Zusammenfassung Numerik von PDEs

Tim Baumann, http://timbaumann.info/uni-spicker

Dies ist ein verkürztes Skript zur gleichnamigen Vorlesung von Frau Prof. Dr. Tatjana Stykel an der Universität Augsburg im WS 15/16.

Def. Sei $\Omega \subseteq \mathbb{R}^n$ offen. Eine DGL der Form

$$F(x, u, \mathcal{D}u, \dots, \mathcal{D}^k u) = 0$$

heißt **partielle DGL/PDE** der Ordnung $k \geq 1$, wobei

$$F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \ldots \times \mathbb{R}^{n^k} \to \mathbb{R}$$

eine gegebene Funktion und $u : \Omega \to \mathbb{R}$ gesucht ist.

Def (**Klassifikation von PDEs**).

- Die PDE heißt **linear**, wenn sie die Form

$$\sum_{|\alpha| \leq k} a_\alpha(x) \mathcal{D}^\alpha u = f(x)$$

  mit Funktionen a_α, $f : \Omega \to \mathbb{R}$ besitzt.
- Die PDE heißt **semilinear**, wenn sie die Form

$$\sum_{|\alpha|=k} a_\alpha(x) \mathcal{D}^\alpha u + a_0(x, u, \mathcal{D}u, \dots, \mathcal{D}^{k-1}u) = 0$$

  besitzt, wobei $a_\alpha : \Omega \to \mathbb{R}$ und $a_0 : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \ldots \times \mathbb{R}^{n^{k-1}} \to \mathbb{R}$ gegeben sind.
- Die PDE heißt **quasilinear**, wenn sie die Form

$$\sum_{|\alpha|=k} a_\alpha(x, u, \mathcal{D}u, \dots, \mathcal{D}^{k-1}u) \mathcal{D}^\alpha u + a_0(x, u, \mathcal{D}_u, \dots, \mathcal{D}^{k-1}u) = 0$$

  hat, wobei $a_\alpha, a_0 : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \ldots \times \mathbb{R}^{n^k}$ gegeben sind.
- Die PDE heißt **nichtlinear**, falls die Ableitungen der höchsten Ordnung nicht linear vorkommen.

Def. Sei $\Omega \subseteq \mathbb{R}^n$ offen und $F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \to \mathbb{R}$ gegeben. Eine **PDE zweiter Ordnung** ist eine PDE der Form

$$F(x, u, \partial_{x_1} u, \dots, \partial_{x_n} u, \partial_{x_1}\partial_{x_1} u, \dots, \partial_{x_1}\partial_{x_n} u, \dots, \partial_{x_n}\partial_{x_n} u) = 0.$$

Notation. Für eine PDE 2. Ordnung sei $p_i := \partial_{x_i} u$, $p_{ij} := \partial^2_{x_i x_j} u$,

$$M(x) := \begin{pmatrix} \frac{\partial F}{\partial p_{11}} & \cdots & \frac{\partial F}{\partial p_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial p_{n1}} & \cdots & \frac{\partial F}{\partial p_{nn}} \end{pmatrix} = M(x)^T.$$

Def (**Typeneinteilung für PDEs der 2. Ordnung**). Obige PDE zweiter Ordnung heißt

- **elliptisch** in x, falls die Matrix $M(x)$ positiv o. negativ definit ist.
- **parabolisch** in x, falls genau ein EW von $M(x)$ gleich null ist und alle anderen EWe dasselbe Vorzeichen haben.
- **hyperbolisch** in x, falls genau ein EW von $M(x)$ ein anderes Vorzeichen als die anderen EWe hat.

Lösungstheorie elliptischer PDEs

Def. Sei $\Omega \subset \mathbb{R}^n$ offen, zusammenhängend und beschränkt.

- $\mathcal{C}(\overline{\Omega}, \mathbb{R}^m) := \{u : \overline{\Omega} \to \mathbb{R}^m \mid u \text{ stetig}\}$, $\mathcal{C}(\overline{\Omega}) := \mathcal{C}(\overline{\Omega}, \mathbb{R})$, mit Norm

$$\|u\|_{\mathcal{C}(\overline{\Omega}, \mathbb{R}^m)} := \sup_{x \in \overline{\Omega}} \|u(x)\|. \qquad (\textit{Supremumsnorm})$$

- $\mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m)$, $k \in \mathbb{N}$ ist der Raum aller auf Ω k-mal stetig diff'baren Funktionen $u : \Omega \to \mathbb{R}^m$, die zusammen mit ihren Ableitungen bis zur Ordnung k stetig auf $\overline{\Omega}$ fortgesetzt werden können mit Norm

$$\|u\|_{\mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m)} := \sum_{|\alpha| \leq k} \|\mathcal{D}^\alpha u\|_{\mathcal{C}(\overline{\Omega}, \mathbb{R}^m)}.$$

- Für $\alpha \in (0, 1]$ ist $\mathcal{C}^{0,\alpha}(\overline{\Omega}, \mathbb{R}^m) := \{u \in \mathcal{C}(\overline{\Omega}, \mathbb{R}^m) \mid H_\alpha(u, \overline{\Omega}) < \infty\}$ mit

$$H_\alpha(u, \overline{\Omega}) := \sup_{x,y \in \overline{\Omega}, x \neq y} \frac{\|u(x) - u(y)\|}{\|x - y\|^\alpha} \qquad (\textit{Hölder-Koeffizient})$$

  der **Raum der glm. Hölder-stetigen Fktn** zum Exponent α. Der Hölder-Koeffizient ist dabei eine Seminorm auf $\mathcal{C}^{0,\alpha}(\overline{\Omega}, \mathbb{R}^m)$.
- $\mathcal{C}^{k,\alpha}(\overline{\Omega}, \mathbb{R}^m) := \{u \in \mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m) \mid \forall |\gamma| = k : \mathcal{D}^\gamma u \in \mathcal{C}^{0,\alpha}(\overline{\Omega}, \mathbb{R}^m)\}$ heißt **Hölder-Raum**. Eine Norm ist gegeben durch

$$\|u\|_{\mathcal{C}^{k,\alpha}(\overline{\Omega}, \mathbb{R}^m)} := \|u\|_{\mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m)} + \sum_{|\gamma|=k} H_\alpha(\mathcal{D}^\gamma u, \overline{\Omega}).$$

Bem.
- Jede Hölder-stetige Funktion ist gleichmäßig stetig.
- $\mathcal{C}^{0,1}(\overline{\Omega}, \mathbb{R}^m)$ heißt **Raum der Lipschitz-stetigen Funktionen**.
- $\mathcal{C}$, $\mathcal{C}^k$ und $\mathcal{C}^{k,\alpha}$ sind Banach-Räume mit den jeweiligen Normen.

Def. Sei $\Omega \subset \mathbb{R}^n$ offen, zusammenhängend und beschränkt. Das Gebiet Ω gehört zur **Klasse $\mathcal{C}^{k,\alpha}$**, wenn in jedem Punkt $x \in \partial\Omega$ eine Umgebung in $\partial\Omega$ existiert, die sich in einem geeigneten Koordinatensystem als ein Graph einer Funktion aus $\mathcal{C}^{k,\alpha}$ darstellen lässt und Ω lokal immer auf einer Seite von $\partial\Omega$ liegt.

Satz (**Gauß'scher Integralsatz**). Sei $\Omega \subset \mathbb{R}^n$ ein Lipschitz-Gebiet und $u \in \mathcal{C}(\overline{\Omega}, \mathbb{R}^n) \cap \mathcal{C}^1(\Omega, \mathbb{R}^n)$. Dann gilt

$$\int_\Omega \operatorname{div} u \, \mathrm{d}x = \int_\Omega \sum_{i=1}^n \frac{\partial u_i}{\partial x_i} \, \mathrm{d}x = \int_{\partial\Omega} \sum_{i=1}^n u_i \nu_i \, \mathrm{d}\rho(x) = \int_{\partial\Omega} u \cdot \nu \, \mathrm{d}\rho(x),$$

wobei ν der äußere Normalenvektor an an den Rand von Ω ist.

Problem. Wir betrachten das Randwertproblem

$$\text{(RWP)} \quad \begin{cases} \mathcal{L}u = f & \text{in } \Omega \quad \text{(PDE)} \\ \mathcal{R}u = g & \text{auf } \partial\Omega \quad \text{(Randbedingung)} \end{cases}$$

wobei $\mathcal{L}$ der lineare Differentialoperator

$$\mathcal{L}u = -\sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial u}{\partial x_i} + c(x) u$$

mit Fktn $a_{ij}, b_i, c, f : \Omega \to \mathbb{R}$, $g : \partial\Omega \to \mathbb{R}$ ist, sodass $A(x) := (a_{ij}(x))$ symmetrisch ist. Als Randbedingung (RB) verlangen wir:

$$\begin{aligned} &\textbf{Dirichlet-RB:} & u &= g & &\text{auf } \partial\Omega, \\ &\textbf{Neumann-RB:} & (A(x)\nabla u) \cdot \nu &= g & &\text{auf } \partial\Omega \text{ oder} \\ &\textbf{Robin-RB:} & (A(x)\nabla u) \cdot \nu + \delta u &= g & &\text{auf } \partial\Omega. \end{aligned}$$

Bemn.
- Man kann auch auf verschiedenen Teilstücken des Randes verschiedene Bedingungen stellen.
- Falls die Funktionen a_{ij} differenzierbar sind, so kann $\mathcal{L}$ in **Divergenzform** geschrieben werden:

$$\begin{aligned} \mathcal{L}u &= -\sum_{i,j=1}^n \frac{\partial}{\partial x_j} \left(a_{ij}(x) \frac{\partial u}{\partial x_i} \right) + \sum_{i=1}^n \underbrace{\left(\left(\sum_{j=1}^n \frac{\partial}{\partial x_j} a_{ij}(x) \right) + b_i(x) \right)}_{\tilde{b}(x) :=} \frac{\partial u}{\partial x_i} + c(x) u \\ &= -\operatorname{div}(A(x)\nabla u) + \tilde{b}(x) \cdot \nabla u + c(x) u \end{aligned}$$

Voraussetzung. Wir nehmen im Folgenden an:

- $\mathcal{L}$ ist **gleichmäßig elliptisch**, d. h.

$$\exists \lambda_0 > 0 : \forall \xi \in \mathbb{R}^n : \forall x \in \Omega : \xi^T A(x) \xi \geq \lambda_0 \|\xi\|^2.$$

  Dabei heißt λ_0 *Elliptizitätskonstante.*
- $a_{ij}, b_i, c, f \in \mathcal{C}(\overline{\Omega})$, $g \in \mathcal{C}(\partial\Omega)$

Bem. $\mathcal{L}$ ist elliptisch auf Ω $:\iff$ $A(x) > 0$ (spd) für alle $x \in \Omega$

Def. Eine Fkt $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ heißt **klassische Lsg** vom (RWP) mit $\mathcal{R}u := u$, wenn die beiden Gleichungen in (RWP) in jedem Punkt von Ω bzw. des Randes $\partial\Omega$ erfüllt sind.

Satz (**Maximumprinzip**). Sei $\Omega \subset \mathbb{R}^n$ offen, zshgd u. beschränkt. Sei $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ eine Lösung vom (RWP), $f \leq 0$ in Ω und $c \equiv 0$. Dann nimmt u sein Maximum auf dem Rand $\partial\Omega$ an, d. h.

$$\sup_{x \in \overline{\Omega}} u(x) = \sup_{x \in \partial\Omega} u(x) = \sup_{x \in \partial\Omega} g(x).$$

Kor. Sei $c \geq 0$ und $f \leq 0$. Dann gilt $\sup_{x \in \overline{\Omega}} u(x) \leq \max\{\sup_{x \in \partial\Omega} u(x), 0\}$.

Kor (**Vergleichsprinzip**). Für $u_1, u_2 \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ und $c \geq 0$ gelte $\mathcal{L}u_1 \leq \mathcal{L}u_2$ in Ω und $u_1 \leq u_2$ auf $\partial\Omega$. Dann gilt $u_1 \leq u_2$ auf $\overline{\Omega}$.

Kor (Eindeutigkeit). Sei $c \geq 0$. Dann hat (RWP) höchstens eine Lösung $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$.

Satz. Sei Ω ein beschr. Lipschitz-Gebiet, $a_{ij}, b_i, c \in \mathcal{C}(\overline{\Omega})$, $c \geq 0$, $\mathcal{L}$ glm. elliptisch, $f \in \mathcal{C}^{0,\alpha}(\overline{\Omega})$ für ein $\alpha \in (0, 1)$ und $g \in \mathcal{C}(\partial\Omega)$. Dann besitzt (RWP) genau eine Lsg $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$.

Achtung. Es muss aber nicht $u \in \mathcal{C}^2(\overline{\Omega})$ gelten!

