Beweis: Wenn $T \in \mathcal{T}$ mit $\hat{\mathcal{M}} \cap \mathcal{E}(T) \neq \emptyset$, existiert nach 1. eine Kante $E \in \mathcal{E}(T) \cap \hat{\mathcal{N}} \neq \emptyset$. Nach 2. ist dann $E(T) \in \hat{\mathcal{M}}$.

(c): Für $\tilde{\mathcal{M}}$ mit (R) und $\mathcal{M}^{(0)} = \mathcal{M} \subseteq \tilde{\mathcal{M}} \subseteq \mathcal{E}$ wird $\mathcal{M}^{(j)} \subseteq \tilde{\mathcal{M}}$ durch vollständige Induktion über $j = 0, 1, 2, \ldots, k$ gezeigt.

$j = 0$: klar.

$j \to j+1$: $j \leq k-1$, d.h. zu zeigen ist $E(K) \in \tilde{\mathcal{M}}$ für $K \in \mathcal{T}$ mit $E(K) \in \mathcal{M}^{(j+1)} \setminus \mathcal{M}^{(j)}$ (sonst $E(K) \in \mathcal{M}^{(j)} \subseteq \tilde{\mathcal{M}}$ nach Induktionsvoraussetzung). Es gilt $E \in \mathcal{N}^{(j)} \cap \mathcal{E}(K)$, $\mathcal{M}^{(j+1)} \neq \mathcal{M}^{(j)}$, $\mathcal{N}^{(j+1)} = \left(\mathcal{N}^{(j)} \cup \{E_+, E_-\}\right) \setminus \{E\}$.

o.B.d.A. $E_+ = E(K)$ und $E \in \mathcal{E}(K)$.

Im Fall $E \in \mathcal{M}^{(0)} \subseteq \tilde{\mathcal{M}}$ ist $\tilde{\mathcal{M}} \cap E(K) \neq \emptyset$, also $E(K) \in \tilde{\mathcal{M}}$ wegen (R).

Sei nun also $E \notin \mathcal{M}^{(0)} = \mathcal{N}^{(0)}$. Dann existiert ein $\ell \in \{0, 1, \ldots, j\}$, $T \in \mathcal{T}$ mit $E = E(T)$ und $E \in \left(\mathcal{N}^{(\ell)} \setminus \mathcal{N}^{(\ell-1)}\right) \cap \mathcal{E}(T)$.

Wegen $E \in \mathcal{M}^{(\ell)} \subseteq \mathcal{M}^{(j)} \subseteq \tilde{\mathcal{M}}$ nach Induktionsvoraussetzung folgt $E \in \tilde{\mathcal{M}} \cap \mathcal{E}(K) \neq \emptyset$. Mit (R) dann $E(K) \in \tilde{\mathcal{M}}$. $\square$

Algorithmus 2.22 (Adaptive Netzverfeinerung).

Input: Eine reguläre Triangulierung $\mathcal{T}^{(0)}$.
Eine Referenzkantenabbildung $E^{(0)} : \mathcal{T}^{(0)} \to \mathcal{E}^{(0)}$.
Mengen $\mathcal{M}^{(0)} \subseteq \mathcal{E}^{(0)}$, $\mathcal{M}^{(1)} \subseteq \mathcal{E}^{(1)}$, …, $\mathcal{M}^{(L)} \subseteq \mathcal{E}^{(L)}$

Compute:
for $\ell = 1, \ldots, L-1$ **do**
$\hat{\mathcal{M}}^{(\ell)}$:=Abschluß von $\mathcal{M}^{(\ell)}$ in regulärer Triangulierung $\mathcal{T}^{(\ell)}$ mit Algorithmus 2.20
Verallgemeinere die Rot-Grün-Blau-Netzverfeinerung mit Bisektion aller Kanten in $\hat{\mathcal{M}}^{(\ell)}$ mit Output $\mathcal{T}^{(\ell+1)}$ und $E^{(\ell+1)} : \mathcal{T}^{(\ell+1)} \to \mathcal{E}^{(\ell+1)}$
end for

Output: Reguläre Triangulierungen $\mathcal{T}^{(0)}, \mathcal{T}^{(1)}, \ldots, \mathcal{T}^{(L)}$

Notation. $\mathcal{T}^{(0)}$ heißt *grobe Triangulierung* oder *Anfangstriangulierung*. Idee: $\mathcal{T}^{(0)}$ hat möglichst wenige Elemente, so daß $\bar{\Omega} = \cup\mathcal{T}^{(0)}$ beschrieben wird.

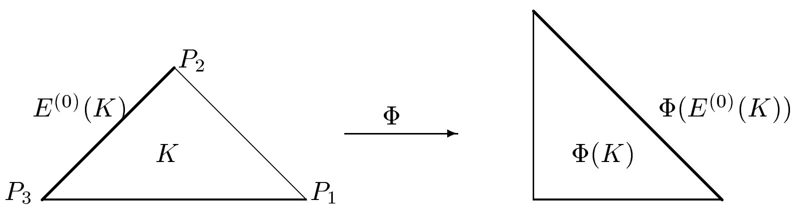
Die Elemente in $\mathcal{T}^{(0)}$ heißen *grobe Elemente* oder *Makroelemente*. (Diese Bezeichnungen sind <u>nicht</u> standardisiert.)

Satz 2.23. *Für eine reguläre Triangulierung $\mathcal{T}^{(0)}$ (und geeignete Referenzkantenabbildung $E^{(0)}$: $\mathcal{T}^{(0)} \to \mathcal{E}^{(0)}$) und geeignete Mengen $\mathcal{M}^{(0)}, \ldots, \mathcal{M}^{(L-1)}$ markierter Kanten sei $\mathcal{T} := \mathcal{T}^{(k)}$ eine der durch den Algorithmus 2.22 definierte Triangulierung. Dann gelten (a)-(e):*

(a) $\mathcal{T}$ ist reguläre Triangulierung.

(b) Für alle $T \in \mathcal{T}$ existiert genau ein $K \in \mathcal{T}^{(0)}$ mit $T \subseteq K$.

Für alle $K \in \mathcal{T}^{(0)}$ existiert genau eine Affinität (d.h. affine Bijektion) $\Phi := \Phi_K : K \to T_{\mathrm{ref}}$, $T_{\mathrm{ref}} := \mathrm{conv}\left\{(0,0),(1,0),(0,1)\right\}$, $\Phi(x) = Ax + b$, $A \in \mathbb{R}^{2\times 2}$ regulär, wobei $\Phi\left(E^{(0)}(K)\right) = \mathrm{conv}\left((1,0),(0,1)\right)$. $\Phi^{-1}(x) := P_1 + (P_2 - P_1, P_3 - P_1)\,x$, $\det D\Phi > 0$.



Dann ist das affine Bild $\{\Phi(T)\,|T \in \mathcal{T}, T \subseteq K\}$ von $\mathcal{T}|_K := \{T \in \mathcal{T}\,|T \subseteq K\}$ eine reguläre Triangulierung von T_{ref} in lauter gleichschenklige rechtwinklige Dreiecke.

(c) Die Referenzkante $E(T)$ wird für $T \in \mathcal{T}|_K$, $K \in \mathcal{T}^{(0)}$, durch Φ_K auf die Hypotenuse von $\Phi_K(T)$ abgebildet.

(d) Die L^2-Projektion $\Pi : H^1(\Omega) \to L^2(\Omega)$ auf den FER $P_1(\mathcal{T}) \cap C(\bar{\Omega})$ ist H^1-stabil in dem Sinne, daß ein $C := C\left(\mathcal{T}^{(0)}, E^{(0)}\right)$ existiert, so daß für alle $v \in H^1(\Omega)$

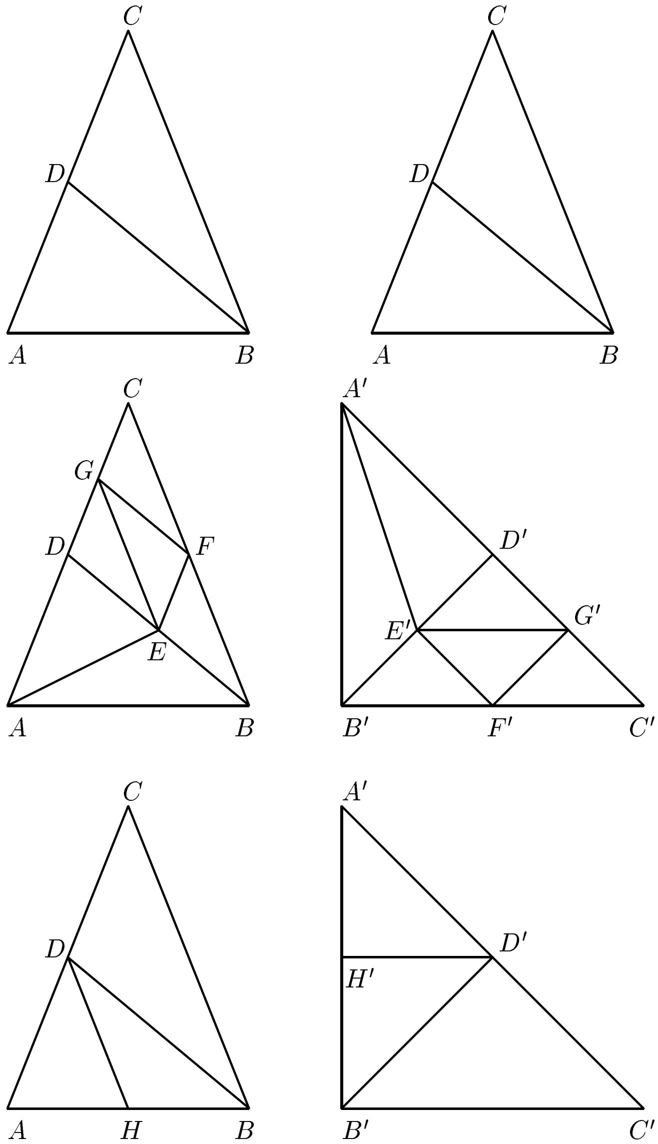
$$\|D\Pi v\|_{L^2(\Omega)} \leq C\,\|Dv\|_{L^2(\Omega)}\,.$$

(e) Es existiert ein $C := C\left(\mathcal{T}^{(0)}, E^{(0)}\right)$, so daß

$$\mathrm{card}\left(\mathcal{T} \setminus \mathcal{T}^{(0)}\right) \leq C \sum_{\ell=0}^{k-1} \mathrm{card}\left(\mathcal{M}^{(\ell)}\right).$$

Dabei ist C von L, k, etc. unabhängig.

Bemerkung 2.24 (Vergleich zur longest-edge-bisection). Statt Referenzkanten zu definieren, hat M.C. Rivara (1984) vorgeschlagen, eine längste Kante als Bezugskante für rot-grün-blau-Verfeinerung zu benutzen. Dies ist nicht dasselbe wie die hier betrachtete verallgemeinerte rot-grün-blau-Verfeinerung (Algorithmus 2.20).



Eine grün-Verfeinerung des Dreiecks (A, B, C) (oben), gefolgt von einer grün-Verfeinerung von (A, B, D) in Bezug auf die Referenzkante $\overline{BD}$ (Mitte) oder $\overline{AB}$ (unten). Die rechte Spalte zeigt das affine Bild auf dem Referenzdreieck (A', B', C'). Der Algorithmus 2.22 vermeidet die mittlere Konfiguration durch $E\,(A, B, D) = \overline{AB}$.

Bemerkung 2.25. Als Folgerung von Satz 2.23c treten in Dreiecken von $\mathcal{T}$ nur endlich viele Winkel auf, unabhängig von k, $\operatorname{card}(\mathcal{T})$, etc. und nur abhängig von $\mathcal{T}^{(0)}, \mathcal{E}^{(0)}$.

Z.B. ist

$$h_T/\rho_T \leq C\left(\mathcal{T}^{(0)}, \mathcal{E}^{(0)}\right)$$

für $T \in \mathcal{T}$ mit $h_T = \operatorname{diam}(T) = 2\cdot$Außenkreisradius und ρ_T :=Innenkreisradius.

Man sagt: $\mathcal{T}$ ist *lokal quasi-uniform.* (Denn benachbarte Elemente $T, K \in \mathcal{T}$ mit $T \cap K \neq \emptyset$ erfüllen $\frac{h_T}{\rho_K} + \frac{h_K}{\rho_T} \lesssim 1$ und $A \lesssim B$ bezeichnet $A \leq C\left(\mathcal{T}^{(0)}, \mathcal{E}^{(0)}\right) B$).

Beweisskizzen: (a): Das Konzept der Kantenmarkierungen vermeidet automatisch das Auftreten hängender Knoten.

