NMPDE/ATSC 2025
Lecture 5

The Galerkin Method (Q, chap 4)

Given $(V, (\cdot,\cdot))$ Hilbert, $\mathcal{A}(\cdot,\cdot)$, $F(\cdot)$
$\mathcal{A}$: bilinear, F: linear

(WP) Find $u \in V$: $\mathcal{A}(u,v) = F(v) \quad \forall v \in V$

Galerkin:

- consider $V_n \subseteq V$ (subspace), $\dim V_n = n$
- restrict (WP) to V_n:

(GM) Find $u_n \in V_n$: $\mathcal{A}(u_n, v_n) = F(v_n) \quad \forall v_n \in V_n$

- If V is separable $\rightarrow$ $\exists$ orthonormal basis,

$\rightarrow$ construct $V_n = \mathrm{span}\langle \varphi_1, \dots, \varphi_n \rangle$

- If (WP) is well-posed (Lax-Milgram applies) then also the resctricted problem still well-posed, so $\exists!\, u \in V_n$ solution of (GM), $\|u_n\|_V \le 2_0 \|F\|_{V'}$

- Question: $u_n \xrightarrow{n \to +\infty} u$

- Suppose $\mathcal{A}$ symmetric, then

u = minimizer over V of $J(v) = \frac{1}{2}\mathcal{A}(v,v) - F(v)$

also

u_n = minimizer over V_n of ″

- Algebraic point of view:

$$u_n(x) = \sum_{j=1}^{n} U_j \, \varphi_j(x)$$

↪ unknowns

(GM) $\Leftrightarrow$ Find $U = \{U_j\}_{j=1}^{n} \in \mathbb{R}^n$:

test with $\varphi_i \; \forall i$

$$\boxed{\sum_{j=1}^{n} U_j \, \underbrace{\mathcal{A}(\varphi_j, \varphi_i)}_{= A_{ij}} = \underbrace{F(\varphi_i)}_{= F_i} \quad \forall i = 1, \dots, n}$$

$\Leftrightarrow$ $\boxed{AU = F}$

- $\mathcal{A}$ coercive $\Rightarrow$ A positive definite $\Rightarrow$ invertible

- $\mathcal{A}$ symetric $\Rightarrow$ A symetric

and $u_h \in V_h$ is the sol. of (GM) $\Leftrightarrow$ U solves $AU = F$

$\Updownarrow$ $\qquad\qquad$ $\Updownarrow$

$J(u_h) \leq J(v) \quad \forall v \in V_h$ $\qquad$ U minimizer of $\phi(V) = \frac{1}{2} V^T A V - V^T F$

useful to devise solver techniques for the linear system

Analysis of GM

Assume Lax-Milgram hypothesis apply

($\mathcal{A}$ is coercive (α_0), continuous (γ)
F continuous)

Stability: $\|u_n\|_V \le \alpha_0 \|F\|_{V'}$

Consistency (Full):

Theorem (Galerkin Orthogonality):

If $u \in V$ solves (WP), $V_n \subseteq V$ subspace
$u_n \in V_n$ solves (GM), then

$$\rightarrow \boxed{\mathcal{A}(u-u_n, v_n) = 0} \qquad \forall v \in V_n$$

Proof: take $v_n \in V_n$,

- $\mathcal{A}(u, v_n) = F(v_n)$ $\qquad (\Leftarrow (WP))$
- $\mathcal{A}(u_n, v_n) = F(v_n)$ $\qquad (\Leftarrow (GM))$

Convergence :

Lemma of Cèa : Under some assumptions

$$\boxed{\|u-u_n\|_V \leq \frac{\gamma}{\alpha_0} \inf_{v_n \in V_n} \|u-v_n\|_V}$$

QUASI OPTIMALITY

Proof :

$$\alpha_0 \|u-u_n\|_V^{\cancel{2}} \overset{\text{coercivity}}{\leq} a(u-u_n, u-\underset{\in V_n}{u_n})$$

$$\overset{(G.O.)}{=} a(u-u_n, u)$$

$$\overset{(G.O.)}{=} a(u-u_n, u-v_n) \qquad \forall v_n \in V_n$$

$$\overset{\text{Continuity}}{\leq} \gamma \cancel{\|u-u_n\|_V} \|u-v_n\|_V \qquad \square$$

Convergence follows from Cèa if

$$\lim_{n \to +\infty} \inf_{v_n \in V_n} \|u-v_n\| = 0$$

True if $\{\varphi_j\}_1^\infty$ used to construct V_n is a complete orthonormal system or $\lim_n V_n$ is dense in V, so

$\forall u \in V, \ \forall \varepsilon > 0, \exists n: \ \exists w_n \in V_n : \|u - w_n\|_V < \varepsilon$