Differenzenverfahren

Problem (**Poisson-Problem**).

$$(\mathrm{RWP}_1)\quad \begin{cases} -\Delta u = f & \text{in } \Omega = (0,1) \\ u(0) = g_0, u(1) = g_1 & \text{auf } \partial\Omega \end{cases}$$

Verfahren (**DV**). Wir führen folgende Schritte durch:

1. Diskretisierung: Wähle $n \in \mathbb{N}$, setze $h := \frac{1}{n}$ und
$$\Omega_h := \{x_i := ih \,|\, i = 1,\dots,n-1\} \qquad \text{(innere Gitterpunkte)}$$
$$\partial\Omega_h := \{x_0 = 0, x_n = 1\} \qquad \text{(Randpunkte)}$$

2. Approx. der Ableitungen durch Differenzenquotienten (DQ)
$$u'(x_i) \approx \tfrac{1}{h}(u(x_i+h) - u(x_i)) \qquad \text{(\textbf{Vorwärts-DQ})}$$
$$u'(x_i) \approx \tfrac{1}{h}(u(x_i) - u(x_i-h)) \qquad \text{(\textbf{Rückwärts-DQ})}$$
$$u'(x_i) \approx \tfrac{1}{2h}(u(x_i+h) - u(x_i-h)) \qquad \text{(\textbf{zentraler DQ})}$$
Für die zweite Ableitung ergibt sich
$$\begin{aligned} u''(x_i) = (u'(x_i))' &\approx \tfrac{1}{h}(u'(x_i+h) - u'(x_i)) \approx \\ &\approx \tfrac{1}{h}\cdot\left(\tfrac{1}{h}(u(x_i+h) - u(x_i)) - \tfrac{1}{h}(u(x_i) - u(x_i-h))\right) \\ &= \tfrac{1}{h^2}(u(x_i+h) - 2\cdot u(x_i) + u(x_i-h)) =: \Delta_h u \end{aligned}$$
Dabei heißt Δ_h der diskrete eindim. Laplace-Operator.
Das diskretisierte Randwertproblem ist nun
$$(\mathrm{RWP}_1)_\mathrm{h}\quad \begin{cases} -\Delta_h u_h = f & \text{in } \Omega_h, \\ u_h(0) = g_0, u_h(1) = g_1 & \text{auf } \partial\Omega_h. \end{cases}$$

3. Aufstellen des linearen Gleichungssystems
$$\tfrac{1}{h^2}(2u_h(x_1) - u_h(x_2)) = f(x_1) + \tfrac{g_0}{h^2} \qquad (i=1)$$
$$\tfrac{1}{h^2}(-u_h(x_{i-1}) + 2u_h(x_i) - u_h(x_{i+1})) = f(x_i) \qquad (i = 2,\dots,n-2)$$
$$\tfrac{1}{h^2}(-u_h(x_{n-2}) + 2u_h(x_{n-1})) = f(x_{n-1}) + \tfrac{g_1}{h^2} \qquad (i=n-1)$$
Als lineares Gleichungssystem: $-\tilde{\Delta}_h \tilde{u}_h = \tilde{f}_h$ mit
$$-\tilde{\Delta}_h = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & & & 0 \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ 0 & & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)},$$
$$\tilde{u}_h = \begin{pmatrix} u_h(x_1) \\ \vdots \\ u_h(x_{n-1}) \end{pmatrix}, \quad \tilde{f}_h = \begin{pmatrix} f(x_1) + \frac{g_0}{h^2} \\ f(x_2) \\ \vdots \\ f(x_{n-2}) \\ f(x_{n-1}) + \frac{g_1}{h^2} \end{pmatrix}$$

Konvergenz, Konsistenz und Stabilität des DV

Ziel. Herausfinden, was die Lösung u_h von $(\mathrm{RWP})_\mathrm{h}$ (die man durch Lösen von (LGS) erhält) mit der Lösung u zum ursprünglichen Problem (RWP) zu tun hat. Ist etwa u_h die Einschränkung von u, oder zumindest annäherungsweise? Wenn ja, wie klein muss man h wählen, damit die Approximation gut wird?

$$(\mathrm{RWP})\quad \begin{cases} -\mathcal{L}u = f & \text{in } \Omega, \\ u = g & \text{auf } \partial\Omega \end{cases}$$
$$(\mathrm{RWP})_\mathrm{h}\quad \begin{cases} -\mathcal{L}_h u = f_h & \text{in } \Omega_h, \\ u_h = g_h & \text{auf } \partial\Omega_h \end{cases}$$
$$(\mathrm{LGS})\quad \tilde{\mathcal{L}}_h \tilde{u}_h = \tilde{f}_h$$

Notation. $U_h := \{\Omega_h \to \mathbb{R}\}, \quad R_h : \mathcal{C}(\overline{\Omega}) \to U_h, \; u \mapsto u|_{\Omega_h}$

Def. Das Differenzenverfahren $(\mathrm{RWP})_\mathrm{h}$ heißt

- **konvergent** von der *Ordnung* p, falls $C > 0$, $h_0 > 0$ existieren, sodass für die Lsg u von (RWP) und die Lsg u_h von $(\mathrm{RWP})_\mathrm{h}$ gilt:
$$\|u_h - R_h u\|_h \le Ch^p \quad \text{für alle } 0 < h \le h_0,$$
wobei $\|-\|_h$ die Maximums-Norm ist, d. h. $\|u_h\|_h := \max\limits_{x\in\Omega_h} |u_h(x)|$.
- **konsistent** von der *Ordnung* p, falls
$$\|\mathcal{L}_h R_h u - R_h \mathcal{L} u\|_h \le ch^p \|u\|_{\mathcal{C}^{p+2}(\overline{\Omega})} \quad \forall u \in \mathcal{C}^{p+2}(\overline{\Omega}).$$
- **stabil**, falls $\tilde{\mathcal{L}}_h$ invertierbar ist und ein $h_0 > 0$ existiert mit
$$\sup_{0<h\le h_0} \|\tilde{\mathcal{L}}_h^{-1}\|_h < \infty, \quad \text{wobei } \|\tilde{\mathcal{L}}_h^{-1}\|_h := \sup_{f\neq 0} \frac{\|\tilde{\mathcal{L}}_h^{-1} f\|_h}{\|f\|_h}.$$

Bem. Die ind. Matrixnorm ist $\|\tilde{\mathcal{L}}_h^{-1}\|_h = \|\tilde{\mathcal{L}}_h^{-1}\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{j=1}^n |l_{ij}|$.

Satz. Ist das DV $(\mathrm{RWP})_\mathrm{h}$ konsistent und stabil, so auch konvergent. Genauer gilt: Ist $(\mathrm{RWP})_\mathrm{h}$ stabil und konsistent von der Ordnung p und $u \in \mathcal{C}^{p+2}(\overline{\Omega})$, dann ist $(\mathrm{RWP})_\mathrm{h}$ konvergent von der Ordnung p.

Beweis. Setze $w_h := u_h - R_h u$. Für $x \in \partial\Omega_h$ gilt dann $w_h(x) = 0$ und für $x \in \Omega_h$ gilt
$$\begin{aligned} \tilde{\mathcal{L}}_h w_h(x) = \mathcal{L}_h w_h(x) &= \mathcal{L}_h u_h(x) - \mathcal{L}_h R_h u(x) \\ &= f_h(x) - \mathcal{L}_h R_h u(x) = R_h f(x) - \mathcal{L}_h R_h u(x) \\ &= R_h \mathcal{L} u(x) - \mathcal{L}_h R_h u(x) \end{aligned}$$
Somit gilt $w_h = \tilde{\mathcal{L}}_h^{-1}(R_h\mathcal{L}u - \mathcal{L}_h R_h u)$ in Ω_h, also
$$\begin{aligned} \|w_h\|_h = \|\tilde{\mathcal{L}}_h^{-1}(R_h\mathcal{L}u - \mathcal{L}_h R_h u)\| &\le \|\tilde{\mathcal{L}}_h^{-1}\|_h \cdot \|R_h\mathcal{L}u - \mathcal{L}_h R_h u\|_h \\ &\le c_1\cdot c_2\cdot h^p \cdot \|u\|_{\mathcal{C}^{p+2}(\overline{\Omega})} \le Ch^p \qquad \text{für } 0 < h \le h_0. \end{aligned}$$
□

Lem. Das DV $(\mathrm{RWP}_1)_\mathrm{h}$ ist konsistent von der Ordnung 2. Es gilt
$$\|\Delta_h R_h u - R_h \Delta u\|_h \le \tfrac{1}{12}\|u\|_{\mathcal{C}^4(\overline{\Omega})} h^2 \quad \forall u \in \mathcal{C}^4(\overline{\Omega}).$$

Bem. Um zu zeigen, dass $(\mathrm{RWP}_1)_\mathrm{h}$ konvergent ist, müssen wir noch zeigen, dass $\tilde{L}_h = -\tilde{\Delta}_h$ invertierbar ist und dass $\sup\limits_{0<h\le h_0} \|\tilde{\Delta}_h\| < \infty$.

Def. Eine Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ heißt **M-Matrix**, falls

a) $a_{ii} > 0$ für $i = 1,\dots,n$, b) $a_{ij} \le 0$ für $i \neq j$, $i,j = 1,\dots,n$,
c) A invertierbar ist und d) für $A^{-1} =: B = (b_{ij})$ gilt $b_{ij} \ge 0$.

Lem. Erfülle $A \in \mathbb{R}^{n\times n}$ die Bedingungen a) und b). Zerlege $A = D + L + R$ in eine Diagonalmatrix und strikte untere/obere Dreiecksmatrizen. Dann ist A genau dann eine M-Matrix wenn
$$\rho(D^{-1}(L+R)) < 1.$$

Bem. Es gilt folgende Monotonie-Eigenschaft für M-Matrizen:
$$x \le y \implies A^{-1}x \le A^{-1}y.$$

Def. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt **reduzibel** (oder *zerlegbar*), wenn es eine Permutationsmatrix $P \in \mathbb{R}^{n\times n}$ gibt, sodass
$$PAP^T = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad \text{mit } A_{11} \in R^{k\times k},\ 0 < k < n.$$

Lem (Gerschgorin)**.** Alle EWe einer Matrix $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ liegen in der Menge
$$\bigcup_{i=1}^n \overline{B_{r_i}(a_{ii})} \quad \text{mit} \quad r_i := \sum_{j=1}^n |a_{ij}|.$$
Falls A irreduzibel ist, so liegen sie sogar in
$$\left(\bigcup_{i=1}^n B_{r_i}(a_{ii})\right) \cup \left(\bigcap_{i=1}^n \partial B_{r_i}(a_{ii})\right)$$

Def. Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ eine Matrix.

- A heißt **(schwach) diagonaldominant**, falls für $i = 1,\dots,n$ gilt:
$$\sum_{\substack{j=1 \\ j\neq i}}^n |a_{ij}| \;\underset{(\le)}{<}\; |a_{ii}|. \qquad (\ddagger)$$
- A heißt **irreduzibel diagonaldominant**, falls A irreduzibel und schwach diagonaldominant ist und für mind. ein $i \in \{1,\dots,n\}$ die strikte Ungleichung in $(\ddagger)$ gilt.

Lem. Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ eine Matrix mit $a_{ii} > 0$, $i = 1,\dots,n$ und $a_{ij} \le 0$, $i,j = 1,\dots,n$, $i \neq j$, die diagonaldominant oder irreduzibel diagonaldominant ist. Dann ist A eine M-Matrix.

Bem. $-\tilde{\Delta}_h$ ist irreduzibel diagonaldominant, also eine M-Matrix.

Lem. Sei A eine irreduzible M-Matrix. Dann gilt $A^{-1} > 0$.

Lem. Sei $A \in \mathbb{R}^{n\times n}$ eine M-Matrix und es existiere ein Vektor v, sodass $(Av)_j \ge 1$, $j = 1,\dots,n$. Dann gilt $\|A^{-1}\|_\infty \le \|v\|_\infty$.

Lem. $\|\tilde{\Delta}_h^{-1}\|_\infty \le \frac{1}{8}$

Satz. Das DV $(\mathrm{RWP}_1)_\mathrm{h}$ ist konvergent von der Ordnung 2, falls die Lösung von (RWP_1) zu $\mathcal{C}^4([0,1])$ gehört. Es gilt die Abschätzung
$$\|u_h - R_h u\|_\infty \le \frac{h^2}{96}\|u\|_{\mathcal{C}^4([0,1])}.$$

Differenzenverfahren in $(0,1)\times(0,1)\subset\mathbb{R}^2$

Problem. Wir betrachten nun

$$(\mathrm{RWP}_2)\quad \begin{cases} -\Delta u = f & \text{in } \Omega=(0,1)\times(0,1)\\ \quad\; u=g & \text{auf } \partial\Omega\end{cases}$$

Verfahren. 1. Diskretisierung: Setze $h := \frac{1}{n}$, $n\in\mathbb{N}$ und

$$\Omega_h := \{(x,y)\in\Omega \mid x=ih,\, y=jh,\, i,j=1,\dots,n-1\}$$
$$\partial\Omega_h := \{(x,y)\in\partial\Omega \mid x=ih,\, y=jh,\, i,j=1,\dots,n-1\}$$

2. Approximation der Ableitungen

$$\begin{aligned}-\Delta u(x,y) &= -\tfrac{\partial^2 u}{\partial x^2}(x,y) - \tfrac{\partial^2 u}{\partial y^2}(x,y)\\ &\approx -\tfrac{u(x+h,y)-2u(x,y)+u(x-h,y)}{h^2} - \tfrac{u(x,y+h)-2u(x,y)+u(x,y-h)}{h^2}\\ &= -\tfrac{u(x+h,y)+u(x-h,y)-4u(x,y)+u(x,y+h)+u(x,y-h)}{h^2} =: -\Delta_h u\end{aligned}$$

Dabei hat der diskrete Laplace-Operator Δ_h die Form eines *Differenzensterns*. Gesucht ist die Lsg $u_h : \Omega_h\cup\partial\Omega_h\to\mathbb{R}$ von

$$(\mathrm{RWP}_2)_\mathrm{h}\quad \begin{cases} -\Delta_h u_h = f_h & \text{in } \Omega_h\\ \quad\;\; u_h = g & \text{auf } \partial\Omega_h.\end{cases}$$

3. Aufstellen des linearen Gleichungssystems $-\tilde\Delta_h\tilde u_h = \tilde f_h$:

$$\tilde u_h = \begin{pmatrix} u_{11}\\ u_{12}\\ \vdots\\ u_{n-1,n-2}\\ u_{n-1,n-1}\end{pmatrix}\in\mathbb{R}^{(n-1)^2},\quad \tilde f_h = \ldots\in\mathbb{R}^{(n-1)^2},$$

$$-\tilde\Delta_h = \frac{1}{h^2}\begin{pmatrix} A & -I & & & 0\\ -I & A & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & A & -I\\ 0 & & & -I & A\end{pmatrix}\in\mathbb{R}^{(n-1)^2\times(n-1)^2},$$

$$A = \begin{pmatrix} 4 & -1 & & & 0\\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 4 & -1\\ 0 & & & -1 & 4\end{pmatrix}\in\mathbb{R}^{n-1\times n-1}$$

Lem. Das DV $(\mathrm{RWP}_2)_\mathrm{h}$ ist konsistent von der Ordnung 2. Es gilt

$$\|\Delta_h R_h u - R_h\Delta u\|_h \le \tfrac{1}{6}\|u\|_{\mathcal{C}^r(\overline{\Omega})}h^2.$$

Lem. Das DV $(\mathrm{RWP}_2)_\mathrm{h}$ ist stabil. Es gilt $\|\tilde D_h^{-1}\|_\infty \le 1/8$.

Satz. Das DV $(\mathrm{RWP}_2)_\mathrm{h}$ ist konvergent von der Ordnung 2, falls die Lösung von (RWP_2) zu $\mathcal{C}^4(\overline{\Omega})$ gehört. Es gilt

$$\|u_h - R_h u\|_h \le 1/48\|u\|$$

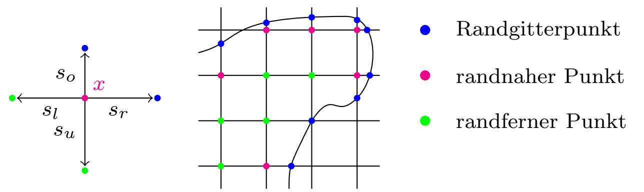
Bem. Ein *9-Punkte-Stern* bezieht weitere Gitterpunkte zur Approximation des Differentialoperators ein und erhöht die Konvergenzordnung auf 4:

$$\begin{aligned}-\Delta_h^{(9)}u(x,y) = \tfrac{1}{12h^2}\,\big(&u(x-2h,y) - 16u(x-h,y) + 30u(x,y)\\ &-16u(x+h,y) + u(x+2h,y) + u(x,y-2h) - 16u(x,y-h)\\ &+30u(x,y) - 16u(x,y+h) + u(x,y+2h)\big) \approx -\Delta u(x,y)\end{aligned}$$

Differenzenverfahren in allg. Gebieten $\Omega\subset\mathbb{R}^2$

Situation. Sei $\Omega\subset\mathbb{R}^2$ beschränkt.