(b),(c): Die Konstruktion von Φ_K ist offensichtlich und sofern $\mathcal{T}^{(\ell)}|_K$ und $\mathcal{E}^{(\ell)}|_K$ durch Φ_K auf eine Triangulierung von T_{ref} und Referenzkantenabbildung abgebildet wird, vererben sich die Referenzkanten wie in $\left\{\Phi_K(T) \,\middle|\, T \in \mathcal{T}^{(\ell+1)}|_K\right\}$ durch Wahl der längsten Kanten.

(d): Die wesentlichen Beweisideen sind nachfolgend skizziert.

Definition 2.26 (Knotenabstand). Zu $a, b \in \mathcal{N}$:=Menge der Knoten in $\mathcal{T} = \mathcal{T}^{(k)}$ definiere $\delta(a,b)$ als die kleinste ganze Zahl J von Elementen $T_1, \ldots, T_J \in \mathcal{T}$ mit $a \in T_1$, $T_1 \cap T_2 \neq \emptyset$, $T_2 \cap T_3 \neq \emptyset$, …, $T_{J-1} \cap T_J \neq 0$, $b \in T_J$, falls $a \neq b$, und $\delta(z,z) := 0$ für $z \in \mathcal{N}$.

Bemerkung. $(\mathcal{N}, \delta)$ ist ein metrischer Raum.

Definition 2.27 (Verfeinerungslevel). Zu $T \in \mathcal{T}$ definiere $\ell(T) \geq 0$ vermöge $K \in \mathcal{T}^{(0)}$, $T \subseteq K$ und

$$\ell(T) := \log_2 \sqrt{\frac{|K|}{|T|}}.$$

Bemerkung. $\ell(T) \in \{0, 1/2, 1, 3/2, 2, 5/2, \ldots\} =: 1/2\mathbb{N}_0$.

Proposition 2.28. *Für $a, b \in \mathcal{N} \setminus \mathcal{N}^{(0)}$ mit $\delta(a,b) \leq 1$, $a \in T$, $b \in K$, $T, K \in \mathcal{T}$ ist*

$$|\ell(T) - \ell(K)| \leq 3.$$

Beweisidee. Patche im affinen Bild ansehen: Betrachte a, b im Innern eines Makroelementes und analysiere 2 benachbarte Patche. Vorsicht bei Randpatchen. $\square$

Definition 2.29. Für $z \in \mathcal{N}$ definiere

$$d_z := \min_{T \in \mathcal{T}} 2^{\frac{3}{2}\delta(z,T) - \ell(T)} \qquad \text{mit } \delta(z,T) := \min_{x \in T} \delta(z,x).$$

Proposition 2.30. *(a) Für $a, b \in \mathcal{N}$ mit $\delta(a,b) = 1$ ist $d_a/d_b \leq \sqrt{8}$.*

(b) Für $z \in \mathcal{N}(T)$, $T \in \mathcal{T}$ gilt

$$\frac{d_z^2}{|T|} + \frac{|T|}{d_z^2} \lesssim 1.$$

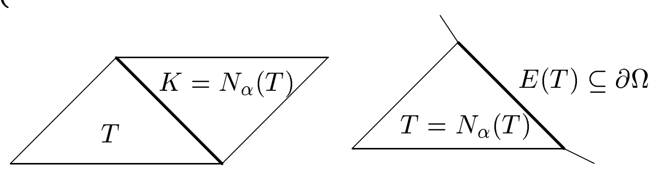
Als letztes Argument verbleibt einem die Analysis von lokalen Massenmatrizen (siehe auch C. Carstensen, Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomee criterion for H^1-stability of the L^2-projection onto finite element spaces, Math. Comp., 71, 237, 157-163).

(e) Für den Rest des Beweises ersetze $\mathcal{M}^{(k)}$ durch die Menge

$$\tilde{\mathcal{M}}^{(k)} := \left\{ T \in \mathcal{T}^{(k)} \,\middle|\, \text{wenigstens eine der Kanten von } T \text{ ist markiert, d.h. } E(T) \cap \mathcal{M}^{(k)} \neq \emptyset \right\}.$$

Definition 2.31 (N_α^β). In einer Triangulierung $\mathcal{T}^{(\alpha)}$ definiere $N_\alpha^{\beta+1} := N_\alpha \circ N_\alpha^\beta$, $N_\alpha^1 := N_\alpha$, $N_\alpha^0 :=$ Id und für $T \in \mathcal{T}^{(\alpha)}$ bezeichne

$$N_\alpha(T) := \begin{cases} K & \text{falls ex. } K \in \mathcal{T}^{(\alpha)} \setminus \{T\} \text{ mit } E^{(\alpha)}(T) = K \cap T \in \mathcal{E}^{(\alpha)} \text{ bzgl. } \mathcal{T}^{(\alpha)}; \\ T & \text{sonst.} \end{cases}$$



Proposition 2.32. *Für alle $\alpha = 1, 2, \ldots, L$ und alle $T \in \mathcal{T}^{(\alpha)} \setminus \mathcal{T}^{(0)}$ existiert genau ein $\beta = 0, 1, \ldots, \alpha - 1$, genau ein $K \in \mathcal{M}^{(\beta)}$ und genau ein $\gamma \in \mathbb{N}_0$, so daß $T \subseteq N_\beta^\gamma(K)$, d.h. T ist aus der Verfeinerung von $N_\beta^\gamma(K)$ im Schritt zur Erzeugung von $\mathcal{T}^{(\beta)}$ hervorgegangen.*

Beweisidee. N_β beschreibt eine Nachbarschaftsrelation im RGB-Abschluß: Ist $T \in \mathcal{T}^{(\beta)}$ markiert, so auch $E(T)$ (wegen Eigenschaft (R)), und damit ebenfalls $N_\beta(T)$. Setzt man dies fort, so sieht man, daß $\left\{N_\beta^\gamma(T) \,|\, \gamma = 0, 1, 2, \dots\right\}$ markiert sind.

Zur Berechnung von $\mathcal{T}^{(\beta+1)}$ werden genau die Elemente in $\bigcup_{\gamma=0}^{\infty}\left\{N_\beta^\gamma(T) \,\middle|\, T \in \mathcal{M}^{(\beta)}\right\}$ verfeinert. □

Proposition 2.33. *Für $T, K \in \mathcal{T}^{(\beta)}$ sei $T = N_\beta(K)$. Dann ist*

$$-\frac{1}{2} \le \ell(T) - \ell(K) \le 1.$$

Beweis. Falls T und K affine Bilder von rechtwinkligen, gleichschenkligen Dreiecken auf einem Referenzelement sind, dann sind wir fertig.

Alternativ sind T und K stückweise affine Bilder von rechtwinkligen, gleichschenkligen Dreiecken. □

Bequemlichkeitshypothese: Für alle $T, K \in \mathcal{T}^{(0)}$ gilt $(T \cap K \in \mathcal{E}^{(0)} \Rightarrow E(T) = E(K)$ oder $E(T) \ne T \cap K \ne E(K))$.

Dann gilt in der Proposition 2.33 (für $T, K \in \mathcal{T}^{(\beta)}$, $T = N_\beta(K)$) die Ungleichung

$$0 \le \ell(T) - \ell(K) \le \frac{1}{2}.$$

Proposition 2.34. *Zu $T \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}$ existiert genau eine endliche Folge $T_1, \dots, T_s$ (mit $s \ge 1$) in $\mathcal{M} = \bigcup_{\ell=0}^{L-1} \mathcal{M}^{(\ell)}$ mit $T_0 = T$ und für alle $\alpha = 0, 1, \dots, s-1$ existiert ein $\beta \ge 0$ und $\gamma \ge 0$, so daß für $\alpha = 0, 1, 2, \dots, s-1$ gilt*

$$\ell(T_\alpha) \le \ell(T_{\alpha+1}) + \frac{3}{2},$$

T_α ist Verfeinerung von $N_\gamma^\beta(T_{\alpha+1})$ in $\mathcal{T}^{(\gamma)}$ und

$$\ell(T_\alpha) \ge \ell(T), \quad \textit{aber } \ell(T_s) \le \ell(T) - \frac{1}{2}.$$

Beweis. Die Existenz der Folge $T_1, \dots, T_s$ folgt aus sukzessiver Anwendung von Proposition 2.32. Für $\alpha = 0, 1, \dots, s-1$ wird T_α als Verfeinerung eines $\hat{T}_\alpha$ zunächst $\ell(T_\alpha) - \ell\left(\hat{T}_\alpha\right) \in \{1/2, 1, 3/2\}$ erfüllen ($3/2$ ist nur für $T_\alpha \in \operatorname{bisec5}\left(\hat{T}_\alpha\right)$ möglich). Die Level der $N_\gamma^\beta(T_\alpha)$ sind aber echt größer oder gleich $\ell\left(\hat{T}_\alpha\right)$. Also gilt

$$\ell(T_{\alpha+1}) \ge \ell(T_\alpha) - \frac{3}{2}.$$ □

Bemerkung 2.35. Für $\alpha = 0, 1, \dots, s$ gilt

$$\ell(T_\alpha) - \ell(T) \in \mathcal{K} := \{-1, -1/2, 0, 1/2, \dots\} = -1 + 1/2\mathbb{N}_0$$

(Der Beweis kombiniert $\ell(T_s) \le \ell(T) \le \ell(T_{s-1})$ und $\ell(T_{s-1}) \le \ell(T_s) + \frac{3}{2}$).

Proposition 2.36. *Es existiert ein $c_1 = c_1\left(\mathcal{T}^{(0)}, \mathcal{E}^{(0)}, E_0\right) > 0$, so daß für alle $\alpha = 0, 1, \dots, s-1$:*

$$\operatorname{diam}(T_\alpha \cup T_{\alpha+1}) \le c_1 \operatorname{diam}(T)\, 2^{\ell(T) - \ell(T_{\alpha+1})}.$$

Beweis. Es ist

$$\begin{aligned}
\operatorname{diam}(T_\alpha \cup T_{\alpha+1}) &\leq \operatorname{diam}\left(N_\gamma^\beta(T_{\alpha+1}) \cup T_{\alpha+1}\right) \qquad (\text{mit } \beta \text{ minimal})\\
&\leq \sum_{\mu=0}^{\beta} \operatorname{diam}\left(N_\gamma^\mu(T_{\alpha+1}) \cup N_\gamma^{\mu+1}(T_{\alpha+1})\right) \qquad (\text{mit Dreiecksungl.})\\
&\lesssim \sum_{\mu=0}^{\beta} 2^{-\ell\left(N_\gamma^{\mu+1}(T_{\alpha+1})\right)} \qquad (\text{lokale Quasiuniformität})
\end{aligned}$$

Die Bequemlichkeitshypothese zeigt

$$\ell\left(N_\gamma^\mu(T_{\alpha+1})\right)_{\mu=0,\dots,\beta-1} = \left(\ell(T_{\alpha+1}), \ell(T_{\alpha+1}) + \frac{1}{2}, \ell(T_{\alpha+1}) + 1, \dots\right),$$

also ist

$$\begin{aligned}
\operatorname{diam}(T_\alpha \cup T_{\alpha+1}) &\lesssim 2^{-\ell(T_{\alpha+1})}\left(1 + 2^{-\frac{1}{2}} + 2^{-1} + 2^{-\frac{3}{2}} + \dots\right)\\
&\lesssim 2^{-\ell(T_{\alpha+1})}\\
&\lesssim \operatorname{diam}(T)\, 2^{\ell(T)-\ell(T_{\alpha+1})}.
\end{aligned}$$

□

Definition 2.37. Definiere $c_2 := c_1\left(2^{3/2} + \left(1 - 2^{-1/4}\right)^{-1}\right)$ mit c_1 aus Proposition 2.36 und

$$\Lambda(T,K) := \begin{cases} \left(\ell(K) - \ell(T) + \frac{5}{2}\right)^{-2} & \text{falls } T \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}, K \in \mathcal{M}, \ell(K) - \ell(T) \geq \frac{3}{2} \\ & \text{und } \operatorname{diam}(T \cup K) \leq c_2 \operatorname{diam}(T); \\ 0 & \text{sonst.} \end{cases}$$

<u>Beweisidee</u> für den Beweis von *(e)*: Λ erfüllt

$$\begin{aligned}
&1. \quad \forall K \in \mathcal{M} : \sum_{T \in \mathcal{T}^{(k)} \setminus \mathcal{T}^{(0)}} \Lambda(T,K) \lesssim 1;\\
&2. \quad \forall T \in \mathcal{T}^{(k)} \setminus \mathcal{T}^{(0)} : \sum_{K \in \mathcal{M}} \Lambda(T,K) \gtrsim 1.
\end{aligned}$$

Daraus folgt die Behauptung *(e)*, denn

$$\operatorname{card}\left(\mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}\right) = \sum_{T \in \mathcal{T}^{(k)} \setminus \mathcal{T}^{(0)}} 1 \overset{2.}{\lesssim} \sum_{K \in \mathcal{M}} \sum_{T \in \mathcal{T}^{(k)} \setminus \mathcal{T}^{(0)}} \Lambda(T,K) \overset{1.}{\lesssim} \sum_{K \in \mathcal{M}} 1 = \operatorname{card}(\mathcal{M}).$$

Definition 2.38 ($m(j,k)$)**.** Für fixiertes $T = T_0 \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}$ und in Proposition 2.34 definierter Folge $T_1, \dots, T_s$ bezeichne

$$m(j,k) := \operatorname{card}\{T_\nu \,|\, \nu = 0,1,2,\dots,j \text{ mit } \ell(T_\nu) = \ell(T) + k\}$$

für $j = 0,1,2,\dots,s$ und $k \in -1/2 + \mathcal{K} = \{-3/2, -1, -1/2, 0, 1/2, \dots\}$.