Proposition (Céa sym. case):
If, moreover, $\mathcal{A}$ is symmetric, then

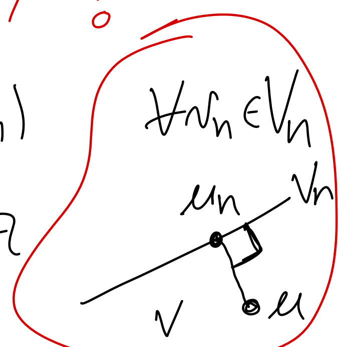
- $\boxed{\|u - u_n\|_V \le \sqrt{\frac{\gamma}{\alpha_0}} \min_{v_n \in V_n} \|u - v_n\|_V}$

- Define $(\cdot,\cdot)_{\mathcal{A}} = \mathcal{A}(\cdot,\cdot) \to \|v\|_{\mathcal{A}} = \sqrt{\mathcal{A}(\cdot,\cdot)}$

$$\|u - u_n\|_{\mathcal{A}} = \min_{v_n \in V_n} \|u - v_n\|_{\mathcal{A}}$$

OPTIMALITY !

Proof: (GO): $0 = \mathcal{A}(u - u_n, v_n)$
$= (u - u_n, v_n)_{\mathcal{A}}$



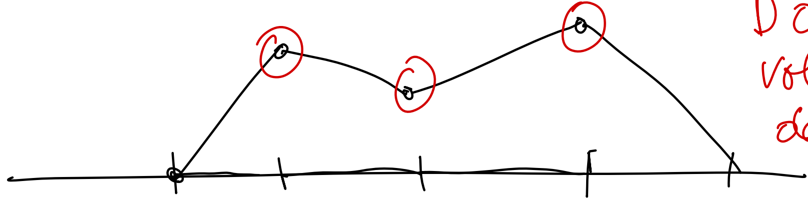
$$\Rightarrow \|u-u_h\|_a^2 = \min_{v_h\in V_h} \|u-v_h\|_a^2 = \min_{v_h\in V_h} a(u-v_h, u-v_h) \underset{\text{cont}}{\le} \gamma \min_{v_h\in V_h} \|u-v_h\|_V^2$$

$$\|u-u_h\|_a^2 = a(u-u_h, u-u_h) \ge \alpha_0 \|u-u_h\|_V^2$$

FEM in 1D

Recall $V_h \subset V = H_0^1(0,1)$

piecewise linear functions



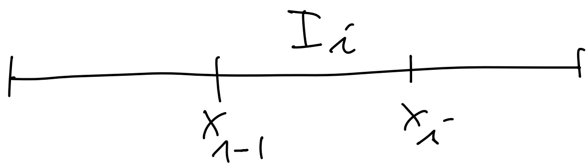
DoF = values defining the FE function

nodes of the partition $\Rightarrow$ FE function

Generalise idea to order k piecewise polynomials, $k \geq 1$:

- Fix partition $0 = x_0 < x_1 < \dots < x_M = 1$

  $x_i = x_{i-1} + h_i, \quad i = 1, \dots, M$

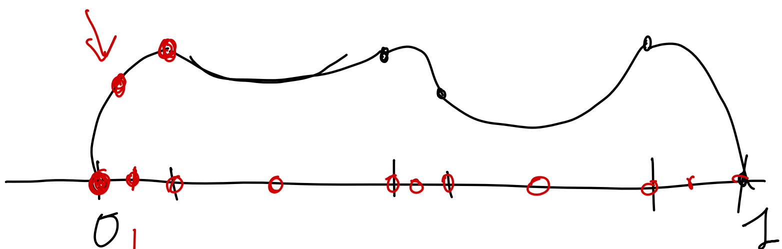


- $h = \max_i h_i$

- $$V_h^k = \left\{ v \in C_0(\underbrace{\Omega}_{=(0,1)}) : v|_{I_i} \in \mathbb{P}^k(I_i) \ \forall i, \ v(0) = 0 = v(1) \right\} \subset H_0^1$$

- FEM(k): Find $u_h \in V_h^k$: $A(u_h, v_h) = F(v_h)$ $\forall v_h \in V_h$

$k = 2$



quadratic

FE nodes $\neq$ Grid nodes

Assessing numerically the rate of convergence

Suppose that an algebric rate of convergence is expected as in:

$$|u - u_h| \leq C\, h^{\alpha}$$

Then, considering a problem for which the exact solution is available, we can run a series of experiment to evaluate the "experimental rate of convergence" (EOC) as follows. We compute the numerical solution with discretisation parameter h and then $h/2$, for which we expect, respectively

$$|u - u_h| \sim h^{\alpha} \quad , \quad |u - u_{h/2}| \sim \left(\frac{h}{2}\right)^{\alpha}$$

$$\Rightarrow \quad \frac{|u - u_h|}{|u - u_{h/2}|} \sim \frac{h^{\alpha}}{\left(\frac{h}{2}\right)^{\alpha}} \sim 2^{\alpha} \quad \Rightarrow \quad \log \frac{|u - u_h|}{|u - u_{h/2}|} \sim \log 2^{\alpha} = \alpha \log 2$$

$$\Rightarrow \quad \boxed{\alpha \sim \frac{\log |u - u_h| - \log |u - u_{h/2}|}{\log 2}}$$