Def.
- $\Omega_h := \{x,y\in\Omega \mid x/h,\, y/h\in\mathbb{Z}\}$ heißen **innere Gitterpkte**.
- Ein Punkt $z_R\in\partial\Omega$ heißt **Randgitterpunkt** (notiert $z_R\in\partial\Omega_h$), falls es einen inneren Gitterpunkt $z\in\Omega_h$ gibt, sodass $z_R = r+\alpha h e_1$ oder $z_R = z+\alpha h e_2$ mit $|\alpha|\le 1$. Die Nachbarn $N(x,y)$ eines Punktes (x,y) sind $(x+s_r h, y)$, $(x-s_l h, y)$, $(x, y+y_o h)$, $(x, y-s_u h)$, falls $s_r, s_l, s_o, s_u\in(0,1]$ und die Verbindungsstrecken zu (x,y) in Ω liegen.
- Ein Punkt $(x,y)\in\Omega_h$ heißt **randnah**, falls (x,y) die Nachbarn $(x-s_l h, y)$, $(x+s_r h, y)$, $(x, y-s_u h)$, $(x, y+s_o h)$ hat mit mindestens einem $s_i<1$. Ansonsten heißt (x,y) **randfern**.



Lem (**Dividierte Differenzen von Newton**).
Für $u\in\mathcal{C}^3([x_l,x_r])$, $x\in(x_l,x_r)$ gilt

$$u''(x) = \underbrace{\frac{2}{x_r-x_l}\left(\frac{u(x_r)-u(x)}{x_r-x} - \frac{u(x)-u(x_l)}{x-x_l}\right)}_{=\frac{2}{x_r-x_l}\left(\frac{1}{x_r-x}u(x_r)+\frac{1}{x-x_l}u(x_l)\right)-\frac{2}{(x_r-x)(x-x_l)}u(x)} + \mathcal{O}(x_r-x_l)$$

Verfahren (**Shortley-Weller-Diskretisierung**).
Dadurch inspiriert approximieren wir den Laplace-Operator durch

$$\begin{aligned}\mathcal{D}_h u(x,y) = \frac{1}{h^2}\Bigg(&\frac{2u(x-s_l h,y)}{s_l(s_r+s_l)} + \frac{2u(x+s_r h,y)}{s_r(s_r+s_l)}\\ &+\frac{2u(x,y-s_u h)}{s_u(s_o+s_u)} + \frac{2u(x,y+s_o h)}{s_o(s_o+s_u)} - \left(\frac{2}{s_l s_r}+\frac{2}{s_o s_u}\right)u(x,y)\Bigg)\end{aligned}$$

wobei $x_r - x = s_r h$, $\;x-x_l = s_l h$, $\;y_o - y = s_o h$, $\;y-y_u = s_u h$.
Wir betrachten nun

$$(\mathrm{RWP}_2)_\mathrm{h}{}'\quad \begin{cases} -\mathcal{D}_h u_h = f_h & \text{in } \Omega_h\\ \quad\;\; u_h = g & \text{auf } \partial\Omega_h\end{cases}$$

$$(\mathrm{LGS}_2)'\quad -\tilde{\mathcal{D}}_h\tilde u_h = \tilde f_h := f_h + g_h$$

$$\text{mit}\quad g_h(x,y) = 1/h^2\sum_{(x_N,y_N)\in N(x,y)\cap\partial\Omega_h} S_{x_N,y_N}\,g(x_N,y_N)$$

wobei

$$S_{x_N,y_N} := \begin{cases} 2/s_l(s_l+s_r) & \text{falls } (x_N,y_N) = (x-s_l h, y),\\ 2/s_r(s_l+s_r) & \text{falls } (x_N,y_N) = (x+s_r h, y),\\ 2/s_o(s_o+s_u) & \text{falls } (x_N,y_N) = (x, y+s_o h),\\ 2/s_u(s_o+s_u) & \text{falls } (x_N,y_N) = (x, y-s_u h),\end{cases}$$

$$-\tilde{\mathcal{D}}_h = (d_{ij})\quad\text{mit}\quad d_{ii} = 1/h^2\left(\tfrac{2}{s_{il}s_{ir}} + \tfrac{2}{s_{iu}s_{io}}\right)\quad\text{und}$$

$$d_{ij} = 1/h^2\begin{cases} -2/s_{il}(s_{il}+s_{ir}) & \text{falls } j \text{ der linke Nachbar von } i \text{ ist},\\ -2/s_{ir}(s_{il}+s_{ir}) & \text{falls } j \text{ der rechte Nachbar von } i \text{ ist},\\ -2/s_{iu}(s_{iu}+s_{io}) & \text{falls } j \text{ der untere Nachbar von } i \text{ ist},\\ -2/s_{io}(s_{iu}+s_{io}) & \text{falls } j \text{ der obere Nachbar von } i \text{ ist}.\end{cases}$$

Lem.
- Die Matrix $-\tilde{\mathcal{D}}_h$ ist eine M-Matrix.
- Sei $\Omega\subset\mathbb{R}^2$ beschränkt und gehöre zu dem Streifen $(x_0,x_0+d)\times\mathbb{R}$ oder $\mathbb{R}\times(y_0,y_0+d)$. Dann gilt $\|\tilde D_h^{-1}\|\le d^2/8$.

Bem. Das DV $(\mathrm{RWP}_2)_\mathrm{h}{}'$ hat in den randnahen Punkten nur die Konsistenzordnung 1. Dennoch gilt:

Satz. Sei $\Omega\subset\mathbb{R}^2$ beschränkt. Dann ist das Verfahren $(\mathrm{RWP}_2)_\mathrm{h}{}'$ konvergent von der Ordnung 2. Genauer gilt: Ist Ω eine Teilmenge des Streifens $(x_0,x_0+d)\times\mathbb{R}$ oder $\mathbb{R}\times(y_0,y_0+d)$, so ist

$$\|u_h - R_h u\|_h \le \left(1/3h^3 + d^2/48h^2\right)\|u\|_{\mathcal{C}^4(\overline{\Omega})}.$$

Ein weiteres Verfahren beruht auf:

Idee. Bestimme den Wert von u bei randnahen Punkten (x,y) durch lineare Interpolation:

- $u(x,y)\approx \dfrac{s_r}{s_r+s_l}u(x-s_l h, y) + \dfrac{s_l}{s_r+s_l}u(x+s_r h, y)$
- $u(x,y)\approx \dfrac{s_o}{s_u+s_o}u(x, y-s_u h) + \dfrac{s_u}{s_u+s_o}u(x, y+s_o h)$

Dies führt auf Gleichungen

$$(\mathrm{RWP}_2)_\mathrm{h}{}''\quad \begin{cases} -\mathcal{D}_h u = f_h & \text{in } \Omega_h\\ \quad\;\; u_h = g & \text{auf } \partial\Omega_h\end{cases}$$

$$(\mathrm{LGS}_2)''\quad -\tilde{\mathcal{D}}_h\tilde u_h = \tilde f_h$$

Lem. Dieses Verfahren besitzt Konsistenzordnung (und somit Konvergenzordnung) 2.

Allgemeine Differentialoperatoren

Problem. Wir betrachten nun

$$(\mathrm{RWP}_3)\quad \begin{cases} -\mathcal{L}u = f & \text{in } \Omega = (0,1)\times(0,1) \\ u = g & \text{auf } \partial\Omega \end{cases}$$

mit

$$\begin{aligned} -\mathcal{L}u = &-(a_{11}(x,y)u_{xx} + 2a_{12}(x,y)u_{xy} + a_{22}(x,y)u_{yy}) \\ &+ b_1(x,y)u_x + b_2(x,y)u_y + c(x,y)u \end{aligned}$$

wobei $c(x,y) \leq 0$, $\xi^T A(x,y)\xi \geq \lambda_0 \|\xi\|^2$, $\lambda_0 > 0$ und

$$A(x,y) = \begin{pmatrix} a_{11}(x,y) & a_{12}(x,y) \\ a_{21}(x,y) & a_{22}(x,y) \end{pmatrix}$$

Verfahren. 1. Diskretisierung: $h = 1/n$, Ω_h, $\partial\Omega_h$ wie früher.
2. Approximation:

$$u_x(x,y) \approx \frac{u(x+h,y)-u(x-h,y)}{2h}, \qquad u_y(x,y) \approx \dots$$

$$u_{xx}(x,y) \approx \frac{u(x+h,y)-2u(x,y)+u(x-h,y)}{h^2}, \quad u_{yy}(x,y) \approx \dots$$

Für die Approx. von u_{xy} haben wir mehrere Möglichkeiten: Wir könnten etwa den zentralen DQ in x- und y-Richrung verwenden und erhalten

$$\begin{aligned} &u_{xy}(x,y) \approx \tfrac{1}{2h}\left(u_x(x,y+h) - u_x(x,y-h)\right) \\ &\approx \tfrac{1}{4h^2}\left(u(x+h,y+h) - u(x+h,y-h) - u(x-h,y+h) + u(x-h,y-h)\right) \end{aligned}$$

Diese Annäherung hat allerdings den Nachteil, dass sie zu keiner M-Matrix führt. Stattdessen nehmen wir

$$u_{xy}(x,y) \approx \quad \frac{1}{2h^2}\begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix} \quad \frac{1}{2h^2}\begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

für $a_{12} \geq 0$, für $a_{12} < 0$.

Wir fassen diese Approx. in folgendem *7-Stern* zusammen:

$$\begin{aligned} -\mathcal{L}_h u := \ &\frac{1}{h^2}\begin{pmatrix} a_{12}^- & |a_{12}| - a_{22} & -a_{12}^+ \\ |a_{12}| - a_{11} & 2(a_{11} + a_{22} - |a_{12}|) & |a_{12}| - a_{11} \\ -a_{12}^+ & |a_{12}| - a_{22} & a_{12}^- \end{pmatrix} \\ &+ \frac{1}{2h}\begin{pmatrix} & b_2 & \\ -b_1 & 0 & b_1 \\ & -b_2 & \end{pmatrix} + \begin{pmatrix} & & \\ & c & \\ & & \end{pmatrix} \end{aligned}$$

Dabei ist $a_{ij}^+ := \max(a_{ij}, 0)$ und $a_{ij}^- := \min(a_{ij}, 0)$.

$$(\mathrm{RWP}_3)_\mathrm{h}\quad \begin{cases} -\mathcal{L}_h u_h = f_h & \text{in } \Omega_h \\ u_h = g & \text{auf } \partial\Omega_h \end{cases}$$

$$(\mathrm{LGS}_3)\qquad -\tilde{\mathcal{L}}_h \tilde{u}_h = \tilde{f}_h$$

Satz. Sei $c \geq 0$ in Ω und $\mathcal{L}$ gleichmäßig elliptisch.
Falls $a_{ii} > |a_{12}| + \frac{h}{2}|b_i|$ für $i = 1,2$ in Ω und $u \in \mathcal{C}^4(\overline{\Omega})$, so ist das DV $(\mathrm{RWP}_3)_\mathrm{h}$ konvergent von der Ordnung 2.

Differenzenverfahren für parabolische DGLn

Problem. Wärmeleitungsgleichung

$$(\mathrm{RWP}_4)\quad \begin{cases} u_t(x,t) - \Delta_x u(x,t) = f(x,t) & \text{in } \Omega = (0,1)\times(0,T) \\ u(x,0) = g(x) & \text{für } x \in (0,1) \\ u(0,t) = g_0(t) & \text{für } t \in [0,T] \\ u(1,t) = g_1(t) & \text{für } t \in [0,T] \end{cases}$$

Verfahren. 1. Diskretisierung mit n Raum- und m Zeitschritten:

$$x_i = ih, \quad h = 1/n, \quad t_k = k\tau, \quad \tau = T/m, \quad u(x_i, t_k) \approx u_i^k$$

2. Approximation der Ableitungen:

$$u_{xx}(x,t) \approx \frac{1}{h^2}\left(u(x-h,t) - 2u(x,t) + u(x+h,t)\right) =: \Delta_h u(x,t)$$

Wir wollen nun eine Lösung von

$$\begin{cases} \dot{u}_h(t) - \tilde{\Delta}_h u_h(t) = f_h(t) \\ u_h(0) = g_h \end{cases}$$

für alle Zeiten t mit

$$u_h(t) = \begin{pmatrix} u_h(h,t) \\ u_h(2h,t) \\ \vdots \\ u_h(1-h,t) \end{pmatrix}, \quad f_h(t) = \begin{pmatrix} f(h,t) + \frac{1}{h^2}g_0(t) \\ f(2h,t) \\ \vdots \\ f(1-h,t) + \frac{1}{h^2}g_1(t) \end{pmatrix}$$

berechnen. Dazu verwenden wir ein *Einschrittverfahren*, wie das *expl./impl. Gauß-Verfahren* oder das *Crank-Nicolson-Verfahren*:

$$(\mathrm{EEV})\quad \begin{cases} \frac{1}{\tau}(u_i^{k+1} - u_i^k) - \tilde{\Delta}_h u_i^k = f_i^k \\ u_i^0 = g_h \end{cases}$$

$$(\mathrm{IEV})\quad \begin{cases} \frac{1}{\tau}(u_i^{k+1} - u_i^k) - \tilde{\Delta}_h u_i^{k+1} = f_i^{k+1} \\ u_i^0 = g_h \end{cases}$$

$$(\mathrm{CNV})\quad \begin{cases} \frac{1}{\tau}(u_i^{k+1} - u_i^k) - \frac{1}{2}\tilde{\Delta}_h(u_i^k + u_i^{k+1}) = f(x_i, t_k + \frac{\tau}{2}) \\ u_i^0 = g_h \end{cases}$$

Lem. Sei $f(x,-) \in \mathcal{C}^1([0,T])$ für alle $x \in [0,1]$.
Dann gilt für die Approximation von (RWP_4):

- Die Verfahren (EEV) und (IEV) besitzen einen Konsistenzfehler von $\mathcal{O}(h^2 + \tau)$, falls $u \in \mathcal{C}^4([0,1]\times[0,T])$
- Das Verfahren (CNV) besitzt einen Konsistenzfehler von $\mathcal{O}(h^2 + \tau^2)$, falls $u \in \mathcal{C}^4([0,1]\times[0,T])$.