Proposition 2.39. *Wenn ein* $j \in \{1,2,\dots,s\}$ *mit*

$$\sup_{\mu \in 1/2\mathbb{N}_0} m(j,\mu)\, 2^{-\frac{\mu}{2}} \leq 1$$

existiert, dann gilt

$$\operatorname{diam}(T \cup T_j) \leq c_2 \operatorname{diam}(T).$$

Beweis. Mit der Dreiecksungleichung folgt für $j \in \{1, 2, \ldots, s-1\}$

$$\begin{aligned}
\operatorname{diam}(T_j \cup T_0) &\leq \sum_{\nu=0}^{j-1} \underbrace{\operatorname{diam}(T_{\nu+1} \cup T_\nu)}_{\leq c_1 \operatorname{diam}(T) 2^{\ell(T)-\ell(T_{\nu+1})} \text{ mit Prop. 2.36}} \\
&\leq c_1 \operatorname{diam}(T) \underbrace{\sum_{\nu=0}^{j-1} 2^{-\mu(\nu+1)}}_{\sum_{\nu=1}^{j} 2^{-\mu(\nu)}}, \qquad \mu(\nu+1) := \ell(T_{\nu+1}) - \ell(T) \geq 0 \\
&\leq c_1 \operatorname{diam}(T) \sum_{k \in 1/2\mathbb{N}} \underbrace{m(j,k) 2^{-k}}_{\leq 2^{-k/2}} \\
&\leq c_1 \operatorname{diam}(T) \sum_{k \in 1/2\mathbb{N}} 2^{-k/2} \\
&\leq c_2 \operatorname{diam}(T) \qquad \text{mit } c_2 \text{ wie oben definiert.}
\end{aligned}$$

Für$j = s$ ist $\mu(s) \in \{-3/2, -1, -1/2\}$ und $m(s, \mu(s)) = 1$. Die obige Abschätzung modifiziert sich daher und zeigt

$$\operatorname{diam}(T_s \cup T_0) \leq c_1 \operatorname{diam}(T) \left(\frac{1}{1 - 2^{-1/4}} + 2^{3/2} \right) \leq c_2 \operatorname{diam}(T). \qquad \square$$

Proposition 2.40 (Eigenschaft 2. von Λ). *Für alle $T \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}$ ist*

$$1 \lesssim \sum_{K \in \mathcal{M}} \Lambda(T, K).$$

Beweis. 1. Fall: Es existiert ein $j \in \{1, \ldots, s\}$ mit

$$1 < \sup_{k \in 1/2\mathbb{N}_0} 2^{-k/2} m(j,k). \tag{2.1}$$

O.B.d.A. sei $j \in \{1, \ldots, s\}$ minimal mit (2.1) gewählt. Zu diesem fixierten j existiert ein k mit

$$2^{k/2} < m(j,k). \tag{2.2}$$

O.B.d.A. sei k minimal mit (2.2) gewählt. Folglich gilt für alle $\nu = 0, 1, \ldots, j-1$

$$1 \geq \sup_{k \in 1/2\mathbb{N}_0} 2^{-k/2} m(\nu, k).$$

Mit Proposition 2.39 folgt $\operatorname{diam}(T_\nu \cup T) \leq c_2 \operatorname{diam}(T)$. Also

$$\sum_{\nu=1}^{j-1} \Lambda(T_\nu, T) = \sum_{\nu=1}^{j-1} \left(\underbrace{\ell(T_\nu) - \ell(T)}_{=:\mu(\nu)} + \frac{5}{2} \right)^{-2} = \sum_{\mu=\frac{1}{2}}^{\infty} \left(\mu + \frac{5}{2} \right)^{-2} m(j-1, \mu).$$

Wegen $m(j-1,\mu) \geq m(j,\mu) - 1 > 2^{k/2} - 1$ für $\mu = k > 0$ mit (2.2) folgt

$$\sum_{\nu=1}^{j-1} \Lambda(T_\nu, T) \geq \left(k + \frac{5}{2} \right)^{-2} \left(2^{k/2} - 1 \right) \geq \min_{\mu \in 1/2\mathbb{N}} \left(\mu + \frac{5}{2} \right)^{-2} \left(2^{\mu/2} - 1 \right) = \frac{(2^{1/4} - 1)}{16} > 0,$$

denn die Funktion $x \mapsto \left(x + \frac{5}{2}\right)^{-2} \left(2^{x/2} - 1\right)$ ist streng monoton wachsend für $x \geq 0$ weil die Ableitung erfüllt

$$\frac{\partial}{\partial x} \left(x + \frac{5}{4} \right)^{-1} (2^x - 1) = \left(x + \frac{5}{4} \right)^{-3} \left(2^x \left(\ln 2 \left(x + \frac{5}{4} \right) - 1 \right) + 1 \right) > 0.$$

2. Fall: Für alle $\mu \in 1/2\mathbb{N}_0$ ist

$$m(5,\mu) \leq 2^{\mu/2}.$$

Nach Proposition 2.39 folgt $\operatorname{diam}(T \cup T_s) \leq c_2 \operatorname{diam}(T)$ und so

$$\Lambda(T_s, T) = \left(\ell(T_s) - \ell(T) + \frac{3}{2}\right)^{-2} \geq 1 > 0. \qquad \square$$

Proposition 2.41. *Für alle $K \in \mathcal{M}$ gilt*

$$\sum_{T \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}} \Lambda(T, K) \lesssim 1.$$

Beweis. Für fixiertes $K \in \mathcal{M}$ und $k \in \mathcal{K}$ definiere man

$$\mathcal{A}_k := \left\{ T \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)} \,|\operatorname{diam}(T \cup K) \leq c_2 \operatorname{diam}(T) \text{ und } \ell(T) = k \right\}.$$

Da $|T| \gtrsim 2^{-2k}$ und $\bigcup \mathcal{A}_k \subseteq$Kugel mit Mittelpunkt $\operatorname{mid}(T)$ und Radius

$$R := \max_{T \in \mathcal{A}} c_2 \operatorname{diam}(T) \lesssim 2^{-k}$$

folgt

$$a_k := \operatorname{card}(\mathcal{A}_k) \leq \frac{|\bigcup \mathcal{A}|}{\min_{T \in \mathcal{A}_k} |T|} \lesssim \pi R^2 2^{-2k} \lesssim 1.$$

Folglich gilt

$$\sum_{T \in \mathcal{T}^{(L)} \setminus \mathcal{T}^{(0)}} \Lambda(T, K) \leq \sum_{k \in \mathcal{K}} \frac{a_k}{(k + 5/2)^2} \lesssim \sum_{k=-\frac{3}{2}}^{\infty} \frac{1}{(k + 5/2)^2} \leq \infty \qquad \square$$

Dies beendet den Beweis des Satzes 2.23. $\square$

2.6 Nichtkonforme und gemischte FEM

Notation. Ein FER $V_h \subseteq V = H_0^1(\Omega)$ heißt *konform* im Zusammenhang mit elliptischen partiellen Differentialgleichungen 2ter Ordnung.

Die bisherigen FEM waren alle konform.

Nun betrachten wir einen FER $V_h \not\subseteq V$, der dann nicht konform ist, und *nichtkonform* heißt.

Abkürzung: NC:=nichtkonform (Englisch: nonconforming).

Gemischte FEM (Abkürzung MFEM, Englisch: mixed FEM) diskretisieren $-\Delta u = f$ als System von partiellen Differentialgleichungen erster Ordnung: $\nabla u = p,\ f + \operatorname{div} p = 0$.

2.6.1 Gebrochene Sobolev-Räume

Englisch: broken Sobolev spaces

Besser ist vielleicht die Bezeichnung *stückweise Sobolev-Räume.*

Definition 2.42 ($H^1(\omega)$). Für ein nichtleeres Gebiet ω im $\mathbb{R}^n$ definiere

$$H^1(\omega) := \left\{ f \in L^2(\omega) \,\middle|\, f \text{ ist schwach diff.bar und } Df \in L^2(\omega;\mathbb{R}^n) \right\}$$

mit Sobolev-Norm

$$\|\cdot\|_{1,2} := \|\cdot\|_{H^1(\omega)}$$

und Halbnorm

$$|\cdot|_{1,2} := |\cdot|_{H^1(\omega)} := \|D\cdot\|_{L^2(\omega)}.$$

Bemerkung 2.43. 1. $H^1(\omega)$ ist ein Hilbert-Raum.

2. $H^1(\omega) \nsubseteq C(\omega)$ für $n \geq 2$.

3. Für zwei Dreiecke T_+, T_- im $\mathbb{R}^2$, $n = 2$, mit gemeinsamer Kante $E = T_+ \cap T_-$ und 2 Funktionen f_+ und f_-, $f_\pm \in H^1(T_\pm) := H^1(\text{Inneres}(T_\pm))$, $\omega_E := \text{Inneres}(T_+ \cup T_-)$ definiere vermöge

$$f(x) := \begin{cases} f_+^*(x) & \text{falls } x \in T_+ \setminus E \\ f_-^*(x) & \text{falls } x \in T_- \setminus E \end{cases}$$

ein f in $L^2(\omega_E)$.

Dann gilt $f \in H^1(\omega_E) \Leftrightarrow [f]_E(x) = 0$ für alle x in E außerhalb einer Menge vom eindimensionalen Lebesgue-Maß 0. Für den Sprung $[f]_E$ von f über E definiere

$$\begin{aligned} [f]_E(x) &:= \lim_{\substack{T_+ \ni x_+ \to x \\ T_- \ni x_- \to x}} \left(f^*(x_+) - f^*(x_-)\right), \\ [f]_E(x) &= f_+^*(x) - f_-^*(x), \end{aligned}$$

wobei

$$f_\pm^*(x) = \begin{cases} \lim\limits_{\delta \to 0} \fint_{B(x,\delta) \cap T_\pm} f(y)\,dy & \text{falls der Limes existiert,} \\ 0 & \text{sonst.} \end{cases}$$

Beweisidee. (a) Sprungsatz (vgl. unten)

(b) Partielle Integration: Für alle $\varphi \in \mathcal{D}(\omega_E)^n$:

$$-\int_{\omega_E} f \operatorname{div} \varphi dx = -\int_{T_+} f_+ \operatorname{div} \varphi dx - \int_{T_-} f_- \operatorname{div} \varphi dx$$

Bezeichne mit ν_+ die äußere Normale. Dann

$$\int_{T_+} f_+ \operatorname{div} \varphi dx = \int_{T_+} (Df_+) \cdot \varphi dx - \int_{\partial T_+} f_+^* \varphi \cdot \nu_+ ds,$$

also, für $D_h f$ stückweise durch $Df_\pm$ definiert,

$$-\int_{\omega_E} f \operatorname{div} \varphi dx = \int_{\omega_E} D_h f \cdot \varphi dx - \int_E \underbrace{\left(f_+^* \nu_+ + f_-^* \nu_-\right)}_{=\left(f_+^* - f_-^*\right)\nu_E = 0 \text{ für } \nu_+ =: \nu_E = -\nu_- \text{ auf } E} \varphi ds.$$

Folglich ist der stückweise definierte Gradient gleich der schwachen Ableitung. □

Definition 2.44 (NCFER). Für eine reguläre Triangulierung $\mathcal{T}$ von $\Omega \subseteq \mathbb{R}^2$ mit Kantenmenge $\mathcal{E}$ und $\mathcal{E}_{\partial\Omega} := \{E \in \mathcal{E} \,|\, E \subseteq \partial\Omega\}$ definiere

$$\begin{aligned} H^1(\mathcal{T}) &:= \left\{ f \in L^2(\Omega) \,\middle|\, \forall T \in \mathcal{T} \quad f|_T \in H^1(T) \right\} \equiv \prod_{T \in \mathcal{T}} H^1(T) \\ P_1(\mathcal{T}) &:= \left\{ f \in L^2(\Omega) \,|\, \forall T \in \mathcal{T} \quad f|_T \in P_1(T) \right\} \\ CR(\mathcal{T}) &:= \{ f \in P_1(\mathcal{T}) \,|\, \forall E \in \mathcal{E} \quad f \text{ ist stetig bei } \operatorname{mid}(E) \} \\ CR_0(\mathcal{T}) &:= \{ f \in CR(\mathcal{T}) \,|\, \forall E \in \mathcal{E}_{\partial\Omega} \quad f(\operatorname{mid}(E)) = 0 \} \end{aligned}$$

Bemerkung 2.45. 1. CR steht für *Crouzeix-Raviart*. Der Raum $CR_0(\mathcal{T}) =: V_h \not\subset V$ (für $\operatorname{card}\mathcal{T} > 1$) ist der einfachste NCFER.