Lem. Es gelte $2\tau \leq h^2$ für (EEV). Die Verfahren (EEV), (IEV) und (CNV) sind stabil.

Differenzenverfahren für hyperbolische DGLn

Problem. Wellengleichung

$$\begin{cases} \partial_{tt}u - c^2\partial_{xx}u = f(x,t) & \text{in } \Omega = (0,1)\times[0,T] \\ u(0,t) = g_0(t), \quad u(1,t) = g_1(t) & \text{für } t \in [0,T] \\ u(x,0) = q_0(x), \quad u_t(x,0) = q_1(x) & \text{für } x \in (0,1) \end{cases}$$

Verfahren. 1. Diskretisierung: $x_i = ih$, $h = \frac{1}{n}$, $t_k = k\tau$, $\tau = \frac{T}{m}$
2. Approximation:

$$\begin{aligned} \partial_{xx}u(x_i,t_k) &\approx \tfrac{1}{h^2}\left(u(x_{i-1},t_k) - 2u(x_i,t_k) + u(x_{i+1},t_k)\right) \\ \partial_{tt}u(x_i,t_k) &\approx \tfrac{1}{\tau^2}\left(u(x_i,t_{k-1}) - 2u(x_i,t_k) + u(x_i,t_{k+1})\right) \\ \partial_t u(x_i,0) &\approx \tfrac{1}{2\tau}\left(u(x_i,t_1) - u(x_i,t_{-1})\right) \end{aligned}$$

Wir erhalten das lineare Gleichungssystem

$$\begin{cases} \frac{1}{\tau^2}(u_i^{k-1} - 2u_i^k + u_i^{k+1}) - \frac{c^2}{h^2}(u_{i-1}^k - 2u_i^k + u_{i+1}^k) = f_i^k \\ \text{für } i = 1,\dots,n-1 \text{ und } k = 0,\dots,m. \\ u_0^k = g_0^k = g_0(t_k), \quad u_n^k = g_1^k = g_1(t_k), \\ u_i^0 = q_{0,i} = q_0(x_i), \quad \frac{1}{2\tau}(u_i^1 - u_i^{-1}) = q_{1,i} = q_1(x_i) \end{cases}$$

Fazit

Das Differenzenverfahren …

☺ … ist einfach in der Herleitung und Implementierung.

😐 … besitzt eine gute Konvergenz (z. B. Ordnung 2) bei genügend glatter Lösung.

☹ … ermöglicht Adaptivität bzw. unregelm. Gitter nur schwer.

Schwache Lsgstheorie für elliptische DGLn

Def. Der **L^p-Raum** ist für $1 \le p < \infty$ definiert durch

$$L^p(\Omega) := \{v : \Omega \to \mathbb{R} \mid \|v\|_p < \infty\} \quad \text{mit} \quad \|v\|_p := \Big(\int_\Omega |v(x)|^p \,\mathrm{d}x\Big)^{1/p},$$

für $p = \infty$ durch

$$L^\infty(\Omega) := \{v : \Omega \to \mathbb{R} \mid \|v\|_\infty < \infty\} \quad \text{mit} \quad \|v\|_\infty := \operatorname*{ess\,sup}_{x\in\Omega} |v(x)|.$$

Bem. $(L^p(\Omega), \|-\|_p)$ ist ein Banachraum, für $p = 2$ sogar ein Hilbertraum mit Skalarprodukt $\langle u, v\rangle_{L^2(\Omega)} := \int_\Omega u(x)v(x)\,\mathrm{d}x$.

Satz (Höldersche Ungleichung). Sei $u \in L^p(\Omega)$ und $v \in L^q(\Omega)$ mit $1 \le p, q, r \le \infty$ und $1/p + 1/q = 1/r$. Dann ist $uv \in L^r(\Omega)$ mit

$$\|uv\|_r \le \|u\|_p \cdot \|v\|_q.$$

Def. Der Raum aller k-mal stetig differenzierbaren Funktionen auf Ω mit **kompaktem Träger** ist

$$\mathcal{C}_0^k(\Omega) := \{\varphi \in \mathcal{C}^k(\Omega) \mid \operatorname{supp}(\varphi) := \overline{\{x \in \Omega \mid \varphi(x) \ne 0\}} \text{ ist kompakt}\}.$$

Def. $\mathcal{D}(\Omega) := \mathcal{C}_0^\infty(\Omega)$ heißt Raum der **Testfunktionen** in Ω.

Lem (Partielle Integration). Für $u, v \in \mathcal{C}^1(\overline{\Omega})$ gilt

$$\int_\Omega v(x)\mathcal{D}_i u(x)\,\mathrm{d}x = \int_{\partial\Omega} v(x)u(x)\eta_i(x)\,\mathrm{d}x - \int_\Omega \mathcal{D}_i v(x)u(x)\,\mathrm{d}x.$$

Für $u \in \mathcal{C}^k(\overline{\Omega})$, $\varphi \in \mathcal{C}_0^k(\Omega)$ und $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| \le k$ gilt

$$\int_\Omega \varphi(x)\mathcal{D}^\alpha u(x)\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega \mathcal{D}^\alpha\varphi(x)u(x)\,\mathrm{d}x.$$

Def. $L^1_{\mathrm{loc}}(\Omega) := \{v : \Omega \to \mathbb{R} \mid v|_K \in L^1(K) \text{ für jedes kpkte } K \subset \Omega\}$ heißt Raum der **lokal integrierbaren Funktionen**.

Def. Sei $u \in L^1_{\mathrm{loc}}(\Omega)$ und $\alpha \in \mathbb{N}^n$. Eine Funktion $v \in L^1_{\mathrm{loc}}(\Omega)$ heißt **schwache (partielle) Ableitung** von u (oder die Ableitung von u im distributionellen Sinn) der Ordung α, wenn

$$\int_\Omega \varphi(x)v(x)\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega \mathcal{D}^\alpha\varphi(x)u(x)\,\mathrm{d}x \quad \text{für alle } \varphi \in \mathcal{D}(\Omega).$$

Bem. Ist eine Funktion im klassischen Sinne differenzierbar, so auch im schwachen mit derselben Ableitung.

Lem (Fundamentallemma der Variationsrechung). Für $u \in L^1_{\mathrm{loc}}(\Omega)$:

$$\int_\Omega u(x)\varphi(x)\,\mathrm{d}x = 0 \;\; \forall \varphi \in \mathcal{D}(\Omega) \quad \Longrightarrow \quad u \equiv 0 \text{ (fast-überall)}.$$

Kor. Die schwache Ableitung ist eindeutig bestimmt, d. h. sind $v, w \in L^1_{\mathrm{loc}}(\Omega)$ schwache Ableitungen von u, so gilt $v \equiv w$ f. ü. in Ω.

Bsp. Die schw. Abl. von $u(x) = |x|$ ist $v(x) = \mathbb{1}_{(0,\infty)} - \mathbb{1}_{(-\infty,0)}$.

Lem. • $\mathcal{D}^\alpha(u + \lambda v) = \mathcal{D}^\alpha u + \lambda\mathcal{D}^\alpha v$ • $\mathcal{D}^{\alpha+\beta}u = \mathcal{D}^\alpha(\mathcal{D}^\beta u)$

Def. Der **Sobolev-Raum** für $1 \le p < \infty$ ist

$$W^{k,p}(\Omega) := \left\{ u \in L^p(\Omega) \;\middle|\; \begin{array}{l} \forall\, \alpha \in \mathbb{N}^n \text{ mit } |\alpha| \le k : \\ \exists \text{schwache Ableitung } \mathcal{D}^\alpha u \in L^p(\Omega) \end{array} \right\}$$

$$\|u\|_{k,p} := \Big(\sum_{|\alpha|\le k} \|\mathcal{D}^\alpha u\|_p^p\Big)^{1/p}.$$

Notation. $H^k(\Omega) := W^{k,2}(\Omega)$

Satz. $(W^{k,p}(\Omega), \|-\|_{k,p})$ ist ein Banachraum.

Bem. $H^k(\Omega)$ ist sogar ein Hilbertraum mit Skalarprodukt

$$\langle u, v\rangle_{H^k(\Omega)} := \int_\Omega \sum_{\alpha \le k} D^\alpha u D^\alpha v\,\mathrm{d}x.$$

Satz (Meyers/Serrin). $W^{k,p}(\Omega) \cap \mathcal{C}^\infty(\Omega)$ liegt dicht in $W^{k,p}(\Omega)$,

$$\overline{W^{k,p}(\Omega) \cap \mathcal{C}^\infty(\Omega)}^{\|-\|_{k,p}} = W^{k,p}(\Omega).$$

Def. $W_0^{k,p}(\Omega) := \overline{\mathcal{D}(\Omega)}^{\|-\|_{k,p}}, \quad H_0^1(\Omega) := W_0^{1,2}(\Omega)$

Satz. Sei Ω ein beschränktes $\mathcal{C}^1$-Gebiet und $1 \le p < \infty$. Dann existiert eine lineare stetige Abbildung $\tau : W^{1,p}(\Omega) \to L^p(\partial\Omega)$, sodass für alle $u \in W^{1,p}(\Omega) \cap \mathcal{C}(\overline{\Omega})$ gilt: $\tau(u) = u|_{\partial\Omega}$.

Def. Die Abbildung τ heißt **Spuroperator**, $\tau(u)$ heißt die **Spur** von $u \in W^{1,p}(\Omega)$ auf $\partial\Omega$.

Satz. Sei Ω ein beschränktes $\mathcal{C}^1$-Gebiet. Dann gilt

$$W_0^{1,p}(\Omega) = \{v \in W^{1,p}(\Omega) \mid \tau(v) = 0\}.$$

Def. Der **Dualraum** eines Banachraums $(U, \|-\|_U)$ ist

$$U' := \{\text{lineare, stetige Abbildungen } \psi : U \to \mathbb{R}\} \quad \text{mit}$$

$$\|\psi\|_{U'} := \sup_{u \in U\setminus\{0\}} \frac{|\psi(u)|}{\|u\|_U} = \sup_{u \in U,\, \|u\|_U = 1} \psi(u).$$

Bsp. Gelte $1/p + 1/q = 1$ mit $p, q \in (1, \infty)$. Dann ist die Abbildung

$$j : L^q(\Omega) \to (L^p(\Omega))', \quad f \mapsto (g \mapsto \int_\Omega f(x)g(x)\,\mathrm{d}x)$$

ein isometrischer Isomorphismus.

Notation. $\langle \psi, u\rangle_{U',U} := \psi(u)$ für $\psi \in U'$, $u \in U$.

Satz (Riesz'scher Darstellungssatz). Sei $(H, \langle -, -\rangle_H)$ ein Hilbertraum. Dann ist

$$j : H \to H', \quad \psi \mapsto (\phi \mapsto \langle \psi, \phi\rangle_H)$$

ein isometrischer Isomorphismus.

Def. $W^{-1,q} := (W_0^{1,p}(\Omega))'$, wobei $1/p + 1/q = 1$,
$H^{-1}(\Omega) := W^{-1,2} = (H_0^1(\Omega))'$.

Def. Für $\Omega \subset \mathbb{R}^d$ heißt $\gamma := k - d/p$ **Sobolev-Zahl** zu $W^{k,p}(\Omega)$.

Satz. Für $m < \gamma$ ist $W^{k,p}(\Omega) \hookrightarrow \mathcal{C}^m(\Omega)$ wohldefiniert und stetig.

Variationsgleichungen

Situation. Sei $\Omega \subset \mathbb{R}^d$ beschränkt. Wir betrachten nun wieder

$$(\mathrm{RWP}_1) \quad \begin{cases} \mathcal{L}u = f & \text{in } \Omega \\ u = 0 & \text{auf } \partial\Omega \end{cases}$$

$$\text{mit} \quad \mathcal{L}u(x) = -\sum_{i=1}^d \mathcal{D}_i\Big(\sum_{j=1}^d a_{ij}(x)\mathcal{D}_j u\Big) + \sum_{i=1}^d b_i(x)\mathcal{D}_i u + c(x)u$$
$$= -\operatorname{div}(A(x)\mathcal{D}u) + b(x)\cdot\mathcal{D}u + c(x)u.$$

Sei u eine Lösung von (RWP_1) und $\varphi \in \mathcal{D}(\Omega)$. Dann gilt

$$\int_\Omega f(x)\varphi(x)\,\mathrm{d}x = \int_\Omega \mathcal{L}u(x)\varphi(x)\,\mathrm{d}x$$
$$= -\int_\Omega \operatorname{div}(A(x)\mathcal{D}u(x))\varphi(x)\,\mathrm{d}x + \int_\Omega (b(x)\cdot\mathcal{D}u(x) + c(x)u(x))\cdot\varphi(x)\,\mathrm{d}x$$
$$= \int_\Omega A(x)\mathcal{D}u(x)\cdot\mathcal{D}\varphi(x)\,\mathrm{d}x + \int_\Omega (b(x)\cdot\mathcal{D}u(x) + c(x)u(x))\cdot\varphi(x)\,\mathrm{d}x$$

Def. Eine Funktion $u \in H_0^1(\Omega)$ heißt **schwache Lösung** von (RWP_1), wenn u folgende Variationsgleichung erfüllt:

$$\int_\Omega A(x)\mathcal{D}u(x)\cdot\mathcal{D}\varphi(x) + b(x)\cdot\mathcal{D}u(x)\varphi(x) + c(x)u(x)\varphi(x)\,\mathrm{d}x$$
$$= \int_\Omega f(x)\varphi(x)\,\mathrm{d}x \quad \text{für alle } \varphi \in H_0^1(\Omega). \qquad (\mathrm{VGL}_1)$$

Sprechweise. Der Raum, in dem man u sucht, heißt **Lsgsraum** (oder *Ansatzraum*), der Raum von φ heißt **Testraum**.

Problem (Allg. Variationsproblem). Seien Abb. $\ell : H_0^1(\Omega) \to \mathbb{R}$ und $B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ gegeben. Gesucht ist $u \in H_0^1(\Omega)$, sodass

$$(\mathrm{VGL}_1)' \qquad B(u, \varphi) = \ell(\varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

Bem. Im obigen Setting ist $\ell(\varphi) := \int_\Omega f(x)\varphi(x)\,\mathrm{d}x$ und

$$B(u, \varphi) := \int_\Omega A(x)\mathcal{D}u(x)\cdot\mathcal{D}\varphi(x) + b(x)\cdot\mathcal{D}u(x)\varphi(x) + c(x)u(x)\varphi(x)\,\mathrm{d}x$$

Def. Sei X ein Banachraum. Eine bilin. Abb. $B : X \times X \to \mathbb{R}$ heißt

- **positiv**, falls $B(u, u) > 0$ für alle $u \in X \setminus \{0\}$,
- **stark positiv** (*koerziv*, *koerzitiv*), falls ein $\lambda > 0$ existiert, sodass
$$\forall\, u \in X \,:\, B(u, u) \ge \lambda\|u\|_X^2,$$
- **beschränkt** (oder *stetig*), falls ein $\mu > 0$ existiert, sodass
$$\forall\, u, \varphi \in X \,:\, |B(u, \varphi)| \le \mu\|u\|_X\|\varphi\|_X.$$

Lem. • Die Abbildung B in $(\mathrm{VGL}_1)'$ ist bilinear und beschränkt.
• Die Abbildung ℓ in $(\mathrm{VGL}_1)'$ ist linear und stetig.

Satz. Sei Ω ein beschr. Lipschitz-Gebiet. Dann ist jede klassische Lsg $u \in \mathcal{C}^2 \cap \mathcal{C}^1(\partial\Omega)$ von (RWP_1) eine schwache Lsg von $(\mathrm{VGL}_1)'$.

Eindeutige Lösung elliptischer DGLn

Satz (Lax-Milgram). Sei H ein Hilbertraum und $B : H \times H \to \mathbb{R}$ eine beschränkte, koerzitive Bilinearform. Dann gibt es für jedes $\ell \in H'$ eine eindeutige Lösung $u \in H$ von $\forall \varphi \in H : B(u,\varphi) = \ell(\varphi)$. Es gilt $\|u\|_H \leq 1/\lambda \|\ell\|_{H'}$ mit der Koerzitivitätskonstante λ von B.

Lem (Poincaré-Ungleichung). Sei Ω beschränkt. Dann $\exists C > 0$:

$$\|u\|_{L^2(\Omega)} \leq C\|\nabla u\|_{L^2(\Omega,\mathbb{R}^d)} = C\left(\int_\Omega \sum_i |\mathcal{D}_i u|^2 \,\mathrm{d}x\right)^{1/2} \quad \forall\, u \in H^1_0(\Omega).$$

Kor. Mit $C_1 := (1+C^2)^{-1/2}$ und $C_2 := 1$ gilt für alle $u \in H^1_0(\Omega)$:

$$C_1\|u\|_{H^1(\Omega)} \leq \|\nabla u\|_{L^2(\Omega,\mathbb{R}^d)} \leq C_2\|u\|_{H^1(\Omega)}.$$

Lem. Falls $\mathcal{L}$ glm. elliptisch ist und $b(x) \equiv 0$ sowie $c(x) \geq 0$ in Ω gelten, so ist B in $(\mathrm{VGL}_1)'$ koerziv.

Satz. Sei $\Omega \subset \mathbb{R}^d$ beschränkt und sei $\mathcal{L}u = -\operatorname{div}(A(x)\mathcal{D}u) + c(x)u$ glm. elliptisch, $c(x) \geq 0$ in Ω, $a_{ij}, c_j \in L^\infty(\Omega)$ und $f \in L^2(\Omega)$. Dann besitzt $(\mathrm{VGL}_1)'$ eine eindeutige Lösung $u \in H^1_0(\Omega)$. Außerdem existiert ein $\hat{C} > 0$, sodass $\|u\|_{H^1(\Omega)} \leq \hat{C} \cdot \|f\|_{L^2(\Omega)}$.

Def. Sei $f \in H^{-1}(\Omega)$. Eine Fktn $u \in H^1_0(\Omega)$ heißt *schwache Lösung* von (RWP_1), falls $B(u,\varphi) = \langle f, \varphi\rangle_{H^{-1}(\Omega),H^1_0(\Omega)}$ $\forall \varphi \in H^1_0(\Omega)$.

Bem. Gelte $b \equiv 0$, $c \geq 0$, glm. Elliptizität, $c, a_{ij} \in L^\infty(\Omega)$. Dann existiert nach Lax-Milgram genau eine schwache Lösung.

Lem. Sei A gleichmäßig elliptisch und $a_{ij}, b_i, c \in L^\infty(\Omega)$. Dann existiert ein $\mu_0 > 0$, sodass für alle $\mu > \mu_0$ das RWP

$$\begin{cases} \mathcal{L}u + \mu u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{auf } \partial\Omega \end{cases}$$

für alle $f \in H^{-1}(\Omega)$ eine eind. schwache Lösung $u \in H^1_0(\Omega)$ besitzt.

RWPe mit anderen Randbedingungen

Problem. Wir untersuchen nun das inhomogene Randwertproblem

$$(\mathrm{RWP}_2) \quad \begin{cases} \mathcal{L}u &= f \quad \text{in } \Omega, \\ u &= g \quad \text{auf } \partial\Omega. \end{cases}$$

Angenommen, $g \in \mathcal{C}(\partial\Omega) \cap L^2(\partial\Omega)$ besitzt eine Fortsetzung $\tilde{g} \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ mit $\tilde{g}|_{\partial\Omega} = g$. Dann ist $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ genau dann eine Lösung von (RWP_2), wenn $v := u - \tilde{g}$ eine Lösung von

$$(\mathrm{RWP}_2)' \quad \begin{cases} \mathcal{L}v &= f - \mathcal{L}\tilde{g} \quad \text{in } \Omega, \\ v &= 0 \quad \text{auf } \partial\Omega \end{cases}$$

ist. Schwache Formulierung von $(\mathrm{RWP}_2)'$: Ges. ist $v \in H^1_0(\Omega)$ mit

$$\int_\Omega (A(x)\mathcal{D}v(x) \cdot \mathcal{D}\varphi(x) + b(x) \cdot \mathcal{D}v(x)\varphi(x) + c(x)v(x)\varphi(x))\,\mathrm{d}x$$
$$= \int_\Omega f \cdot \varphi \,\mathrm{d}x + \int_\Omega (A(x)\mathcal{D}\tilde{g} \cdot \mathcal{D}\varphi + b \cdot \mathcal{D}\tilde{g}\varphi + c\tilde{g}\varphi)\,\mathrm{d}x \quad \forall \varphi \in H^1_0(\Omega).$$

Voraussetzungen: $a_{ij}, b_i, c \in L^\infty(\Omega)$, $f \in L^2(\Omega)$, $\tilde{g} \in \mathcal{C}(\overline{\Omega}) \cap H^1(\Omega)$ und $g \in L^2(\partial\Omega)$. Ges. ist ein $u \in U := \{w \in H^1(\Omega) \mid \tau(w) = g\}$ mit

$$\underbrace{\int_\Omega A(x)\mathcal{D}u \cdot \mathcal{D}\varphi + b\mathcal{D}u\varphi + cu\varphi \,\mathrm{d}x}_{B(u,\varphi):=} = \underbrace{\int_\Omega f\varphi\,\mathrm{d}x}_{\ell(\varphi):=} \quad \forall \varphi \in H^1_0(\Omega). \quad (\mathrm{VGL}_2)$$

Für $f \in H^{-1}(\Omega)$ verwendet man

$$B(u,\varphi) = \ell'(\varphi) := \langle f,\varphi\rangle_{H^{-1}(\Omega),H^1_0(\Omega)} \quad \forall \varphi \in H^1_0(\Omega). \quad (\mathrm{VGL}_2)'$$

Satz. Sei $B : H^1_0(\Omega) \times H^1_0(\Omega) \to \mathbb{R}$ beschränkt und koerziv. Dann besitzt (VGL_2) genau dann eine eindeutige Lösung $u \in U$, wenn ein $u_0 \in H^1(\Omega)$ existiert, sodass $\tau(u_0) \equiv g$ f. ü. auf $\partial\Omega$.

Problem. Wir betrachten nun die Randbedingung

$$(\mathrm{RWP}_3) \quad \begin{cases} \mathcal{L}u &= f \quad \text{in } \Omega, \\ A(x)\mathcal{D}u \cdot \nu + \mu u &= g \quad \text{auf } \partial\Omega \text{ (glatt)} \end{cases}$$

Falls $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ eine Lösung ist und $\varphi \in \mathcal{C}^\infty(\overline{\Omega})$, so gilt

$$\int_\Omega f\varphi\,\mathrm{d}x = -\int_\Omega \operatorname{div}(A(x)\mathcal{D}u)\varphi + b(x)\mathcal{D}u\varphi + c(x)u\varphi\,\mathrm{d}x$$
$$= -\int_{\partial\Omega} A(x)\mathcal{D}u \cdot \nu\varphi\,\mathrm{d}s + \int_\Omega A(x)\mathcal{D}u\mathcal{D}\varphi + b(x)\mathcal{D}u\varphi + c(x)u\varphi\,\mathrm{d}x.$$

Aus der Randbedingung bekommen wir

$$\int_{\partial\Omega} A(x)\mathcal{D}u \cdot \nu\varphi\,\mathrm{d}s + \mu\int_{\partial\Omega} u\varphi\,\mathrm{d}s = \int_{\partial\Omega} g\varphi\,\mathrm{d}s.$$

Zusammengesetzt erhalten wir die Variationsgleichung

$$\mu\int_{\partial\Omega} u\varphi\,\mathrm{d}s + \int_\Omega A\mathcal{D}u\mathcal{D}\varphi + b \cdot \mathcal{D}u\varphi + cu\varphi\,\mathrm{d}x = \int_\Omega f\varphi\,\mathrm{d}x + \int_{\partial\Omega} g\varphi\,\mathrm{d}s.$$

Wegen Dichtheit von $\mathcal{C}^\infty(\overline{\Omega})$ in $H^1(\Omega)$ ist diese Gleichung nicht nur für $\varphi \in \mathcal{C}^\infty(\overline{\Omega})$ sondern allgemeiner für $\varphi \in H^1(\Omega)$ erfüllt.

Def. Sei $\mu \in \mathbb{R}$, $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega)$. Eine Fktn $u \in H^1(\Omega)$ heißt *schwache Lösung* von (RWP_3), falls für alle $\varphi \in H^1(\Omega)$ gilt:

$$\underbrace{\mu\int_{\partial\Omega} u\varphi\,\mathrm{d}s + \int_\Omega A\mathcal{D}u\mathcal{D}\varphi + b\mathcal{D}u\varphi + cu\varphi\,\mathrm{d}x}_{B(u,\varphi):=} = \underbrace{\int_\Omega f\varphi\,\mathrm{d}x + \int_{\partial\Omega} g\varphi\,\mathrm{d}s}_{\ell(\varphi):=}.$$

Approximation von Variationsgleichungen

Verfahren. Gegeben sei ein Hilbertraum H, eine beschränkte, koerzive Bilinearform $B : H \times H \to \mathbb{R}$ und ein $\ell \in H'$. Gesucht ist eine Lösung $u \in H$ der Variationsgleichung

$$(\mathrm{VGL}) \qquad B(u,\varphi) = \ell(\varphi) \quad \forall \varphi \in H.$$

Wir wollen diese Lösung annähern durch die Lösung eines möglichst ähnlichen, aber *endlichdim.* Problems. Dazu wählen wir einen endlichdim. Unterraum $U_n \subset H$ (dieser ist wieder ein Hilbertraum), eine beschränkte, koerzive Bilinearform $B_n : U_n \times U_n \to \mathbb{R}$ und ein Element $\ell_n \in U_n'$. Wir bestimmen dann die Lösung u_n von

$$(\mathrm{VGL})_\mathrm{n} \qquad B_n(u_n,\varphi) = \ell_n(\varphi) \quad \forall \varphi \in U_n.$$

Fragen. 1. Wie berechnet man die Lösung u_n von $(\mathrm{VGL})_\mathrm{n}$?
2. Wie kann man U_n, B_n und ℓ_n wählen, sodass $u_n \xrightarrow[n\to\infty]{} u$?

Def. Die Approximation von (VGL) mittels $(\mathrm{VGL})_\mathrm{n}$ mit $U_n \subset H$ und $B_n = B|_{U_n \times U_n}$ heißt **konforme Approximation** von (VGL). Eine solche Methode wird als **Verfahren von Ritz** bezeichnet.

Vorgehen. Um die Lösung u_n von $(\mathrm{VGL})_\mathrm{n}$ zu berechnen, wählen wir zunächst eine Basis $w_1, \ldots, w_{d_n}$ von U_n. Wir setzen

$$\hat{\ell} := (\ell_n(w_1), \ldots, \ell_n(w_{d_n}))^T \in \mathbb{R}^{d_n}, \quad \text{und}$$
$$\hat{B} := (B_{ij}) \in \mathbb{R}^{d_n \times d_n} \quad \text{mit} \quad B_{ij} := B(w_i, w_j).$$

Dann ist $u_n = \gamma_1 w_1 + \ldots + \gamma_{d_n} w_{d_n} \in U_n$ genau dann eine Lösung von $(\mathrm{VGL})_\mathrm{n}$, wenn $\gamma = (\gamma_1, \ldots, \gamma_{d_n})^T \in \mathbb{R}^{d_n}$ erfüllt:

$$\hat{B}\gamma = \hat{\ell} \qquad \text{(Galerkin-Gleichung)}.$$

Lem (Céa). Für die Lösungen u_n der konformen Approximation $(\mathrm{VGL})_\mathrm{n}$ und die Lösung u von (VGL) gilt

$$\|u_n - u\|_H \leq C\left(\inf_{v \in U_n} \|u - v\|_H + \|\ell_n - \ell\|_{U_n'}\right)$$

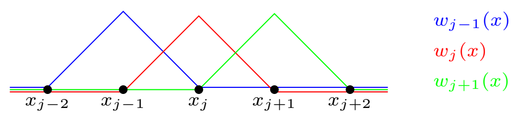
Folgerung. Für $u_n \xrightarrow[n\to\infty]{} u$ genügt es, U_n u. ℓ_n so zu wählen, dass

$$\forall\, u \in H : \inf_{v_n \in U_n} \|u - v_n\|_H \xrightarrow[n\to\infty]{} 0, \qquad \|\ell_n - \ell\|_{U_n'} \xrightarrow[n\to\infty]{} 0.$$

Bem. Ist $\ell(\varphi)$ durch Integration einer Funktion gegeben, so kann man für $\ell_n(\varphi)$ eine Annäherung dieses Integrals etwa mittels der summierten Trapezregel verwenden.

Bem. Wir betrachten (VGL_1). Es gibt mehrere sinnvolle Möglichkeiten, die Basiselemente w_i zu wählen. Man versucht dabei zu erreichen, dass die Matrix $\hat{B}$ möglichst einfach (wenige von null verschiedene Einträge) und gut konditioniert ist.