2. Die Stetigkeitsbedingungen an f in V sind für $f_h \in CR_0(\mathcal{T})$ abgeschwächt im Integralmittel:

$$\begin{aligned} \forall E \in \mathcal{E}_\Omega := \mathcal{E} \setminus \mathcal{E}_{\partial\Omega} \qquad & \int_E [f_h]_E \, ds = 0 \\ \text{und } \forall E \in \mathcal{E}_{\partial\Omega} \qquad & \int_E f ds = 0. \end{aligned}$$

3. Df_h existiert im Distributionensinne, ist i.a. aber keine L^p-Funktion, d.h. i.a. ist f_h nicht schwach differenzierbar.

   $D_h f_h$:=stückweiser Gradient existiert für alle $f_h \in H^1(\mathcal{T})$, z.B. für $f \in CR(\mathcal{T}) \subsetneq H^1(\mathcal{T})$. (alternative Notation: $D_\mathcal{T} f_h$)

4. $P_1(\mathcal{T}) \cap C(\Omega) =: P_1$-FER $\subset CR(\mathcal{T}) =: P_1^{\mathrm{NC}}$-FER.

5. $P_1(\mathcal{T}) \cap H_0^1(\Omega) \subset CR_0(\mathcal{T})$.

2.6.2 NCFEM

Definition 2.46 (CRFEM). Die *Crouzeix-Raviart FEM (CRFEM)* zum Poisson-Problem $-\Delta u = f$ in Ω, $u = 0$ auf $\partial\Omega$ bzgl. einer regulären Triangulierung $\mathcal{T}$ von Ω lautet

$$\exists! u_h \in V_h := CR_0 \, \forall v_h \in V_h \quad a_h(u_h, v_h) = F(v_h) \tag{CRFEM}$$

Hier ist

$$\begin{aligned} a_h(\tilde{u}_h, \tilde{v}_h) &:= \int_\Omega D_h\tilde{u}_h \cdot D_h\tilde{v}_h dx, \\ F(\tilde{v}_h) &:= \int_\Omega f(x)\, \tilde{v}_h(x)\, dx \end{aligned}$$

für alle $\tilde{u}_h, \tilde{v}_h \in H^1(\mathcal{T})$.

Beweis der Eindeutigkeit. Falls $u_h \in V_h$ erfüllt $a(u_h, v_h) = 0$ für alle $v_h \in V_h$, folgt $\|D_h u_h\|_{L^2(\Omega)} = 0$. Dann ist $u_h|_T = \mathrm{const}(T)$ für alle $T \in \mathcal{T}$. Da also $[u_h]_E = \mathrm{const}$ und $\int_E [u_h]_E \, ds = 0$ für alle $E \in \mathcal{E}$, folgt $[u_h]_E = 0$ und $u_h = 0$ auf $\partial\Omega$. Also ist $u_h \equiv 0$. □

Bemerkung 2.47. (CRFEM) ist ein lineares Gleichungssystem. Aus der Eindeutigkeit von Lösungen folgt auch deren Existenz!

2.6.3 Nodale Basis von CRFER

Bezeichne mit

$$\mathcal{M} := \{\mathrm{mid}(E) \,|E \in \mathcal{E}\}$$

die Menge der Kantenmittelpunkt in einer Triangulierung $\mathcal{T}$. Die Werte $(u_h(m))_{m\in\mathcal{M}}$ bestimmen u_h eindeutig in $CR(\mathcal{T})$ durch stückweise lineare Interpolation.

Definition 2.48 (nodale Basis von CRFER). Zu $E \in \mathcal{E}$ mit Mittelpunkt $m := \mathrm{mid}(E)$ definiere ψ_E in $CR(\mathcal{T})$ vermöge $\psi_E(m) = 1$ und $\psi_E(n) = 0$ für alle $n \in \mathcal{M} \setminus \{m\}$ (und elementweise lineare Interpolation dieser Werte).

$(\psi_E)_{E\in\mathcal{E}}$ heißt *nodale Basis* von $CR(\mathcal{T})$. $\mathcal{E}_\Omega := \mathcal{E} \setminus \mathcal{E}_{\partial\Omega}$.

$(\psi_E)_{E\in\mathcal{E}_\Omega}$ heißt *nodale Basis* von $CR_0(\mathcal{T})$.

Bemerkung 2.49. Es gilt

$$\begin{aligned} \operatorname{span}\{\psi_E \,|E \in \mathcal{E}\} &= CR(\mathcal{T}), \\ \operatorname{span}\{\psi_E \,|E \in \mathcal{E}_\Omega\} &= CR_0(\mathcal{T}). \end{aligned}$$

Bemerkung 2.50. Es ist $\operatorname{supp}\psi_E = \overline{\omega_E} = T_+ \cup T_-$ für $E \in \mathcal{E}_\Omega$.

Es gilt $u_h \in CR_0(\mathcal{T})$, gdw. ein $(x_E)_{E\in\mathcal{E}_\Omega}$ existiert, so daß $u_h = \sum_{E\in\mathcal{E}_\Omega} x_E \psi_E$.

Definition 2.51 (lokale STEMA). Für ein Dreieck T mit NC nodalen Basisfunktionen ψ_1, ψ_2, ψ_3 ist

$$\mathrm{NCSTEMA}(T) := \left(\int_T D\psi_j \cdot D\psi_k dx\right)_{j,k=1,2,3} \in \mathbb{R}^{3\times 3}.$$

Proposition 2.52. *Es gilt*

$$\mathrm{NCSTEMA}(T) = 4\,\mathrm{STEMA}(T)$$

für die lokale Steifigkeitsmatrix der P_1-FEM STEMA (T) *von* T.

Beweisidee. Auf dem Teildreieck $T_4 := \operatorname{conv}\{M_1, M_2 M_3\}$ ist $(D\psi_j \cdot D\psi_k)$ (auf T konstant) gleich dem Eintrag für die konforme FEM zu den Baryzentrischen Koordinaten ψ_1, ψ_2, ψ_3. Der Faktor 4 erscheint durch Skalierung und die Umrechnung von $M_j = (P_{j+1} + P_{j+2})/2$ auf die Eckpunkte P_j. □

Kapitel 3

Konvergenzanalysis

Ziel. Fehlerabschätzungen für FEM

a˜priori	a˜posteriori
vor der Rechung	nach der Rechnung
u glatt, scheinbar bekannt u_h unbekannt kommt auf RHS nicht vor	u völlig unbekannt und kommt auf RHS nicht vor u_h bekannt und geht in RHS ein

allgemeine Argumente: Quasi-Optimalität

spezielle Argumente: Approximationsaussage $|||u - Iu|||$ etc.

3.1 Sobolev-Räume

Definition 3.1 (H^m). Für ein nichtleeres Gebiet Ω im $\mathbb{R}^n$ mit Rand $\partial\Omega$, $m \in \mathbb{N}$, $n \in \mathbb{N}$, $m, n \geq 1$, $H^0(\Omega) = L^2(\Omega)$, bezeichne den Vektorraum

$$H^m(\Omega) := \left\{ u \in H^{m-1}(\Omega) \,\middle|\, Du \text{ ex. im schwachen Sinn und } Du \in H^{m-1}(\Omega; \mathbb{R}^n) \right\}$$

als den *Sobolev-Raum* m-ter Ordnung.

$\|\cdot\|_{m,2}$ bzw. $|\cdot|_{m,2}$ heißt *Sobolev-Norm* bzw. *-Halbnorm*.

Satz 3.2. *Der Raum $H^m(\Omega)$ ist ein Hilbert-Raum.*

Beweis. Wie im eindimensionalen Fall (Satz 1.27, Seite 12). □

Satz 3.3 (Spursatz). *Für ein Lipschitz-Gebiet ist die Abbildung*

$$\gamma : H^1(\Omega) \to L^2(\partial\Omega), \qquad u \mapsto u^*|_{\partial\Omega},$$

wohldefiniert, linear und stetig.

Bemerkung 3.4. Man sagt, "H^1-Funktionen haben Spuren" und γ heißt *Spuroperator*.

Beweis. Liegt tiefer, insbesondere weil i.a. $L^2(\partial\Omega)$ bzgl. des $(n-1)$-dimensionalen Hausdorff-Maßes $\mathcal{H}_{n-1}$ definiert ist. Für stückweise affine Ränder aber ist $\mathcal{H}_{n-1}$ gleich dem (transformierten) $(n-1)$-dimensionalen Lebesgue-Maß $\mathcal{L}_{n-1}$.

Details z.B. in [EG]. □

Satz 3.5 (partielle Integration). *Für ein beschränktes Lipschitz-Gebiet Ω mit Rand $\Gamma := \partial\Omega$ ist der äußere Einheitsnormalenvektor $\nu(x)$ für $\mathcal{H}_{n-1}$-fast alle x auf Γ im klassischen Sinne als Limes von Differenzenquotienten definiert und für alle $u, v_1, \ldots, v_n \in H^1(\Omega)$ ist*

$$\int_\Omega u \operatorname{div} v \, dx + \int_\Omega v \cdot Du \, dx = \int_{\partial\Omega} u \, (v \cdot \nu) \, ds,$$

wobei u und $v = (v_1, \ldots, v_n)$ auf $\partial\Omega$ im oben genannten Sinn als Spuren verstanden werden (d.h. RHS$= \int_{\partial\Omega} u^*(x) \, v^*(x) \cdot \nu(x) \, ds$, $ds = d\mathcal{H}_{n-1}(x)$*)*

Beweis. Beweisideen in der Übung. Für Details siehe [EG] etc. □

Satz 3.6 (Meyers). *Für*

$$\mathcal{D}(\Omega) := \{ f \in C^\infty(\mathbb{R}^n) \, | \operatorname{supp} f \Subset \Omega \} = C_c^\infty(\Omega)$$

gilt

(a) $\mathcal{D}(\mathbb{R}^n)$ ist dicht in $H^1(\Omega)$ und

$$\overline{\mathcal{D}(\mathbb{R}^n)}^{\|\cdot\|_{H^1(\Omega)}} = H^1(\Omega).$$

(b) Es gilt

$$H_0^1(\Omega) = \left\{ u \in H^1(\Omega) \, | \gamma u = 0 \right\} \stackrel{!}{=} \overline{\mathcal{D}(\Omega)}^{\|\cdot\|_{H^1(\Omega)}}.$$

Beweis. Siehe [EG], es werden Regularisierungen benutzt. □

Bemerkung 3.7. Es ist

$$H^1(\mathbb{R}^n) = \left\{ f \in L^2(\mathbb{R}^n) \;\middle|\; \begin{array}{l} \forall j = 1, 2, \ldots n, \text{ für fast alle } x' \in \mathbb{R}^{n-1} \text{ ist die Funktion} \\ f\left(x_1', \ldots, x_{j-1}', \cdot, x_j', \ldots, x_{n-1}'\right) \text{ absolut stetig und} \\ \frac{\partial f}{\partial x_j} := \text{Ableitung von } f\left(x_1', \ldots, x_{j-1}', \cdot, x_j', \ldots, x_{n-1}'\right) \\ \text{ist definiert als Funktion in } L^2(\mathbb{R}^n) \end{array} \right\}$$

Beweis mit Fubini in [EG].