1. Angenommen, es gibt eine Basis von Eigenfunktionen w_j von $\mathcal{L}$, $\mathcal{L}w_j = \lambda_j w_j$, die in $L^2(\Omega)$ eine Orthonormalbasis bilden. Dann ist $\hat{B}$ eine Diagonalmatrix mit Einträgen λ_j. Beispielsweise ist $w_j(x) := \sin(\pi j x)$ eine EF von $\mathcal{L} := -\Delta$ auf $\Omega = (0,1)$ zum EW $\pi^2 j^2$ und es gilt $\langle w_i, w_j\rangle_{L^2(\Omega)} = 0$ für $i \neq j$.
2. $U_n := \operatorname{span}\{w_j(x) = x^j(1-x) \mid j = 1, \ldots, n\} \subset H^1_0(\Omega)$ (Das ist eine schlechte Wahl, da dann $\hat{B}$ vollbesetzt.)
3. Wir unterteilen $\Omega = (0,1)$ durch das Gitter $0 = x_0 < x_1 < \ldots < x_n < x_{n+1} = 1$, $x_i = ih$ mit $h = 1/n+1$,
$$U_n := \{v \in \mathcal{C}(0,1) \mid \forall i : v|_{[x_{i-1},x_i]} \in \mathbb{P}_1[x], v(0) = v(1) = 0\}$$
Eine Basis von U_n sind die *Hutfunktionen* $w_1, \ldots, w_n$:



Wir betrachten die VGL zu (RWP_1) mit $g_0 = g_1 = 0$. Wenn wir ℓ durch die Trapezregel approximieren, erhalten wir

$$\hat{B} = \frac{1}{h}\begin{pmatrix} 2 & -1 & & & & 0 \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ 0 & & & & -1 & 2 \end{pmatrix}, \quad \ell_n(w_j) = hf(x_j),$$

also das Finite-Differenzen-Verfahren.

Methode der Finiten Elemente (FEM)

Ziel. (VGL) mittels $(\text{VGL})_n$ approximieren.

Idee. Zerlege Ω in endlich viele Teilgebiete. Wir wählen U_n als Raum der Funktionen, die sich als Linearkombination von Basisfktn schreiben lassen, deren Träger nur wenige Teilgebiete umfasst.

Bem. Als Teilgebiete verwendet man in $\mathbb{R}^1$ regelmäßige Teilintervalle, in $\mathbb{R}^2$ Dreiecke oder Rechtecke und in $\mathbb{R}^3$ Tetraeder. Als lokale Ansatzfunktionen über den Teilgebieten verwenden wir Polynome. Globale Ansatzfktn über Ω sind lokale Ansatzfktn mit bestimmten Glattheitsbedingungen am Rand der Teilgebiete.

Def. Ein **finites Element** (FE) in $\mathbb{R}^d$ ist ein Tripel (K, P, Σ) mit

- $K \subset \mathbb{R}^d$ ist kompakt,
- ∂K ist Lipschitz-stetig,
- P ist ein endlichdim. lin. Raum von Funktionen $p \in \mathcal{C}^s(K, \mathbb{R})$
- $\Sigma = \{b_1, \ldots, b_m\}$ mit $b_j \in (\mathcal{C}^s(K, \mathbb{R}))'$ ist ***P*-unisolvent**, d. h.

$$\forall\, \alpha \in \mathbb{R}^m \;:\; \exists!\, p \in P \;:\; b_j(p) = \alpha_j, \quad j = 1, \ldots, m$$

Bem. Σ ist P-unisolvent $\iff$ Σ ist Basis von $P' \subset (\mathcal{C}^s(K, \mathbb{R}))'$

Lem.
- Sei $\Sigma = \{b_1, \ldots, b_m\}$ P-unisolvent.
  - $b_1, \ldots b_m$ sind linear unabhängig.
  - Sei $p_j \in P$ so gewählt, dass $b_i(p_j) = \delta_{ij}$. Dann ist $\{p_1, \ldots, p_m\}$ eine Basis von P.
- Sei $\{p_1, \ldots, p_m\}$ eine Basis von P und seien $b_i \in (\mathcal{C}^s(K, \mathbb{R}))'$ mit $b_i(p_j) = \delta_{ij}$, $i, j = 1, \ldots, m$. Dann ist $\{b_1, \ldots, b_m\}$ P-unisolvent.

Def. Ein finites Element (K, P, Σ) heißt zu einem anderen FE $(\hat{K}, \hat{P}, \hat{\Sigma})$ **affin äquivalent** *vermöge* einer invertierbaren affinen Abbildung $F : \mathbb{R}^d \to \mathbb{R}^d$, $\hat{x} \mapsto B\hat{x} + b$, falls gilt:

- $K = F(\hat{K})$
- $P = \{p := \hat{p} \circ F^{-1} \,|\, \hat{p} \in \hat{P}\}$
- $\Sigma = \{b_i : p \mapsto b_i(p \circ F) \,|\, \hat{b}_i \in \hat{\Sigma}\}$

Finite Elemente vom Lagrange-Typ

Def. Der ***d*-Simplex** mit Ecken $a_1, \ldots, a_{d+1} \in \mathbb{R}^d$ ist

$$K = \{x = \sum_{j=1}^{d+1} \mu_j a_j \,|\, 0 \leq \mu_j \leq 1,\ j = 1, \ldots, d+1,\ \sum_{j=1}^{d+1} \mu_j = 1\} \subset \mathbb{R}^d.$$

Dabei heißen $\mu_1, \ldots, \mu_{d+1}$ baryzentrische Koordinaten von x. K heißt **nicht entartet**, falls $a_1, \ldots, a_{d+1}$ affin unabhängig sind. Das Simplex mit den Ecken $a_j = e_j \in \mathbb{R}^d$, $j = 1, \ldots, d$, und $a_{d+1} = 0$ heißt ***d*-Einheitssimplex** $\hat{K}$.

Bspe. Ein nichtentartetes Simplex im $\mathbb{R}^1$ ist ein geschlossenes Intervall, im $\mathbb{R}^2$ ein Dreieck und im $\mathbb{R}^3$ ein Tetraeder.

Lem. Jedes nicht entartete d-Simplex K ist affin äquivalent zu $\hat{K}$: Es gibt genau eine Abb. $F : \hat{K} \to K$, $\hat{x} \mapsto A_K \hat{x} + b_K$ mit einer Matrix $A_K \in \mathrm{GL}(d)$, sodass $F(e_j) = a_j$ und $F(0) = a_{d+1}$.

Def. Ein **simpliziales finites Element vom Lagrange-Typ** der Ordnung k ist ein Tripel (K, P, Σ) mit

- einem d-Simplex K mit Ecken $a_1, \ldots, a_{d+1}$,
- $P = \mathbb{P}_k := \mathbb{R}[x_1, \ldots, x_d]_{\leq k} = \{\text{Polynome vom Grad} \leq k\}$
- $\Sigma = \{b_a : P \to \mathbb{R},\ p \mapsto p(a) \,|\, a \in \mathcal{K}_k\}$ mit der *Knotenmenge*

$$\mathcal{K}_k = \{x = \sum_{j=1}^{d+1} \mu_j a_j \,|\, \sum_{j=1}^{d+1} \mu_j = 1,\ \ \mu_j \in \{0, \tfrac{1}{k}, \ldots, \tfrac{k-1}{k}, 1\}\}.$$

Bemn.
- Sei K ein d-Simplex. Dann ist $\dim(\mathbb{P}_k) = |\mathcal{K}_k| = \binom{k+d}{d}$.
- Die kanonische Basis von $\mathbb{P}_1$ für $\hat{K} \subset \mathbb{R}^d$ ist

  $$\hat{p}_i(\hat{x}) = \hat{x}_i \text{ für } i = 1, \ldots, d, \qquad \hat{p}_{d+1}(\hat{x}) = 1 - \hat{x}_1 - \ldots - \hat{x}_d.$$

  Dann bilden die $\hat{b}_j \in \Sigma$ eine Dualbasis der $\hat{p}_i$, d. h. $\hat{b}_j(\hat{p}_i) = \delta_{ij}$.
- Die duale Basis von $\mathbb{P}_2$ zu Σ für $\hat{K} \subset \mathbb{R}^d$ ist

  $$\{\hat{p}_i \,|\, i \in \{1, \ldots, d+1\}\} \cup \{\hat{p}_{ij} \,|\, i \neq j \in \{1, \ldots, d+1\}\} \quad \text{wobei}$$
  $$\hat{p}_i(\hat{x}) = \mu_j(\hat{x})(2\mu_j(\hat{x}) - 1), \quad \hat{p}_{ij}(\hat{x}) = 4\mu_i(\hat{x})\mu_j(\hat{x}),$$
  $$\mu_i(\hat{x}) = \hat{x}_i \text{ für } i = 1, \ldots, d, \quad \mu_{d+1}(\hat{x}) = 1 - \hat{x}_1 - \ldots - \hat{x}_d$$
- Kanonische Basiselemente $p_1, \ldots, p_m \in \mathbb{P}_k$ für einen allgemeinen Simplex $K \subset \mathbb{R}^d$ kann man wie folgt berechnen:
  1. Finde eine affin lineare Bijektion $F : \hat{K} \to K$ wie früher.
  2. Setze $p_i(x) := \hat{p}_i(F^{-1}(x))$ für $i = 1, \ldots, m$.

Räume von Finite-Elemente-Funktionen auf Ω

Def. Sei $\Omega \subset \mathbb{R}^d$ ein beschränktes, polygonal berandetes Gebiet. Eine **Triangulierung** von Ω mit simplizialen finiten Elementen vom Lagrange-Typ der Ordnung k ist eine endliche Menge

$$T(\overline{\Omega}) = \{(K_i, P(K_i), \Sigma(K_i)) \,|\, i = 1, \ldots, N\} \quad \text{mit}$$

- $(K_i, P(K_i), \Sigma(K_i))$ sind simpl. El. vom Lagrange-Typ der Ord. k,
- $\overline{\Omega} = K_1 \cup \ldots \cup K_N$,
- $\mathrm{int}(K_i) \cap \mathrm{int}(K_j) = \emptyset$ für $i \neq j$,
- Jede Seite von K_i, d. h. jedes von d Eckpunkten von K_i aufgesp. $(d-1)$-dimensionale Simplex, ist entweder Teil des Gebietrandes $\partial\Omega$ oder gleichzeitig Seite genau eines anderen Simplex K_j.

Def. Die *Knotenmenge* von $T(\overline{\Omega})$ ist die Vereinigung der Knotenmengen der simpl. FE, d. h. $\mathcal{K}_k = \mathcal{K}_k(K_1) \cup \ldots \cup \mathcal{K}_k(K_N)$.

Def. Der **Raum der finiten Elemente** zu einer Triang. $T(\overline{\Omega})$ von Ω mit simplizialen FE vom Lagrange-Typ der Ordnung k und Knotenmenge $\mathcal{K}_k = \{\tilde{a}_1, \ldots, \tilde{a}_n\}$ ist

$$U_n = \{v \in \mathcal{C}(\overline{\Omega}) \,|\, v|_{K_i} \in \mathbb{P}_1(K_i) \text{ für } i = 1, \ldots, N\}.$$

Satz $(k = 1)$**.** Sei U_n der Raum der *linearen finiten Elemente* zu einer Triang. $T(\overline{\Omega})$ mit simpl. FE vom Lagrange-Typ der Ord. 1.

- Sei K ein nichtentart. d-Simplex mit Ecken $a_1, \ldots, a_{d+1}$. Dann ist durch $p(a_j), \ldots, p(a_{d+1})$ ein Polynom $p \in \mathbb{P}_1(K)$ eindeutig bestimmt. Für alle $p \in \mathbb{P}_1(K)$ und $x \in K$ gilt

  $$p(x) = p(a_1)p_1(x) + \ldots + p(a_{d+1})p_{d+1}(x). \text{ wobei } p_i(a_j) = \delta_{ij}.$$
- Sind $\mathcal{K} = \{\tilde{a}_1, \ldots, \tilde{a}_n\}$ die Knoten der Triangulierung, so ist eine Funktion $v \in U_n$ durch die Vorgabe von $v(\tilde{a}_1), \ldots, v(\tilde{a}_n)$ eindeutig definiert. Es gilt $U_n \subset H^1(\Omega)$.
- Eine Basis von U_n ist gegeben durch die Funktionen $p_j \in U_n$ mit $p_j(\tilde{a}_i) = \delta_{ij}$ für $i, j = 1, \ldots, n$. Insbesondere gilt $\dim U_n = n$.

Satz $(k \geq 1)$**.** Sei $T(\overline{\Omega})$ eine Triangulierung mit finiten Lagrange-Elementen der Ordnung k, U_n der zugehörige Raum der FE und $\mathcal{K}_k$ die Knotenmenge von $T(\overline{\Omega})$. Dann ist durch Vorgabe von $v|_{\mathcal{K}_k}$ eindeutig ein $v \in U_n \subset H^1(\Omega)$ bestimmt. Eine Basis von U_n ist durch $p_j \in U_n$ mit $p_j(\tilde{a}_i) = \delta_{ij}$, $i, j = 1, \ldots, n$, gegeben.

Verfahren (Realisierung der Finite-Elemente-Methode)**.**

1. Eingabe und Beschreibung des RWPs
2. Umformulierung in ein Variationsproblem
3. Generierung einer Triangulierung. Entweder *uniforme Zerlegung* oder Zerlegung mit *lokaler Verfeinerung* von Ω.
4. Erzeugung eines endlich-dim. Problems, d. h. Berechnen der Koeffizientenmatrix und der rechten Seite der Galerkin-Gleichung
5. Lösung der Galerkin-Gleichung

Konvergenz der FE-Methode

Def. Sei (K, P, Σ) ein finites Element, $P \subset \mathcal{C}^s(K)$. Dann heißt $\Pi_K w$ ***P*-Interpolierende** einer Fktn $w \in \mathcal{C}^s(K)$, falls

- $\Pi_K w \in P$,
- $b_j(\Pi_K w) = b_j(w)$ für jedes $b_j \in \Sigma$.

Bem.
- Ist $p_1, \ldots, p_m$ eine zu Σ duale Basis von P, d. h. $b_i(p_j) = \delta_{ij}$, so gilt $\Pi_K w = \sum_{i=1}^m b_i(w) p_i$.
- Für Lagrange-FE gilt $w(\tilde{a}_j) = b_j(w) = b_j(\Pi_K w) = \Pi_K w(\tilde{a}_j)$
- $\forall\, p \in P \;:\; \Pi_K p = p$

Def. Sei U_n ein FE-Raum zu einer Triangulierung $T(\overline{\Omega})$ und sei $\{p_1, \ldots, p_n\}$ die kanonische Basis von U_n, d. h. $b_i(p_j) = \delta_{ij}$. Die ***U_n*-Interpolierende** einer Funktion $w \in \mathcal{C}^s(\overline{\Omega})$ ist dann

$$\Pi w := \sum_{i=1}^n b_i(w) p_i \in U_n.$$

Lem. Für alle $K_i \in T(\overline{\Omega})$ und $w \in \mathcal{C}^s(\overline{\Omega})$ gilt $(\Pi w)|_K = \Pi_{K_i}(w|_{K_i})$ und somit $\|w - \Pi w\|_{H^1(\Omega)} = \sum_{K_i \in T(\overline{\Omega})} \|w - \Pi_{K_i} w\|_{H^1(K_i)}$.

Lem. Sei $F : \hat{K} \to K$ mit $F(\hat{x}) = A\hat{x} + b$, $A \in \mathrm{GL}(d)$, und $l \in \mathbb{N}$.

- Es gilt $\ v \in H^l(K) \iff v \circ F \in H^l(\hat{K})$
- Es existiert eine Konstante $c > 0$, sodass

  $$|v \circ F|_{H^l(\hat{K})} \leq c \cdot \|A\|_2^l \cdot \frac{1}{\sqrt{|\det(A)|}} \cdot |v|_{H^l(K)}$$
  $$|\hat{v} \circ F^{-1}|_{H^l(K)} \leq c \cdot \|A^{-1}\|_2^l \cdot \sqrt{|\det(A)|} \cdot |\hat{v}|_{H^l(\hat{K})}$$

  für alle $v \in H^l(K)$ bzw. $\hat{v} \in H^l(\hat{K})$ gilt, wobei

  $$|v|_{H^l(K)} := \left(\int_K \sum_{|\alpha| = l} \|\mathcal{D}^\alpha v\|^2 \,\mathrm{d}x \right)^{1/2}$$

  eine Seminorm auf $H^l(K)$ ist.

Def. Sei K ein d-Simplex mit Ecken $a_1, \dots, a_{d+1}$. Wir definieren:

$h(K) := \max_{i,j=1}^{d+1} |a_i - a_j|$ — **Durchmesser**

$\rho(K) := 2 \sup \{ R > 0 \mid \exists x \in K : B_R(x) \subseteq K \}$ — **Innendurchmesser**

$\sigma(K) := h(K)/\rho(K) > 1$ — (misst „Spitzheit")

Lem. Sei der d-Simplex K affin äquivalent zu $\hat{K}$ vermöge $F : \hat{K} \to K$, $\hat{x} \mapsto A\hat{x} + b$, $A \in GL(d)$. Dann gilt:

$$\|A\|_2 \leq h(K)/\rho(\hat{K}), \qquad \|A^{-1}\|_2 \leq h(\hat{K})/\rho(K).$$

Lem. Sei $k \geq 0$. Dann existiert eine Konstante $c > 0$, sodass

$$\inf_{\hat{p} \in \mathbb{P}_k(\hat{K})} \|\hat{v} - \hat{p}\|_{H^{k+1}(\hat{K})} \leq c \cdot |\hat{v}|_{H^{k+1}(\hat{K})} \qquad \forall \hat{v} \in H^{k+1}(\Omega).$$

Satz (Abschätzung des lokalen Interpolationsfehlers)**.** Sei $(\hat{K}, P(\hat{K}), \Sigma(\hat{K}))$ ein finites Element mit $\Sigma(\hat{K}) \subset \mathcal{C}^s(\hat{K})'$. Angenommen, für $r, k \in \mathbb{N}$ mit $0 \leq r \leq k+1$ gilt

$$H^{k+1}(\hat{K}) \hookrightarrow \mathcal{C}^s(\hat{K}) \quad \text{und} \quad \mathbb{P}_k(\hat{K}) \subseteq P(\hat{K}) \subset H^r(\hat{K}).$$

Dann existiert ein $C > 0$, sodass für alle zu $(\hat{K}, P(\hat{K}), \Sigma(\hat{K}))$ affin äquivalente FE $(K, P(K), \Sigma(K))$ und für alle $v \in H^{k+1}(K)$ gilt:

$$|v - \Pi_K v|_{H^r(K)} \leq C \cdot \frac{h(K)^{k+1}}{\rho(K)^r} |v|_{H^{k+1}(K)}.$$

Kor. Es gibt eine Konstante $\tilde{C} > 0$, sodass für jedes zu $(\hat{K}, P(\hat{K}), \Sigma(\hat{K}))$ affin äquivalente FE $(K, P(K), \Sigma(K))$ mit $h(K) \leq 1$ und jedes $v \in H^{k+1}(K_i)$ gilt:

$$\begin{aligned} \|v - \Pi_K v\|_{H^r(K_i)} &\leq \tilde{C} \cdot \frac{h(K_i)^{k+1}}{\rho(K_i)^r} |v|_{H^{k+1}(K_i)} \\ &= \tilde{c}_K \sigma(K_i)^r h(K_i)^{k+1-r} |v|_{H^{k+1}(K_i)}. \end{aligned}$$

Bspe. Das letzte Korollar liefert für FE vom Lagrange-Typ:

	Ordnung	$k=1$	$k=2$
Voraussetzungen	Regularität für v	$H^2(K)$	$H^3(K)$
	Beschränkung für d	$d \leq 3$	$d \leq 5$
	Beschränkung für r	$0 \leq r \leq 2$	$0 \leq r \leq 3$
Konvergenz	$\|v - \Pi v\|_{H^r(K)}$	$\mathcal{O}(h^{2-r})$	$\mathcal{O}(h^{3-r})$

Voraussetzungen. Wir suchen die Lösung von (VGL) im Lösungsraum U mit $H_0^1(\Omega) \subseteq U \subseteq H^1(\Omega)$. Es gelte:

- (V_1) $\overline{\Omega}$ ist ein Polyeder und $\mathcal{T} = (T_n(\overline{\Omega}))_{n \in \mathbb{N}}$ ist eine Familie von Triangulierungen von Ω mit finiten Elementen. Es sei $\mathcal{T}$ **regulär**, d. h. es existiert eine Konstante $\sigma_0 > 0$, sodass $\sigma(K) \leq \sigma_0$ für alle $n \in \mathbb{N}$ und $K \in T_n(\overline{\Omega})$ und es gilt $h_n := \max_{K \in T_n(\overline{\Omega})} h(K) \xrightarrow[n \to \infty]{} 0$.
- (V_2) Alle FE $(K, P(K), \Sigma(K))$ der Familie $\mathcal{T}$ sind affin äquiv. zu einem Referenzelement $(\hat{K}, P(\hat{K}), \Sigma(\hat{K}))$ mit $\Sigma(\hat{K}) \subset \mathcal{C}^s(\hat{K})'$.
- (V_3) $U_n \subset \mathcal{C}(\overline{\Omega})$ ist der FE-Funktionenraum zu $T_n(\overline{\Omega})$.
- (V_4) Für $r, k \in \mathbb{N}$ mit $0 \leq r \leq k+1$ gilt

$$H^{k+1}(\hat{K}) \hookrightarrow \mathcal{C}^s(\hat{K}) \quad \text{und} \quad \mathbb{P}_k(\hat{K}) \subseteq P(\hat{K}) \subset H^r(\hat{K}).$$

Satz. Seien obige Voraussetzungen erfüllt. Sei $\Pi_n : U \to U_n$ der zu $T_n(\overline{\Omega})$ gehörende U_n-*Interpolationsoperator*. Dann existiert ein $c > 0$, sodass für alle $0 \leq l \leq r$, $n \in \mathbb{N}$ und $v \in H^{k+1}(\Omega) \cap U$ gilt:

$$\left(\sum_{K \in T_n(\overline{\Omega})} \|v - \Pi_n v\|_{H^l(K)}^2 \right)^{1/2} \leq c \cdot h_n^{k+1-l} \cdot |v|_{H^{k+1}(\Omega)}.$$

Bemn. Für $l = 0 \leq r$ gilt $U_n \subset L^2(\Omega)$ und

$$\|v - \Pi_n v\|_{L^2(\Omega)} = \left(\sum_{K \in T_n(\overline{\Omega})} \|v - \Pi_K v\|_{L^2(K)}^2 \right)^{1/2} \leq c \cdot h_n^{k+1} \cdot |v|_{H^{k+1}(\Omega)}.$$

Für $l = 1 \leq r$ gilt $U_n \subset H^1(\Omega)$ und

$$\|v - \Pi_n v\|_{H^1(\Omega)} = \left(\sum_{K \in T_n(\overline{\Omega})} \|v - \Pi_K v\|_{H^1(K)}^2 \right)^{1/2} \leq c \cdot h_n^{k} \cdot |v|_{H^{k+1}(\Omega)}.$$

Satz (Konvergenz der konformen Approximation)**.** Seien die Voraussetzungen (V_1) – (V_3) und (V_4) mit $r = 1$ erfüllt. Dann gilt für die Lösung u von (VGL) und die Lösung u_n der konformen Approximation $(\text{VGL})_n$ mit $\ell_n = \ell|_{U_n}$ die Abschätzung

$$\|u - u_n\|_{H^1(\Omega)} \in \mathcal{O}(h_n^k \cdot |u|_{H^{k+1}(\Omega)}) \qquad \text{falls } u \in H^{k+1}(\Omega).$$

Bsp. Für $k = 1$, $u \in H^2(\Omega)$ ist die Konvergenzordnung bloß 1.

Bem. Die konforme Approximation von u hat gemessen mit der L^2-Norm (nach bisherigem Kenntnissstand) im Vergleich zur Interpolation von u eine um eins schlechtere Konvergenzordnung:

$$\begin{aligned} \|u - u_n\|_{L^2(\Omega)} \leq \|u - u_n\|_{H^1(\Omega)} &\in \mathcal{O}(h_n^k \cdot |u|_{H^{k+1}(\Omega)}) \\ \|u - \Pi_n u\|_{L^2(\Omega)} &\in \mathcal{O}(h_n^{k+1} \cdot |u|_{H^{k+1}(\Omega)}) \end{aligned}$$

Dieser Missstand lässt sich mit weiteren Voraussetzungen beheben:

Lem (**Aubin-Nitsche**)**.** Seien U und H Hilberträume, $E : U \to H$ eine stetige inj. Einbettung und $U_n \subset U$ ein endlichdim. Teilraum. Sei $B : U \times U \to \mathbb{R}$ eine stetige, koerzitive Bilinearform und $\ell \in U'$. Seien $u \in U$, $u_n \in U_n$ die Lösungen (VGL) bzw. $(\text{VGL})_n$. Dann gilt

$$\|E(u - u_n)\|_H \leq c_B \cdot \|u - u_n\|_U \cdot \sup_{r \in H \setminus \{0\}} \frac{\inf_{w \in U_n} \|w(r) - w\|_U}{\|r\|_H},$$

wobei $w(r) \in U$ für $r \in H$ die Lösung des **adjungierten Problems**

$$B(\varphi, w(r)) = \langle E(\varphi), r \rangle_H \quad \forall \varphi \in U \qquad \text{ist.}$$

Kor. Seien die Voraussetzungen aus dem Satz zur Konvergenz der konformen Approximation erfüllt. Zusätzlich existiere ein $c_a > 0$, sodass für alle $r \in L^2(\Omega)$ die Lösung $w(r)$ vom adjungierten Problem

$$B(\varphi, w(r)) = \langle \varphi, r \rangle_{L^2(\Omega)} \quad \forall \varphi \in H^1(\Omega)$$

die Abschätzung $\|w(r)\|_{H^2(\Omega)} \leq c_a \cdot \|r\|_{L^2(\Omega)}$ erfüllt. Dann gilt

$$\|u - u_n\|_{L^2(\Omega)} \in \mathcal{O}(h_n^{k+1} \cdot |u|_{H^{k+1}(\Omega)}) \qquad \text{falls } u \in H^{k+1}(\Omega).$$

Rechteckige finite Elemente

Def. Ein **rechteckiges finites Element vom Lagrange-Typ** der Ordnung k ist ein Tupel $(K, P(K), \Sigma(K))$ mit

- $K = [c_1, c_1 + r_1] \times \ldots \times [c_d, c_d + r_d] \subset \mathbb{R}^d$ ist ein Rechteck,
- $P(K) = \mathbb{Q}_k(K) := \{ p(x) = \sum_{\alpha \in \{0, \dots, k\}^d} \lambda_\alpha x_1^{\alpha_1} \cdot \ldots \cdot x_d^{\alpha_d} \} \subset \mathbb{P}_{dk}(K)$
- $\Sigma(K) = \{ b : P(K) \to \mathbb{R},\ p \mapsto p(a) \mid a \in \mathcal{K}_k \}$, wobei $\mathcal{K}_k = \{ (c_1 + i_1 \frac{r_1}{k}, \dots, c_d + i_d \frac{r_d}{k}) \mid i_j \in \{0, \dots, k\},\ j = 1, \dots, d \}$.

Satz. Jedes Polynom $p \in \mathbb{Q}_k(K)$ ist eindeutig durch die Werte auf der Knotenmenge $\mathcal{K}_k$ definiert.

Def. $T_n(\overline{\Omega}) := \{ (K_i, P(K_i), \Sigma(K_i)) \mid i = 1, \dots, N \}$ heißt Triangulierung von $\overline{\Omega}$ mit rechteckigen FE vom Lagrange-Typ, wenn

- $\overline{\Omega} = K_1 \cup \ldots \cup K_N$,
- $\operatorname{int}(K_i) \cap \operatorname{int}(K_j) = \emptyset$ für $i \neq j$,
- Jede Seite von K_i ist entweder eine Teilmenge von $\partial\Omega$ oder die Seite von genau einem anderen K_j.

Def. Der *Finite-Element-Raum* zur Triangulierung $T_n(\overline{\Omega})$ ist

$$U_n := \{ v \in \mathcal{C}(\overline{\Omega}) \mid v|_{K_i} \in \mathbb{Q}_k(K_i),\ i = 1, \dots, N \}.$$

Lem.
- $U_n \subset H^1(\Omega)$
- Eine Basis von U_n ist durch Polynome $p_j \in U_n$ mit $p_i(a_j) = \partial_{ij}$ für alle $ia_j \in \mathcal{K}_k$ gegeben.