Satz 3.8 (Rellich). *Die Einbettung $H^1(\Omega) \hookrightarrow L^2(\Omega)$ ist kompakt, d.h. jede in $H^1(\Omega)$ beschränkte Folge (f_j) hat eine in $L^2(\Omega)$ konvergente Teilfolge (d.h. es existiert ein $g \in H^1(\Omega)$ und eine Teilfolge (f_{j_k}) mit $\lim_{k\to\infty} \|g - f_{j_k}\|_{L^2(\Omega)} = 0$).*

Beweisidee für den eindimensionalen Fall. Sei (f_j) eine beschränkte Folge in $H^1(\Omega)$, $\Omega = (a, b)$.

$$m_j := \fint_\Omega f_j(x) \, dx \to m$$

in $\mathbb{R}$. Durch Übergang zu $f_j - m_j$ sieht man, daß man o.B.d.A. $m_j = 0$ annehmen kann.

Nun ist

$$|f_j(x) - f_j(y)| = \left| \int_x^y f_j'(z) \, dz \right| \leq |x - y|^{\frac{1}{2}} \underbrace{\left\| f_j' \right\|_{L^2(\Omega)}}_{\leq M}.$$

Dann ist (f_j) <u>gleichgradig stetig</u> auf dem Kompaktum $[a, b] = \bar{\Omega}$.

Da f_j eine Nullstelle hat, ist f_j beschränkt unabhängig von j (denn für eine Nullstelle ξ von f_j ist $|f_j(x)| = \left| \int_\xi^x f_j'(z) \, dz \right| \leq (b - a)^{\frac{1}{2}} M$).

Mit dem Satz von Arzela-Ascoli existiert eine Teilfolge $(f_{j_k}) \to g$ in $C[a, b]$. □

Satz 3.9 (Poincaré-Ungleichung). *Für ein beschränktes, nichtleeres Gebiet Ω existiert eine Konstante $c_P(\Omega) > 0$, so daß*

$$\forall f \in H^1(\Omega)\ \exists m \in \mathbb{R} \quad \|f - m\|_{L^2(\Omega)} \le c_P(\Omega)\, \|Df\|_{L^2(\Omega)}\,.$$

Beweis (indirekt). Andernfalls existiert eine Folge $f_j \in H^1(\Omega)$ mit

$$\fint_\Omega f_j(x)\, dx = 0$$

(dies fixiert m_j) und

$$1 = \|f_j\|_2\,,$$

aber

$$|f_j|_{1,2} \le {}^1\!/_j, \qquad \text{für } j = 1, 2, \ldots$$

Also ist (f_j) beschränkt in $H^1(\Omega)$. Aus dem Satz 3.8 von Rellich folgt, daß eine Teilfolge (f_{j_k}) und ein $f \in H^1(\Omega)$ mit

$$\|f - f_{j_k}\|_2 \to 0.$$

Dann ist $\|f\|_2 = 1$. Wegen $|f_{j_k}|_{1,2} \to 0$ folgt $Df = 0$ fast überall. Da Ω zusammenhängend ist, folgt $f =$ konst. (Technisch müßte man diese Aussage noch für den schwachen Ableitungsbegriff zeigen (z.B. mit Regularisierung)). Wegen

$$\fint_\Omega f(x)\, dx = 0$$

folgt dann $f \equiv 0$, also $\|f\|_2 = 0$. Widerspruch. □

Satz 3.10 (Friedrichs-Ungleichung). *Für ein Gebiet Ω der Breite $\le b$ (d.h. Ω liegt zwischen 2 parallelen Hyperebenen mit Abstand $b > 0$) und $f \in H_0^1(\Omega)$ gilt*

$$\|f\|_{L^2(\Omega)} \le \frac{b}{\pi}\, \|Df\|_{L^2(\Omega)}\,.$$

Korollar 3.11. *Die Abbildung*

$$a(u,v) := \int_\Omega Du \cdot Dv dx$$

ist ein Skalarprodukt auf $H_0^1(\Omega)$ mit Norm $|\cdot|_{1,2} \approx \|\cdot\|_{1,2}$.

Beweis. O.B.d.A. ist $\Omega \subseteq (0,b) \times \mathbb{R}^{n-1} =: S$ und $f \in \mathcal{D}(\Omega)$ auf S durch Null fortgesetzt.

Für jeden Parameter $x' \in \mathbb{R}^{n-1}$ ist $f(\cdot, x') \in H_0^1(0,b)$. Also ist

$$\|f(\cdot, x')\|_{L^2(a,b)} \le \frac{b}{\pi} \left\| \frac{\partial f}{\partial x_1}(\cdot, x') \right\|_{L^2(a,b)}$$

nach der eindimensionalen Friedrichsungleichung (1.2) (Seite 13, der Beweis von $C_F = \frac{b}{\pi}$ folgt später).

Quadratur und Integration über $x' \in \mathbb{R}^{n-1}$,

$$\int_{\mathbb{R}^{n-1}} \|f(\cdot, x')\|^2_{L^2(a,b)}\, dx' \le \int_{\mathbb{R}^{n-1}} \left(\frac{b}{\pi}\right)^2 \left\| \frac{\partial f}{\partial x_1}(\cdot, x') \right\|^2_{L^2(a,b)} dx'$$

zeigt die Behauptung. □

3.2 Poincaré-Ungleichung

Ziel. Beweis von $c_P(\Omega) \leq \frac{1}{\pi} \operatorname{diam}(\Omega)$ für ein konvexes Gebiet $\Omega \subseteq \mathbb{R}^n$.

Hilfsmittel: Eindimensionale Abschätzungen mit Gewichten.

Satz 3.12 (Payne-Weinberger, 1961). *ρ sei eine konkave Funktion auf (a,b) mit*

$$0 \leq \rho_0 \leq \rho(x) \leq \rho_1 < \infty \qquad \text{für alle } x \in [a,b]$$

und Konstanten $a < b$, $\rho_0 \leq \rho_1$. Dann gilt für alle $f \in H^1(a,b)$

$$\min_{m\in\mathbb{R}} \int_a^b \rho(x)(f(x)-m)^2\,dx \leq \left(\frac{b-a}{\pi}\right)^2 \int_a^b \rho(x)(f'(x))^2\,dx. \tag{3.1}$$

Bemerkung 3.13. $\rho \equiv 1$ zeigt $c_p(a,b) \leq \frac{b-a}{\pi}$ und damit

$$c_P(a,b) = \frac{b-a}{\pi}$$

nach Beispielen aus den Übungen.

Beweisskizze.
1. O.B.d.A. sei $a = 0$ und $b = 1$. Ansonsten wende affine Variablentransformationen an!
2. Das Minimum auf der linken Seite von (3.1) wird für

$$Mf := \frac{\int_0^1 f(x)\rho(x)\,dx}{\int_0^1 \rho(x)\,dx} \in \mathbb{R}$$

angenommen. O.B.d.A. ist $Mf = 0$ möglich.
3. O.B.d.A. sei $\rho \in C^2(a,b)$, denn konkave Funktionen lassen sich bzgl. $\|\cdot\|_\infty$ durch glatte konkave Funktionen approximieren (z.B. $\rho_\varepsilon := \eta_\varepsilon * \rho$ ist konkav).
4. O.B.d.A. sei $\rho_0 > 0$, denn sonst betrachte $\rho + \varepsilon$ für $\varepsilon > 0$ und lasse am Schluß $\varepsilon \to 0$.

Für den Rest des Beweises sei also: $a = 0$, $b = 1$, $0 < \rho_0 \leq \rho$, ρ konkav mit $\rho \in C^2[a,b]$.

Schwache Formen: L_ρ sei definiert als der Hilbert-Raum $L^2(0,1)$ mit Skalarprodukt

$$(\cdot,\cdot)_\rho := \int_0^1 \rho(x)\cdot\cdot\,dx$$

mit Norm

$$\|\cdot\|_\rho := \left\|\rho^{\frac{1}{2}}\cdot\right\|_2.$$

Der Operator $M : L_\rho \to L_\rho$ ist selbstadjungiert.

Definiere nun weiter die Operatoren $S_\pm : L^2(0,1) \to L^2(0,1)$ vermöge

$$\begin{aligned}(S_+f)(x) &:= \int_0^x f(y)\,dy,\\ (S_-f)(x) &:= \int_x^1 f(y)\,dy,\end{aligned}$$

und

$$T := (1-M)S_+ : L^2(0,1) \to L^2(0,1).$$

Dann sind M, S_+, S_-, $T : L^2(0,1) \to L^2(0,1)$ linear und beschränkt.

Übungsaufgaben:

1. $M : L_\rho \to L_\rho$ ist selbstadjungiert und kompakt.
2. $S_\pm$ ist adjungiert zu $S_\mp$in $L^2(0,1)$ und $S_\pm$ ist kompakt.
3. $T^* := \rho^{-1} S_- \rho (1 - M)$ ist adjungiert zu T in L_ρ.
4. $A := T^*T : L_\rho \to L_\rho$ ist linear, beschränkt, selbstadjungiert, kompakt und positiv semidefinit.
5. $\|T\|^2_{L(L_\rho, L_\rho)} \le \|A\|_{L(L_\rho, L_\rho)}$

Spektraltheorie kompakter selbstadjungierter Operatoren:

Die Operatornorm $\lambda := \|A\|_{L(L_\rho, L_\rho)} > 0$ ist ein Eigenwert von A mit Eigenvektor $u \in L_\rho$, $\|u\|_\rho = 1$, und $Au = \lambda u$.

Proposition 3.14. *Für alle $f \in H^1(0,1)$ gilt*

$$\min_{m \in \mathbb{R}} \int_0^1 \rho\,(f - m)^2\, dx \le \lambda \int_0^1 \rho\,(f')^2\, dx.$$

Beweis. O.B.d.A. sei $f \in H^1(0,1)$ mit $f(0) = 0$. Dann ist

$$S_+ f' = f$$

und

$$\|Tf'\|_\rho = \|(1 - M) f\|_\rho = \left(\min_{m \in \mathbb{R}} \int_0^1 \rho\,(f - m)^2\, dx\right)^{1/2}.$$

Weiterhin ist $\|Tf'\|_\rho \le \|T\|_{L(L_\rho, L_\rho)} \|f'\|_\rho$, also

$$\min_{m \in \mathbb{R}} \int_0^1 \rho\,(f - m)^2\, dx \le \|T\|^2_{L(L_\rho, L_\rho)} \|f'\|^2_\rho \overset{5.}{\le} \underbrace{\|A\|_{L(L_\rho, L_\rho)}}_{\lambda} \int_0^1 \rho\,(f')^2\, dx.$$ □

Proposition 3.15. *Der Eigenvektor u und $\mu := \frac{1}{\lambda}$ lösen das Sturm-Liouville-Eigenwertproblem*

$$\left(\rho^{-1} (\rho u)'\right)' + \mu u = 0 \quad \textit{in } (0,1),$$

$$u(0) = u(1) = 0,$$

im klassischen Sinn, d.h. $u, v := \rho u \in H^1_0(0,1) \cap H^2(0,1)$ und

$$\left(\rho^{-1} v'\right)' + \mu \rho^{-1} v = 0$$

gilt wie

$$v'' - \left(\frac{\rho'}{\rho}\right) v' + \mu v = 0$$

punktweise in $(0,1)$.

Beweis. $w := Tu \in H(0,1)$ erfüllt $\lambda u = T^* w$, also

$$\lambda v = \lambda \rho u = \rho T^* w = S_- \rho (1 - M) w \in H^2(0,1)$$

mit

$$\lambda v(1) = \int_1^1 \rho (1 - M) w dx = 0$$

und

$$\lambda v(0) = \int_0^1 \rho(1-M)\, w dx = \int_0^1 \rho w dx - (Mw) \int_0^1 \rho dx = 0.$$

Also ist $v, u \in H_0^1(0,1) \cap H^2(0,1)$ und

$$v'(x) = -\mu(\rho(1-M)\, w)(x),$$

also

$$\rho^{-1} v' + \mu(1-M)\, w = 0,$$

und dann auch

$$\left(\rho^{-1} v'\right)' + \mu w' = 0,$$

$w' = (Tu)' = (S_+ u)' = u$. Also

$$0 = \left(\rho^{-1} v'\right)' + \mu u = \rho^{-1} v'' - \rho^{-2} \rho' v' + \rho^{-1} \mu v. \qquad \square$$

Proposition 3.16. $c_P(0,1) = \frac{1}{\pi}$.

Beweis. Für $\rho \equiv 1$ folgt in Prop. 3.15:

$$u'' + \mu u = 0 \quad \text{mit } u(0) = u(1) = 0.$$

Ein allgemeiner Ansatz mit $\cos -$ und $\sin -$Funktionen zeigt

$$u(x) = \sin(k\pi x) \quad \text{für } k = 1, 2, \ldots$$

und μ ist minimal für $k = 1$, d.h.

$$\mu = \pi^2. \qquad \square$$

Proposition 3.17. $c_F(0,1) = \frac{1}{\pi}$.

Beweis. Ein Transformationsargument führt die Behauptung auf $c_P = \frac{1}{\pi}$ zurück. $\square$

Die folgende Proposition zeigt $\lambda \le \pi^{-2}$ und beschließt damit den Beweis zu Satz 3.12. $\square$

Proposition 3.18. *Für eine konkave Funktion ρ gilt*

$$\pi^2 \le \mu$$

(d.h. $\lambda \le \frac{1}{\pi^2}$ und das zeigt den Satz!).

Beweis. Die Funktion $v \in H_0^1(0,1) \cap H^2(0,1)$ aus Prop. 3.15 erfüllt

$$v'' - \left(\frac{\rho'}{\rho}\right) v' + \mu v = 0,$$

also

$$\underbrace{\int_0^1 v v'' dx}_{-\|v'\|_2^2} \underbrace{- \int_0^1 \frac{\rho'}{\rho} v' v dx}_{-\frac{1}{2}\int_0^1 \frac{\rho'}{\rho}(v^2)' dx} + \underbrace{\int_0^1 \mu v^2 dx}_{\mu\|v\|_2^2} = 0.$$

Wegen

$$-\frac{1}{2}\int_0^1 \frac{\rho'}{\rho}\left(v^2\right)' dx = \frac{1}{2}\int_0^1 \left(\frac{\rho'}{\rho}\right)' v^2 dx = \frac{1}{2}\int_0^1 \left(\underbrace{\frac{\rho''}{\rho}}_{\le 0} - \frac{\rho'^2}{\rho^2}\right) v^2 dx \le 0$$

gilt also mit Proposition 3.17

$$\pi^2 \|v\|_2^2 \leq \|v'\|_2^2 = \int_0^1 \left(\frac{1}{2} \underbrace{\left(\frac{\rho'}{\rho}\right)'}_{\leq 0} + \mu \right) v^2 dx \leq \mu \|v\|_2^2 ,$$

d.h. $\pi^2 \leq \mu$ (da $\|v\|_2 \neq 0$). $\square$

Satz 3.19 (Payne-Weinberger). *Für eine konvexe Menge $\Omega \Subset \mathbb{R}^n$ gilt*

$$c_P(\Omega) \leq \frac{1}{\pi} \operatorname{diam}(\Omega).$$

Diese Abschätzung ist optimal in dem Sinne, daß

$$\sup_{\Omega \text{ konvex}} \frac{c_P(\Omega)}{\operatorname{diam}(\Omega)} = \pi.$$

Beweis. Das Integralmittel von f aus Ω werde mit

$$f_\Omega := \fint_\Omega f(x)\, dx$$

bezeichnet. Dann lautet die Behauptung

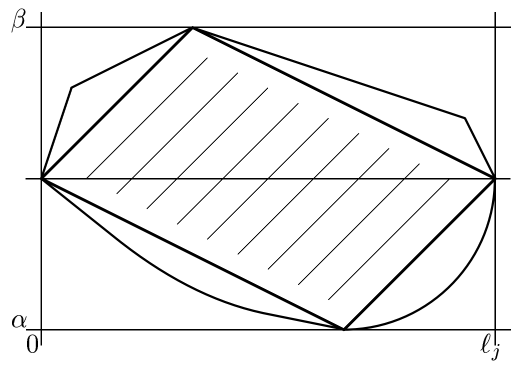
$$\|f - f_\Omega\|_2 \leq \frac{1}{\pi} \operatorname{diam}(\Omega) \, |f|_{1,2} .$$

1. O.B.d.A. kann in der Behauptung $f \in C^\infty(\mathbb{R}^n)|_\Omega$ mit $\int_\Omega f dx = 0$ angenommen werden. Für ein solches f ist
$$M := \|f\|_\infty + \|Df\|_\infty$$
eine fixe Konstante, $|\Omega| =: A$. Der Beweis wird nur für $n = 2$ ausgeführt.

2. Definiere $\Omega_{0,1} := \Omega$ und zu $\Omega_{j,1}, \ldots, \Omega_{j,2^j}$ definiere man $\Omega_{j+1,1}, \ldots, \Omega_{j+1,2^{j+1}}$ durch Bisektion jedes der konvexen und beschränkten $\Omega_{j,1}, \ldots, \Omega_{j,2^j}$. Dabei wird $\Omega_{j,m}$ in zwei flächengleiche konvexe Gebiete $\Omega_{j+1,m}, \Omega_{j+1,m+1}$ mit
$$\int_{\Omega_{j+1,m}} f(x)\, dx = 0 = \int_{\Omega_{j+1,m+1}} f(x)\, dx$$
geteilt. (Sandwich-Problem erlaubt Lösungen!) Dann ist
$$|\Omega_{j,m}| = 2^{-j} A.$$

3. Für alle $\varepsilon > 0$ existieren konvexe, paarweise disjunkte, offene Mengen $\Omega_1, \ldots, \Omega_m$ mit $\bar{\Omega} = \bar{\Omega}_1 \cup \ldots \cup \bar{\Omega}_m$,
$$\frac{A}{m} = |\Omega_1| = \ldots = |\Omega_m| , \qquad \int_{\Omega_j} f(x)\, dx = 0 \quad \text{für alle } j = 1, 2, \ldots, m,$$
und Ω_j ist kongruent zu einer Teilmenge von $(0, \ell_j) \times \left(-\frac{\varepsilon}{2}, +\frac{\varepsilon}{2}\right)$, $\ell_j := \operatorname{diam}(\Omega_j)$.

   Beweis. Für j in 2. groß sei $b := \beta - \alpha \leq \ell_j$.
   Dann gilt
$$\frac{b\ell_j}{2} \leq |\Omega_j| \leq 2^{-j} A,$$
   also wegen
$$\frac{b^2}{2} \leq \frac{b\ell_j}{2} \xrightarrow{j\to\infty} 0$$
   dann $b \xrightarrow{j\to\infty} 0$. $\square$



4. Für $\Omega \subseteq \left(-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right) \times (0, \ell)$ konvex und $f \in H^1(\Omega)$ mit $\int_\Omega f(x)\,dx = 0$ ist

$$\|f\|_{L^2(\Omega)} \leq \frac{\ell}{\pi} \|f\|_{H^1(\Omega)} + \varepsilon \left(2 + \frac{\ell}{\pi}\right) \|f\|_{W^{1,\infty}(\Omega)} |\Omega|^{1/2}$$

für $\ell_j := \operatorname{diam}(\Omega_j) \leq \operatorname{diam}(\Omega)$.

Beweis. O.B.d.A. sei $\ell = \operatorname{diam}(\Omega) = |a - b|$ für $a, b \in \partial\Omega$, $a = (0,0)$, $b = (\ell, 0)$ und

$$\Omega = \left\{(\xi, \eta) \in \mathbb{R}^2 \,\middle|\, 0 < \xi < \ell,\, \underline{p}(\xi) < \eta < \bar{p}(\xi)\right\}$$

für Funktionen $\underline{p}$, $\bar{p}$ die den Rand von Ω parametrisieren mit $\underline{p} \leq 0 \leq \bar{p}$, $(0, \ell) \times \{0\} \subseteq \Omega$. Dann ist

$$g(\xi) := f(\xi, 0)$$

für alle $0 < \xi < \ell$ definiert, die Funktion $\rho(x) := \bar{p}(x) - \underline{p}(x)$ ist konkav und mit

$$m := \frac{\int_0^\ell g\rho d\xi}{\int_0^\ell \rho d\xi}$$

gilt nach dem Satz 3.12

$$\left\|\rho^{1/2}(g - m)\right\|_{L^2(0,\ell)} \leq \frac{\ell}{\pi} \left\|\rho^{1/2} g'\right\|_{L^2(0,\ell)}.$$

Da f und Df Lipschitz-stetig sind, gilt

$$\begin{aligned}
\|f - g\|_{L^2(\Omega)} &\leq \varepsilon |\Omega|^{1/2} \operatorname{Lip}(f), \\
\left\|\frac{\partial f}{\partial x_1} - g'\right\|_{L^2(\Omega)} &\leq \varepsilon |\Omega|^{1/2} \operatorname{Lip}(Df)
\end{aligned}$$

und daher auch

$$\begin{aligned}
\left\|\rho^{1/2} m\right\|_{L^2(0,\ell)} &= \frac{\left|\int_\Omega g(x_1)\,dx\right|}{\left\|\rho^{1/2}\right\|_{L^2(0,\ell)}} \\
&= \frac{\left|\int_\Omega f(x_1, 0) - f(x_1, x_2)\,dx\right|}{|\Omega|^{1/2}} \\
&\leq \varepsilon \operatorname{Lip}(f) |\Omega|^{1/2}.
\end{aligned}$$

Folglich gilt

$$\begin{aligned}
\|f\|_{L^2(\Omega)} &\leq \|f - g\|_{L^2(\Omega)} + \left\|\rho^{1/2}(g - m)\right\|_{L^2(0,\ell)} + \left\|\rho^{1/2} m\right\|_{L^2(0,\ell)} \\
&\leq 2\varepsilon |\Omega|^{1/2} \operatorname{Lip}(f) + \frac{\ell}{\pi} \left\|\rho^{1/2} g'\right\|_{L^2(0,\ell)} \\
&= 2\varepsilon |\Omega|^{1/2} \operatorname{Lip}(f) + \frac{\ell}{\pi} \|g'\|_{L^2(\Omega)} \\
&\leq 2\varepsilon |\Omega|^{1/2} \operatorname{Lip}(f) + \frac{\ell}{\pi} \left\|g' - \frac{\partial f}{\partial x_1}\right\|_{L^2(\Omega)} + \frac{\ell}{\pi} \left\|\frac{\partial f}{\partial x_1}\right\|_{L^2(\Omega)} \\
&\leq \left(2 + \frac{\ell}{\pi}\right) \varepsilon |\Omega|^{1/2} \|f\|_{W^{1,\infty}(\Omega)} + \frac{\ell}{\pi} \|Df\|_{L^2(\Omega)}.
\end{aligned}$$

□

5. Beweis von $c_P(\Omega) \leq \frac{1}{\pi} \operatorname{diam}(\Omega)$:
   Für $\ell_j := \operatorname{diam}(\Omega_j) \leq \operatorname{diam}(\Omega)$ folgt aus 4. für alle j

$$\|f\|_{L^2(\Omega_j)} \leq \frac{\ell_j}{\pi} \|f\|_{H^1(\Omega_j)} + \varepsilon \left(2 + \frac{\ell}{\pi}\right) \|f\|_{W^{1,\infty}(\Omega)} 2^{-j/2} |A|^{1/2}.$$

   Diese Aussage wird quadriert und summiert und liefert für $\varepsilon \to 0$ die Behauptung.

6. Beweis zur Optimalität: Auf dem Gebiet $\Omega_\varepsilon = (0,\ell) \times (0,\varepsilon)$ betrachte man

$$f(\xi,\eta) := \cos\left(\frac{\pi}{\ell}\xi\right),$$

   unabhängig von η. Wie im eindimensionalen Fall folgt

$$\|f\|_2 = \frac{\ell}{\pi} \|Df\|_2.$$

   Also

$$\frac{\ell}{\pi} \leq c_P(\Omega_\varepsilon) \leq \frac{\sqrt{\ell^2 + \varepsilon^2}}{\pi}$$

   und insgesamt

$$\lim_{\varepsilon \searrow 0} \frac{c_P(\Omega_\varepsilon)}{\operatorname{diam}(\Omega_\varepsilon)} = \frac{1}{\pi}.$$

□

Bemerkung. In der Originalarbeit ist $n \geq 3$ unklar (betrachtete Flächenfunktionen sind nicht konkav). Aber: Die Aussage läßt sich verifizieren!

3.3 Interpolationsfehlerabschätzungen

Ziel. Interpolationsfehler $u - Iu$ für abstrakte Fehlerabschätzung

$$\|u - u_h\|_V \lesssim \min_{v_h \in V_h} \|u - v_h\|_V \lesssim \|u - Iu\|_V$$

abschätzen.

Notation. In diesem Abschnitt ist $\mathcal{T}$ eine reguläre Triangulierung mit der Kantenmenge $\mathcal{E}$ und Knotenmenge $\mathcal{N}$ von Ω in lauter Dreiecke.

Definition 3.20 (nodale Interpolation). In der P_1-FEM bezeichne $(\varphi_z)_{z\in\mathcal{N}}$ die nodale Basis von $P_1(\mathcal{T}) \cap C_0^0(\bar{\Omega})$. Für $f \in C^0(\bar{\Omega})$ heißt

$$If := \sum_{z \in \mathcal{N}} f(z)\,\varphi_z \in P_1(\mathcal{T}) \cap C^0(\bar{\Omega})$$

nodaler Interpolant und $f - If$ heißt *Interpolationsfehler.*

Bemerkung. Für $n = 1, 2, 3$ gilt $H^2(\Omega) \hookrightarrow C^0(\bar{\Omega})$ ("Sobolev-Einbettung").