Simpliziale Elemente vom Hermite-Typ

Def. Ein **simpliziales finites Element vom Hermite-Typ** ist ein Tupel $(K, P(K), \Sigma(K))$ mit

- einem d-Simplex K mit Ecken $a_1, \dots, a_{d+1}$,
- $P(K) := \mathbb{P}_3(K)$
- $\Sigma(K) := \{ p \mapsto p(a_i) \mid i = 1, \dots, d+1 \}$
  $\cup \{ p \mapsto p(a_{ijl}) \mid 1 \leq i < j < l \leq d+1,\ a_{ijl} := \frac{1}{3}(a_i + a_j + a_l) \}$
  $\cup \{ p \mapsto \mathcal{D}p(a_i)(a_j - a_i) \mid 1 \leq i \neq j \leq d+1 \}$

Der zugehörige Finite-Elemente-Raum zu einer Triangulierung $T(\overline{\Omega})$ mit simpl. finiten Elementen vom Hermite-Typ ist

$$U_n := \{ v \in \mathcal{C}(\overline{\Omega}) \mid v|_K \in \mathbb{P}_3(K)\, \forall K \in T(\overline{\Omega}) \}$$

Bem. $U_n \subset \mathcal{C}^1(\overline{\Omega})$ gilt (nur) für $d = 1$. $U_n \subset \mathcal{C}^1(\overline{\Omega})$ für $d = 2$ erreicht man mit folgenden Elementen:

- **Argyris-Dreieck**: $(K, P(K), \Sigma(K))$ mit
  - einem 2-Simplex K mit Ecken a_1, a_2, a_3,
  - $P(K) := \mathbb{P}_5(K)$ und
  - $\Sigma(K) := \{ p \mapsto p(a_i) \mid 1 \leq i \leq 3 \} \cup \{ p \mapsto \partial_x p(a_i) \mid 1 \leq i \leq 3 \}$
    $\cup \{ p \mapsto \partial_y p(a_i) \mid 1 \leq i \leq 3 \} \cup \{ p \mapsto \partial_{yy} p(a_i) \mid 1 \leq i \leq 3 \}$
    $\cup \{ p \mapsto \partial_{xx} p(a_i) \mid 1 \leq i \leq 3 \} \cup \{ p \mapsto \partial_{xy} p(a_i) \mid 1 \leq i \leq 3 \}$
    $\cup \{ p \mapsto \mathcal{D}p(\frac{a_i + a_j}{2}) \cdot \nu_{ij} \mid 1 \leq i < j \leq 3 \}$

    (wobei ν_{ij} der äußere Normalenvektor an $(a_i + a_j)/2$ ist)
- **Bell-Dreieck**: $(K, P(K), \Sigma(K))$ mit
  - einem 2-Simplex K mit Ecken a_1, a_2, a_3,
  - $P(K) := \{ p \in \mathbb{P}_5 \mid \frac{\partial p}{\partial \nu_{ij}} \in \mathbb{P}_3(K') \text{ für alle Kanten } K' \text{ von } K \}$
  - $\Sigma(K) := \{ p \mapsto p(a_i) \mid 1 \leq i \leq 3 \} \cup \{ p \mapsto \partial_x p(a_i) \mid 1 \leq i \leq 3 \}$
    $\cup \{ p \mapsto \partial_y p(a_i) \mid 1 \leq i \leq 3 \} \cup \{ p \mapsto \partial_{xx} p(a_i) \mid 1 \leq i \leq 3 \}$
    $\cup \{ p \mapsto \partial_{yy} p(a_i) \mid 1 \leq i \leq 3 \} \cup \{ p \mapsto \partial_{xy} p(a_i) \mid 1 \leq i \leq 3 \}$

Nichtkonforme finite Elemente

Bisher war $B_n = B|_{U_n\times U_n}$. Manchmal ist es schwierig, B exakt auszuwerten und muss daher angenähert werden, etwa wenn B durch Integration definiert ist. Dann hilft folgende Verallgemeinerung des Lemmas von Céa:

Satz (1. Lemma von **Strang**). Sei U ein Hilbertraum und $(U_n)_{n\in\mathbb{N}}$ eine Familie von Unterräumen. Sei $B : U\times U \to \mathbb{R}$ eine beschränkte, koerzitive Bilinearform und $(B_n : U_n \times U_n \to \mathbb{R})_{n\in\mathbb{N}}$ eine *gleichmäßig koerzitive* Familie beschränkter Bilinearformen, d. h.

$$\exists\,\alpha > 0 : \forall n : B_n(u_n, u_n) \geq \alpha \|u_n\|_U^2 \quad \text{für alle } n\in\mathbb{N},\ u\in U_n.$$

Dann existiert eine Konstante $c > 0$, sodass

$$\|u-u_n\|_U \leq c\left(\inf_{v_n\in U_n} \Phi(v_n) + \|\ell-\ell_n\|_{U_n'}\right) \quad \text{wobei}$$

$$\Phi(v_n) \coloneqq \|u-v_n\|_U + \sup_{w_n\in U_n, w_n\neq 0} \tfrac{|B(v_n,w_n)-B_n(v_n,w_n)|}{\|w_n\|_U}$$

Def. Das **Fehlerfunktional der Quadraturformel** mit *Stützpunkten* x_i^K und *Gewichten* w_i^K ist

$$E_K(\psi) \coloneqq \int_K \psi(x)\,\mathrm{d}x - \sum_{i=1}^q w_i^K \psi(x_i^K).$$

Satz (Konvergenzsatz für das Modellproblem). Seien die Voraussetzungen (V_1) – (V_3) erfüllt, $P(\hat{K}) = \mathbb{P}_k(\hat{k})$ mit $k > d/2$ (damit $H^{k+1}(\Omega) \hookrightarrow \mathcal{C}^s(\hat{K})$). Sei die Quadraturformel so gewählt, dass $E_{\hat{K}}(p) = 0$ für alle $p\in\mathbb{P}_{2k-2}(\hat{K})$. Für die Lösung $u\in H_0^1(\Omega)$ von (VGL) mit $B(u,\varphi) \coloneqq \int \mathcal{D}u\cdot\mathcal{D}\varphi\,\mathrm{d}x$, $\ell(\varphi)\coloneqq\int f\varphi\,\mathrm{d}x$ mit $f\in H^k(\Omega)$ gelte $u\in H^{k+1}(\Omega)$. Dann gilt:

- Es gibt eine Konstante $C_1 > 0$, sodass $\forall\, p,q\in\mathbb{P}_k(K)$:
$$E_K(\mathcal{D}p\cdot\mathcal{D}q) \leq C_1\cdot h^k(K)\cdot\|p\|_{H^k(K)}\cdot\|q\|_{H^1(K)}$$
- Es gibt eine Konstante $C_2 > 0$, sodass $\forall\ f\in H^k(K)$, $\varphi\in H^1(K)$:
$$E_K(f\varphi) \leq C_2\cdot h^k(K)\cdot\|f\|_{H^k(K)}\cdot\|\varphi\|_{H^1(K)}$$
- Für die Lsg $u_n\in U_n$ von $(\mathrm{VGL})_n$ mit der nicht-konformen Approx. $B_n(u,\varphi)$ und $\ell_n(\varphi)$ mittels der Quadraturformel gilt:
$$\|u-u_n\|_{H^1(\Omega)} \leq C\cdot\left(\|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^k(\Omega)}\right)$$

Def. Sei $\Omega\subset\mathbb{R}^d$ ein Gebiet mit glattem Rand. Dann heißt $(\Omega_n, T(\overline{\Omega}_n))$ **zulässige Gebietsapproximation**, falls $T_n(\overline{\Omega}_n)$ eine zulässige Triangulierung von Ω_n ist und

$$\partial\Omega_n \cap \{\text{ Eckpunkte der Simplizes in } T(\overline{\Omega}_n)\ \} \subset \partial\Omega.$$

Wir betrachten nun

Problem.

$$(\mathrm{VGL_n})\quad \begin{cases} \overbrace{\int_{\Omega_n} \mathcal{D}u_n\cdot\mathcal{D}\varphi_n\,\mathrm{d}x}^{B_n(u_n,\varphi_n)\coloneqq} = \overbrace{\int_{\Omega_n}\tilde{f}\varphi_n\,\mathrm{d}x}^{\ell_n(\varphi_n)\coloneqq} \quad \forall\,\varphi\in U_n \\ u_n\in U_n \end{cases}$$

wobei $\tilde{f}\in L^2(\Omega_n\cup\Omega)$ mit $\tilde{f}|_\Omega = f$

Isoparametrische Finite Elemente

Def. Sei $(\hat{K},\hat{P},\hat{\Sigma})$ ein gerades (simpliziales oder rechteckiges) FE vom Lagrange-Typ. Ein finites Element (K,P,Σ) vom Lagrange-Typ (k) heißt **isometrisch äquivalent** zu $(\hat{K},\hat{P},\hat{\Sigma})$, falls eine umkehrbare Abbildung

$$F:\hat{K}\to K,\quad \hat{x}\mapsto(F_1(\hat{x}),\dots,F_d(\hat{x}))$$

mit $F_1,\dots,F_d\in\hat{P}$ existiert, sodass

- $K = F(\hat{K})$
- $P = \{p = \hat{p}\circ F^{-1} \mid \hat{p}\in\hat{P}\}$
- $\Sigma = \{p\mapsto p(a_i) \mid$ für Punkte $a_i \coloneqq F(\hat{a}_i), \hat{a}_i\in\mathcal{K}_k\}$

Bem. Die zu (K,P,Σ) isoparametrische Abbildung F ist für isoparametrische FE vom Lagrange-Typ (k) eindeutig durch die Werte an den Knoten $\hat{a}_i$ definiert.

Finite Elemente für parabolische Probleme

Problem. Wärmeleitungsgleichung:

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) &= f(x,t) & \text{in } \Omega\times(0,T) \\ u(x,0) &= u_0(x) & \text{für } x\in\Omega \\ u(x,t) &= 0 & \text{für } x\in\partial\Omega, t\in[0,T] \end{cases}$$

Variationsgleichung: Gesucht ist $u\in\mathcal{C}^1([0,T]\,;H_0^1(\Omega))$ mit

$$\int_\Omega u_t(x,t)\varphi(x)\,\mathrm{d}x = \int_\Omega \Delta u(x,t)\varphi(x)\,\mathrm{d}x + \int_\Omega f(x,t)\varphi(x)\,\mathrm{d}x \quad \forall t\in[0,T]$$

$$= -\int_\Omega \mathcal{D}u(x,t)\cdot\mathcal{D}\varphi(x)\,\mathrm{d}x + \int_\Omega f(x,t)\varphi(x)\,\mathrm{d}x \quad \forall\varphi\in H_0^1(\Omega)$$

und Anfangsbedingung $u(0) = u_0$. Anders geschrieben:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\langle u(t,-),\varphi\rangle + B(u,\varphi) - \langle f(t,-),\varphi\rangle \qquad \forall\varphi\in H_0^1(\Omega).$$

Verfahren (**Vertikale Linienmethode**).
Idee: Erst Diskretisierung im Ort, dann Diskretisierung in der Zeit
Wir wählen dazu einen endlichdim. Teilraum $U_n\subset U = H_0^1(\Omega)$ und betrachten dann die approximierte VGL

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\langle u_n(t),\varphi_n\rangle + B_n(u_n,\varphi_n) = \langle f(t,-),\varphi_n\rangle \quad \forall\,\varphi_n\in U_n$$

Sei $\varphi_1,\dots,\varphi_{d_n}$ eine Basis von U_n. Wir schreiben

$$u_n(t,x) = \textstyle\sum_{j=1}^{d_n}\gamma_j(t)\varphi_j(x).$$

Sei $u_n^0 = \sum\gamma_j^0\varphi_j$ eine Approximation von u_0. Wir erhalten die GGL

$$\begin{cases} \sum_{i=1}^{d_n}\tfrac{\mathrm{d}}{\mathrm{d}t}\gamma_i(t)\langle\varphi_j,\varphi_i\rangle + \sum_{i=1}^{d_n}\gamma_i(t)B_n(\varphi_j,\varphi_i) &= \langle f(t,-),\varphi_j\rangle\ \forall\, j, \\ \gamma_j(0) &= \gamma_j^0 \quad \forall\, j \end{cases}$$

mit $\gamma(t)\in\mathbb{R}^{d_n}$. Wir definieren die *Massenmatrix* $M\in\mathbb{R}^{d_n\times d_n}$, die Matrix $B\in\mathbb{R}^{d_n\times d_n}$ und $g:[0,T]\to\mathbb{R}^{d_n}$ durch

$$M_{ij} = \int_\Omega\varphi_j(x)\varphi_i(x)\,\mathrm{d}x,\quad B_{ij} = \int_\Omega\mathcal{D}\varphi_j\cdot\mathcal{D}\varphi_i\,\mathrm{d}x,\quad g_j(t) = \int_\Omega f(t,x)\varphi_j(x)\,\mathrm{d}x$$

Dann können wir obige GGL wie folgt umschreiben:

$$\begin{cases} M\tfrac{\mathrm{d}}{\mathrm{d}t}\gamma(t) + B\gamma(t) &= g(t) \\ \gamma(0) &= \gamma^0 \end{cases}$$

Durch Modellreduktion erhält man eine Gleichung

$$\begin{cases} \tilde{M}\tfrac{\mathrm{d}}{\mathrm{d}t}\tilde{\gamma}(t) + \tilde{B}\tilde{\gamma}(t) &= \tilde{g}(t) \\ \tilde{\gamma}(0) &= \tilde{\gamma}^0 \end{cases}$$

sodass $\gamma\in\mathbb{R}^r$, $r < d_n$ und $\gamma\approx V\tilde{\gamma}$.

Verfahren (**Horizontale Linienmethode**, *Rothe-Methode*).
Idee: Erst Diskretisierung in der Zeit, dann Diskretisierung im Ort
Sei $\tau\coloneqq T/q$, $t_i\coloneqq i\cdot\tau$ und

$$\psi_i(t)\coloneqq\begin{cases} t-t_{i-1}/\tau & \text{für } t\in[t_{i-1},t_i]\,, \\ t_i-t/\tau & \text{für } t\in[t_i,t_{i+1}]\,, \\ 0 & \text{sonst.} \end{cases}$$

Wir verwenden die Approx. $u(x,t)\approx u_\tau(x,t) = \sum_{j=1}^q\psi_j(t)c_j(x)$, also $u_\tau(x,t_i) = c_i(x)$. Implizites Euler-Verfahren:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}u_\tau(x,t_{i+1}) \approx \tfrac{u_\tau(x,t_{i+1})-u_\tau(x,t_i)}{\tau} = \tfrac{c_{i+1}(x)-c_i(x)}{\tau},$$

Dies führt pro Zeitpunkt t_i zu je einer VGL

$$\left\langle\tfrac{c_{i+1}-c_i}{\tau},\varphi\right\rangle + B(c_{i+1},\varphi) = \langle f(t_{i+1}),\varphi\rangle\ \ \forall\,\varphi \quad \text{bzw.}$$

$$\underbrace{\int_\Omega c_{i+1}\varphi\,\mathrm{d}x + \tau\int_\Omega\mathcal{D}c_{i+1}\cdot\mathcal{D}\varphi\,\mathrm{d}x}_{\hat{B}(c_{i+1},\varphi)} = \underbrace{\tau\int_\Omega f(t_{i+1},x)\varphi(x)\,\mathrm{d}x + \int_\Omega c_i\varphi\,\mathrm{d}x}_{\hat{\ell}(\varphi)} \quad \forall\,\varphi.$$

Diese VGLn lösen wir dann mittels der FEM iterativ für $i = 1,\dots,q$.