Bemerkung. Für $f \in H^2(\Omega)$ ($\Omega \subseteq \mathbb{R}^n$, $n \geq 2$) ist auf $E \in \mathcal{E}$

$$(f - If)|_E \in H_0^1(E)$$

und

$$\left.\frac{\partial}{\partial s}(f - If)\right|_E \in L^2(E) \quad \text{mit} \quad \fint_E \frac{\partial}{\partial s}(f - If)\,ds = 0.$$

Satz 3.21 (lokaler Approximationsfehler). *Für ein Dreieck $T = \operatorname{conv}\{P, E\}$ mit Kante E und gegenüberliegendem Knoten P und eine Funktion $g \in H^1(T)$ mit $\fint_E g ds = 0$ ist*

$$\|g\|_{L^2(T)} \leq h_T \|Dg\|_{L^2(T)} \sqrt{\pi^{-2} + \frac{1}{8}}.$$

Lemma 3.22 (Spuridentität). *Für $f \in W^{1,1}(T)$ gilt*

$$\fint_E f ds - \fint_T f dx = \frac{1}{n} \fint_T (x - P) Df\, dx.$$

Beweis von Satz 3.21. Nach Lemma 3.22 ist wegen $g_E := \fint_E g ds = 0$

$$g_T := \fint_T g dx = \frac{1}{2} \fint_T (P - x) Dg\, dx.$$

Dann ist

$$\begin{aligned}
\|g\|^2_{L^2(T)} &= \|g - g_T\|^2_{L^2(T)} + \|g_T\|^2_{L^2(T)} && \text{(nach Pythagoras)} \\
&\leq \left(\frac{h_T}{\pi}\right)^2 |g|^2_{W^{1,2}(T)} + \frac{1}{4} |T|^{-1} \left| \int_T (P - x) Dg\, dx \right|^2 && \text{(nach Poincaré, (3.1))} \\
&\leq \left(\left(\frac{h_T}{\pi}\right)^2 + \frac{1}{4} |T|^{-1} \|P - \cdot\|^2_{L^2(T)} \right) |g|^2_{W^{1,2}(T)}
\end{aligned}$$

und weiter

$$\|P - \cdot\|^2_{L^2(T)} = \int_0^\omega \int_0^{R(\varphi)} r^2 r\, dr\, d\varphi = \int_0^\omega \frac{R(\varphi)^4}{4} d\varphi \leq \frac{h_T^2}{2} \underbrace{\int_0^\omega \frac{1}{2} R(\varphi)^2\, d\varphi}_{|T|}. \qquad \square$$

Satz 3.23 (lokale Interpolationsfehlerabschätzung). *Für $f \in H^2(T)$ für ein Dreieck $T =$ $\operatorname{conv}\{P_1, P_2, P_3\}$ mit einem Innenwinkel $0 < \omega < \pi$ gilt:*

$$(A) \qquad |f - If|_{H^1(T)} \leq \sqrt{\frac{1}{4} + \frac{2}{\pi^2}} \frac{h_T}{\sqrt{1 - |\cos\omega|}} |f|_{H^2(T)}$$

$$(B) \qquad \|f - If\|_{L^2(T)} \leq \frac{3h_T^2}{\sqrt{1 - |\cos\omega|}} |f|_{H^2(T)} \qquad \textit{(Ist die 3 scharf?)}$$

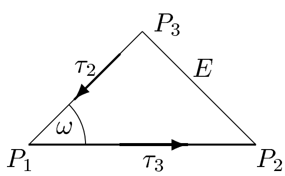
Bemerkung. Bezeichne mit $h_{\mathcal{T}}$ die Netzweite. Für $s = 0, 1$ ist

$$|f - If|_{H^s(T)} \lesssim \left\| h_{\mathcal{T}}^{2-s} D^2 f \right\|_{L^2(\Omega)}.$$

Beweis. O.B.d.A. sei $E = E_1 = \operatorname{conv}\{P_2, P_3\}$, $\cos\omega = -\tau_2 \cdot \tau_3 =: -\gamma$ und $|\tau_2| = |\tau_3| = 1$.

Behauptung: Für alle $a \in \mathbb{R}^2$ ist

$$|a|^2 \leq \frac{(a \cdot \tau_2)^2 + (a \cdot \tau_3)^2}{1 - |\cos\omega|},$$



denn: sei $a = a_2\tau_2 + a_3\tau_3$ für $a_2, a_3 \in \mathbb{R}$.

O.B.d.A. ist $a_2^2 + a_3^2 = 1$. Nach Young-Ungleichung folgt dann $-1 \leq 2a_2a_3 \leq 1$. Dann

$$\begin{aligned}
0 &\leq \underbrace{\gamma^2 - 2a_2a_3\gamma\,|\gamma|}_{\geq 0} + \underbrace{|\gamma| + 2a_2a_3\gamma}_{\geq 0} \\
\Leftrightarrow \underbrace{(2a_2a_3\gamma + 1)}_{=|a|^2}\underbrace{(1 - |\gamma|)}_{>0} &\leq 1 + \gamma^2 + 4a_2a_3\gamma \\
&= (a \cdot \tau_2)^2 + (a \cdot \tau_3)^2 \\
&= (a_2 + a_3\gamma)^2 + (a_3 + a_2\gamma)^2 .
\end{aligned}$$

Zum **Beweis von (A)**: Die Funktion

$$g_j := D\,(f - If) \cdot \tau_j, \qquad j = 2, 3,$$

erfüllt $\int_{E_j} g_j ds = 0$. Dann folgt mit Satz 3.21

$$\|g_j\|_{L^2(T)}^2 \leq \left(\frac{1}{\pi^2} + \frac{1}{8}\right) h_T^2 \,\|Dg_j\|_{L^2(T)}^2 \leq \left(\frac{1}{\pi^2} + \frac{1}{8}\right) h_T^2 \,|f|_{H^2(T)}^2$$

(mit $D\,(f - If) \cdot \tau_j = \frac{\partial}{\partial s_j}(f - If)$).

Die obige Behauptung liefert nun

$$\|D\,(f - If)\|_2^2 \leq \frac{\|g_2\|_2^2 + \|g_3\|_2^2}{1 - |\cos\omega|} \leq \frac{h_T^2\left(\frac{2}{\pi^2} + \frac{1}{4}\right)}{1 - |\cos\omega|}\,|f|_{H^2(T)}^2 .$$

Aufgabe: Man bestimme die Konstanten in (A) und (B) für T_{ref}.

Aufgabe: Sei $g := f - If \in H^2(T)$ und $g|_E \in H_0^1(E)$, $g_T := ⨍_T g dx$. Dann ist

$$\|g_T\|_{L^2(T)} \leq \frac{3}{2} h_T \,|g|_{H^1(T)} + \frac{h_T^2}{2}\,|g|_{H^2(T)} . \tag{3.2}$$

Beweis von (B):

$$\begin{aligned}
\|f - If\|_{L^2(T)}^2 &= \|g\|_{L^2(T)}^2 \\
&= \underbrace{\|g_T\|_{L^2(T)}^2}_{(3.2)} + \underbrace{\|g - g_T\|_{L^2(T)}}_{\leq \left(\frac{h_T}{\pi}\right)^2 |g|_{H^1(T)}^2} \\
&\leq \frac{h_T^4}{2}\,|f|_{H^2(T)}^2 + h_T^2\left(\frac{9}{2} + \frac{1}{\pi^2}\right)|g|_{H^1(T)}^2 \\
&\leq \frac{h_T^4}{2}\,|f|_{H^2(T)}^2 + h_T^2\left(\frac{9}{2} + \frac{1}{\pi^2}\right)\left(\frac{1}{4} + \frac{2}{\pi^2}\right)\frac{h_T^2}{1 - |\gamma|}\,|f|_{H^2(T)}^2
\end{aligned}$$

Nebenrechnung:

$$\frac{1}{2} + \left(\frac{9}{2} + \frac{1}{\pi^2}\right)\left(\frac{1}{4} + \frac{2}{\pi^2}\right)\left(\frac{1}{1 - |\gamma|}\right) \leq \frac{1}{1 - |\gamma|}\left(\frac{1}{2} + \frac{9}{8} + \frac{9}{\pi^2} + \frac{1}{4\pi^2} + \frac{2}{\pi^4}\right) \leq \frac{2.59}{1 - |\gamma|} \qquad \square$$

Definition 3.24 (Maximumwinkelbedingung). Sofern die größten Winkel in jedem Dreieck einer Familie von Triangulierungen $(\mathcal{T}_h)_{h\in H}$ durch eine Konstante $c_\sphericalangle < \pi$ nach oben beschränkt sind, sagt man: $(\mathcal{T}_h)_{h\in H}$ erfüllt die *Maximumwinkelbedingung*.

Bemerkung 3.25. Unter der Maximumwinkelbedingung kann stets ω so gewählt werden, daß im Approximationssatz

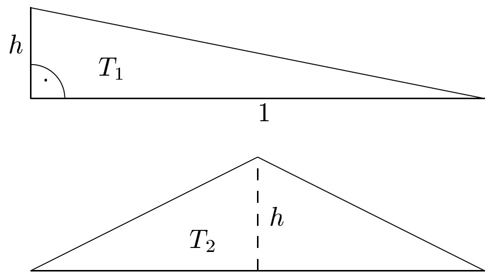
$$(1-|\cos\omega|)^{-1} \le (1-|\cos c_{\sphericalangle}|)^{-1} \lesssim 1$$

Bemerkung 3.26 (Minimumwinkelbedingung). Der kleinste Winkel ist durch eine Konstante $c_{\min\sphericalangle} > 0$ beschränkt.

Es ist klar, daß aus der Minimumwinkelbedingung die Maximumwinkelbedingung folgt. Die Äquivalenz gilt hingegen nicht, wie folgendes Beispiel zeigt.

Beispiel. $f(x,y) = 1-x^2$ für $(x,y) \in T_1 := \operatorname{conv}\{(0,1),(0,1),(h,0)\}$, $\omega = \pi/2$.

Interpolationsfehlerabschätzung: Konstante, unabhängig von $h \searrow 0$.
T_1 erfüllt die Maximumwinkelbedingung, aber für $h \searrow 0$ <u>nicht</u> die Minimumwinkelbedingung.



Auf $T_2 := \operatorname{conv}\{(1,0),(0,h),(-1,0)\}$ wird für $h \searrow 0$ die Maximumwinkelbedingung verletzt.

$$\frac{\partial(f-If)}{\partial y}(x,y) = -\frac{\partial If}{\partial y} = -h^{-1} \xrightarrow{h\searrow 0} \infty$$

3.4 Modellproblem und Regularität

Betrachte ein polygonal berandetes Lipschitz-Gebiet $\Omega \subseteq \mathbb{R}^2$ und Funktionen $f \in L^2(\Omega)$, $u \in H_0^1(\Omega)$ mit $f + \Delta u = 0$ in Ω.

Satz 3.27. *(a) Für $\omega \Subset \Omega$ offen gilt $u|_\omega \in H^2(\omega)$ und es existiert $c(\omega,\Omega) > 0$ mit*

$$\|u\|_{H^2(\omega)} \le c(\omega,\Omega)\left(\|f\|_{L^2(\Omega)} + \|u\|_{H^1(\Omega)}\right).$$

(b) Für Ω konvex gilt $u \in H^2(\Omega)$ und

$$\|u\|^2_{H^2(\Omega)} \le \|f\|^2_{L^2(\Omega)} + \|u\|^2_{H^1(\Omega)}.$$

(Die Konstante ist “1”.)

Bemerkung. Für glatte Ränder, Ω C^2-Gebiet, gilt ebenfalls globale Regularität: $u \in H^2(\Omega)$ und $\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$.

Beweisskizze.

Proposition 3.28. *Es existiert ein $c(\omega,\Omega)$, so daß für alle $v \in H^2(\Omega)$*

$$\|v\|_{2,2,\omega} \le c(\omega,\Omega)\left(\|\Delta v\|_{2,\operatorname{supp}\chi} + \|v\|_{1,2,\Omega}\right)$$

für $\chi \in C_c^\infty(\Omega)$ mit $0 \le \chi \le 1$ auf $\mathbb{R}^2$ und $\chi|_\omega \equiv 1$.

Beweis. Für $w := v\chi$ gilt

$$\|\Delta w\|_{L^2(\Omega)} = \left(\int_\Omega (w_{xx}+w_{yy})^2\, d\Omega\right)^{1/2} \le \left(\int_\Omega (\chi\Delta v)^2\, dx\right)^{1/2} + 4\|\chi\|_{2,\infty,\Omega}\|v\|_{1,2,\Omega},$$

da $|w_{xx}| = |(v\chi)_{xx}| \le |v_{xx}|\,\|\chi\|_\infty + 2|v_x|\,\|\chi_x\|_\infty + v\,\|\chi_{xx}\|_\infty$.

Für $\omega \in H_0^3(\Omega)$ folgte

$$\int_\Omega w_{xx}w_{yy}d\Omega = -\int_\Omega w_{xyx}w_y d\Omega = \int_\Omega w_{xy}^2 d\Omega = \|w_{xy}\|_{2,\Omega}^2 \qquad \text{für alle } v \in H^3(\Omega).$$

Da H^3 dicht in H^2 folgt dieses sogar für alle $v \in H^2(\Omega)$.

Also gilt

$$\begin{aligned}
\left\|D^2 v\right\|_{2,\omega}^2 &\leq \left\|D^2 w\right\|_2^2 = \|w_{xx}\|_{2,\Omega}^2 + \|w_{yy}\|_{2,\Omega}^2 + 2\|w_{xy}\|_{2,\Omega}^2 \\
&= \|\Delta w\|_{2,\Omega}^2 \\
&\leq 2\left\|\underbrace{\chi}_{\leq 1}\Delta v\right\|_2^2 + 8\|\chi\|_{\infty,2,\Omega}^2 \|v\|_{1,2,\Omega}^2
\end{aligned}$$

und das ist die Behauptung. □

Beweis zu *a)*:

Zu $u \in H_0^1(\Omega)$ reguliert man $u\chi$ zu $w_\varepsilon := (u\chi) * \eta_\varepsilon$ für $0 < \varepsilon < \operatorname{dist}(\omega, \partial\Omega)$. Dann ist $w_\varepsilon \in \mathcal{D}(\Omega)$ mit

$$\|f + \Delta w_\varepsilon\|_{2,\Omega} + \|u - w_\varepsilon\|_{1,2,\Omega} \xrightarrow{\varepsilon \searrow 0} 0.$$

Mit Proposition 3.28 für $v = \omega_\varepsilon$ und $\varepsilon \searrow 0$ folgt

$$\limsup_{\varepsilon \searrow 0} \|w_\varepsilon\|_{2,2,\omega} \leq c(\omega, \Omega)\left(\|f\|_{2,\Omega} + \|u\|_{1,2,\Omega}\right).$$

Die Folge (w_ε) ist beschränkt im Hilbert-Raum $H^2(\omega)$ und besitzt eine schwach konvergente Teilfolge $(w_\varepsilon) \rightharpoonup w$ in $H^2(\omega)$, also $w = u|_\omega$ und

$$\|w\|_{2,2,\omega} = \|u\|_{2,2,\omega} \leq c(\omega, \Omega)\left(\|f\|_{L^2(\Omega)} + \|u\|_{1,2,\Omega}\right).$$

Proposition 3.29. *Für alle glatten v mit $v|_{\partial\Omega} = 0$ und $\partial\Omega$ glatt und konvex gilt*

$$|v|_{2,2,\Omega} \leq \|\Delta v\|_{2,\Omega}.$$

Beweisidee. Sei $x(s)$ eine Randparametrisierung von $\partial\Omega$ bzgl. Bogenlängenparameter s.

Für einen Punkt $A \in \partial\Omega$ ist $\tau(A) = \frac{\partial x}{\partial s}(A) \perp \nu(A)$.

Konvexität: $\frac{\partial \tau(x(s))}{\partial s} \cdot \nu(x(s)) \leq 0$

Randbedingung $\frac{\partial v(x(s))}{\partial s} = 0$

Partielle Integration wie in Proposition 3.28 zeigt <u>mit</u> Randtermen

$$\begin{aligned}
\|\Delta v\|_{2,\Omega}^2 - |v|_{2,2,\Omega}^2 &= 2\int_{\partial\Omega}\left(v_{xx}v_y \underbrace{\nu_y}_{-\tau_x} - v_{xy}v_y \underbrace{\nu_x}_{\tau_y}\right)dx \\
&= -2\int_{\partial\Omega} v_y \frac{\partial v_x}{\partial s} ds \\
&= 2\int_{\partial\Omega} \frac{\partial v_y}{\partial s} v_x ds
\end{aligned}$$

Sei $v_\nu := \frac{\partial v}{\partial \nu} := \nabla v \cdot \nu$, also $Dv = v_\nu \cdot \nu$, also $v_x = v_\nu \nu_x$. Weiterhin folgt $\frac{\partial v_y}{\partial s} = \frac{\partial (v_\nu \nu_y)}{\partial s} = \frac{\partial v_\nu}{\partial s}\nu_y + v_\nu \frac{\partial \nu_y}{\partial s}$.

Aus $0 = v_\tau := \frac{\partial v}{\partial \tau} := \nabla v \cdot \tau = \nabla v \cdot (-\nu_y, \nu_x)^T$ folgt $v_x \nu_y = v_y \nu_x$.

Dies ergibt dann weiter

$$
\begin{aligned}
\|\Delta v\|_{2,\Omega}^2 - |v|_{2,2,\Omega}^2 &= 2\int_{\partial\Omega} v_\nu \nu_x \left(\frac{\partial v_\nu}{\partial s}\nu_y + v_\nu \frac{\partial \nu_y}{\partial s}\right) \\
&= 2\int_{\partial\Omega} \frac{1}{2} v_\nu^2 \left(-\frac{\partial}{\partial s}(\nu_x \nu_y)\right) + v_\nu^2 \nu_x \frac{\partial \nu_y}{\partial s} dx \quad \text{(part. Integration für } \frac{\partial}{\partial s}\left(\frac{1}{2}v_\nu\right)^2 \\
&= \int_{\partial\Omega} v_\nu^2 \left(\nu_x \underbrace{\frac{\partial \nu_y}{\partial s}}_{-\frac{\partial \tau_x}{\partial s}} - v_y \underbrace{\frac{\partial \nu_x}{\partial s}}_{\frac{\partial \tau_y}{\partial s}} \right) ds \\
&= -\int_{\partial\Omega} v_\nu^2 \underbrace{\left(\frac{\partial \tau}{\partial s} \cdot \nu\right)}_{\leq 0} ds \\
&\geq 0.
\end{aligned}
$$

□

Beweis von *b)*: Dichtigkeitsargumente für u und konvexe Approximation von Ω durch glatte Gebiete. □

Bemerkung. In der Umgebung $\omega = B(x,\delta) \cap \Omega$ eines Punktes x bei dem $\partial\Omega$ glatt ist, gilt $u \in H^2(\omega)$. Bei konvexen Eckpunkten auch, aber bei stumpfen Ecken nicht.

Definition 3.30 ($\mathcal{H}^2$-regulär). Ein Gebiet Ω (bzw. ein Problem $-\Delta u = f$ in Ω mit $u \in H_0^1(\Omega)$) heißt *$\mathcal{H}^2$-regulär*, falls eine Konstante $c(\Omega) > 0$ existiert, so daß für alle $f \in L^2(\Omega)$ genau ein $u \in H_0^1(\Omega) \cap H^2(\Omega)$ existiert, so daß

$$
f = -\Delta u \quad \text{und} \quad \|u\|_{H^2(\Omega)} \leq c(\Omega)\|f\|_{L^2(\Omega)}.
$$

Folgerung 3.31. Für H^2-reguläre Probleme und deren Finite Elemente-Approximation gilt

$$
\|u - u_h\|_{H^1(\Omega)} \lesssim \left\|h_\mathcal{T} D^2 u\right\|_{L^2(\Omega)} \lesssim \|h_\mathcal{T}\|_\infty \|f\|_{L^2(\Omega)},
$$

also lineare Konvergenz für quasi-uniforme Netze (d.h. für $\|h_\mathcal{T}\|_\infty \|1/h_\mathcal{T}\|_\infty \lesssim 1$).

Satz 3.32. *Es sei $\Omega \subseteq \mathbb{R}^2$ ein Lipschitz-Gebiet mit polygonalem Rand Γ, das durch die Ecken $P_1, \ldots, P_J$ definiert ist vermöge*

$$
\Gamma = \underbrace{\operatorname{conv}\{P_1,P_2\}}_{\Gamma_1} \cup \underbrace{\operatorname{conv}\{P_2,P_3\}}_{\Gamma_2} \cup \ldots \cup \underbrace{\operatorname{conv}\{P_{J-1},P_J\}}_{\Gamma_{J-1}} \cup \underbrace{\operatorname{conv}\{P_J,P_1\}}_{\Gamma_J}.
$$

Die Randbedingungen wechseln höchstens bei $P_1, \ldots, P_J$, d.h. Γ_j ist entweder ganz Dirichlet-Rand, $u|_{\Gamma_j} = 0$, oder Neumann-Rand, $\frac{\partial u}{\partial \nu}\big|_{\Gamma_j} \in H^{s-1/2}(\Gamma_j)$, $s > 0$. Die Funktion $u \in H_D^1(\Omega)$ erfülle

$$
-\Delta u = f \in H^{s-1}(\Omega).
$$

Dann gilt $u = u_r + u_s$ mit $u_r \in H^{1+s}(\Omega)$ und

$$
u_s = \sum_{j=1}^{J} \sum_{\substack{k=1 \\ a_{jk}<s}}^{\infty} a_{jk}\, S_{jk}\, \chi_j \in H^{1+s}(\Omega),
$$

mit Koeffizienten a_{jk}, *Singulärfunktionen*

$$S_{jk}(r,\varphi) := \begin{cases} r^{\alpha_{jk}} \sin(\varphi_j + \alpha_{jk}\varphi), & \textit{falls } \alpha_{jk} \notin \mathbb{N} \textit{ und} \\ & \varphi = \omega_{jk} \textit{ zu } \Gamma_D, \\ r^{\alpha_{jk}} \cos(\varphi_j + \alpha_{jk}\varphi), & \textit{falls } \alpha_{jk} \notin \mathbb{N} \textit{ und} \\ & \varphi = \omega_{jk} \textit{ zu } \Gamma_N, \\ r_j^{\alpha_{jk}} (\log r \sin(\varphi_j + \alpha_{jk}\varphi) + \varphi \cos(\varphi_j + \alpha_{jk}\varphi)), & \textit{falls } \alpha_{jk} \in \mathbb{N} \textit{ und} \\ & \varphi = \omega_{jk} \textit{ zu } \Gamma_D, \\ r_j^{\alpha_{jk}} (\log r \cos(\varphi_j + \alpha_{jk}\varphi) - \varphi \sin(\varphi_j + \alpha_{jk}\varphi)), & \textit{falls } \alpha_{jk} \in \mathbb{N} \textit{ und} \\ & \varphi = \omega_{jk} \textit{ zu } \Gamma_N \end{cases}$$

und $\chi_j \in \mathcal{D}(\mathbb{R}^2)$ *sind Abschneidefunktionen, die auf einer Umgebung von* P_j *identisch* 1 *sind.*

Beispiel. $u = u_r + u_s$ mit $u_r \in H^2(\Omega)$ und

$$u_s(v,\varphi) = v^{2/3} \sin\left(\frac{3}{2}\varphi\right)$$

auf Lipschitz-Gebiet. $u \notin H^2(\Omega)$ für $-\Delta u \equiv 1$ in Ω.

Die empirische Konvergenzrate $2/3$ für ein uniformes Netz ist <u>nicht</u> optimal!

Heuristik: $n^{2/3}$ wird durch lineare Funktion approximiert, mit Fehler

$$\begin{aligned} |u_s - a \cdot x|^2_{H^1(T)} \gtrsim \left\| v^{-1/3} - a \right\|^2_{L^2(T)} &\approx \int_0^\omega \int_0^{h_T} \left(v^{-1/3}\right)^2 v \, dv \, d\varphi \\ &\approx h_T^{4/3} = \text{Gesamtfehler}^2! \end{aligned}$$

Bemerkung (Graduierte Gitter). $\xi_j := \left(\frac{j}{N}\right)^\beta$, $j = 0,1,2,\ldots,N$, z.B. $\beta = 3/2$ für ein Lipschitzgebiet.

$h_T = \left(\frac{1}{N}\right)^\beta = N^{-3/2}$ liefert auf T mit $0 \in T$ Fehler von

$$h_T^{2/3} = \frac{1}{N} \to 0,$$

d.h. lineare Konvergenz bzgl. N!

Geeignete graduierte Gitter zeigen optimale Konvergenzrate!

Index