Why the finite element method works

calculations for week 14

Ed Bueler

UAF Math 617 Functional Analysis

Spring 2026

Outline

1. recall: Poisson problems and their finite element (FE) approximations
2. abstract bilinear form notation
3. Céa's lemma: the FE approximation is (essentially) the best
4. polynomial and piecewise-linear interpolation on $\mathbb{R}^1$
5. convergence of the FE approximation as $h \to 0$
6. lecture content in week 14

recall: the Sobolev spaces $H^1(\Omega)$ and $H_0^1(\Omega)$

Definition

$$H^1(\Omega) = \left\{ f \in L^2(\Omega) \,\middle|\, \begin{array}{c}\text{the weak derivative } f' \text{ exists} \\ \text{and is in } L^2(\Omega)\end{array} \right\}$$

is a Hilbert space with inner product and norm:

$$\langle f, g \rangle_{H^1} = \int_\Omega f(x) g(x)\, dx + \int_\Omega \nabla f(x) \cdot \nabla g(x)\, dx, \qquad \|f\|_{H^1} = \sqrt{\langle f, f \rangle_{H^1}}$$

Definition

$$H_0^1(\Omega) = \overline{C_c^1(\Omega)},$$

with closure in the H^1 norm

- $H_0^1(\Omega) \subset H^1(\Omega)$ is the closed subspace of functions which have zero boundary values along $\partial\Omega$

recall: the Poincaré-Friedrichs inequality

Theorem (Poincaré-Friedrichs inequality)

Let $\Omega \subset \mathbb{R}^d$ be an open, bounded domain with Lipschitz boundary. There is $C > 0$, depending only on Ω, so that

$$\|f\|_{L^2} \le C \|\nabla f\|_{L^2} \quad \text{for all } f \in H_0^1(\Omega).$$

- a main consequence of this inequality is that the simpler formula

$$|f|_{H^1} = \|\nabla f\|_{L^2} = \left(\int_\Omega |\nabla f|^2\right)^{1/2}$$

  is equivalent to the full H^1 norm on $H_0^1(\Omega)$
- the full norm is $\|f\|_{H^1} = \left(\int_\Omega |f|^2 + \int_\Omega |\nabla f|^2\right)^{1/2}$
- ... I still owe you a proof of this inequality

recall: the weak form of example PDE problems

- I'll use these two example elliptic PDEs to illustrate the FE theory

Poisson equation weak form (homogeneous Dirichlet conditions)

Let $V = H_0^1(\Omega)$. Suppose $f \in L^2(\Omega)$ is given. Find $u \in V$ so that

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx \quad \forall v \in V$$

strong form Poisson: $\boxed{-\nabla^2 u = f}$

Helmholtz equation weak form (general Neumann conditions)

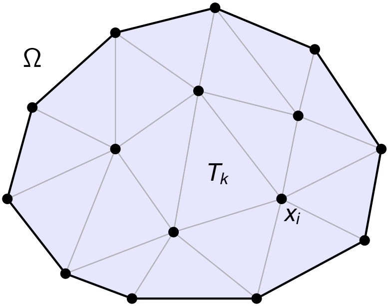
Let $V = H(\Omega)$. Suppose $f \in L^2(\Omega)$ and $g_N \in L^2(\partial\Omega)$ is given and $\gamma > 0$. Find $u \in V$ so that

$$\int_\Omega \nabla u \cdot \nabla v \, dx + \gamma \int_\Omega u \cdot v \, dx = \int_\Omega f v \, dx + \int_{\partial\Omega} g_N v \, ds \quad \forall v \in V$$

strong form Helmholtz: $\boxed{-\nabla^2 u + \gamma u = f}$

recall: the piecewise-linear finite element space

- suppose $\Omega \subset \mathbb{R}^2$ is a polygon
- put a mesh $\mathcal{T}$ of triangles on Ω
- denote nodes x_i and triangles T_k
- $f(x, y) = a + bx + cy$ is called a *linear polynomial* in two variables



Definition (the continuous P_1 FE space over $\mathcal{T}$)

$$P_1 = \left\{ f \in C(\Omega) \,:\, f\big|_{T_k} \text{ is a linear polynomial} \right\}$$

- key idea: $P_1 \subset H^1(\Omega)$ ... I owe you a proof of that, too

recall: the basic idea of the finite element method

weak form (here for illustration: Poisson equation)

Find $u \in V$ so that

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx \quad \forall v \in V$$

- FE method solves same weak form, but over a finite-dim'l subspace V_h

finite element method

Find $u_h \in V_h$ so that

$$\int_\Omega \nabla u_h \cdot \nabla v_h \, dx = \int_\Omega f v_h \, dx \quad \forall v_h \in V_h$$

- for Dirichlet conditions:

$$V_h = \left\{ f \in P_1 \,:\, f\big|_{\partial\Omega} = 0 \right\} \subset V = H_0^1(\Omega)$$

- for Neumann conditions: $V_h = P_1 \subset V = H^1(\Omega)$

Outline

1. recall: Poisson problems and their finite element (FE) approximations
2. abstract bilinear form notation
3. Céa's lemma: the FE approximation is (essentially) the best
4. polynomial and piecewise-linear interpolation on $\mathbb{R}^1$
5. convergence of the FE approximation as $h \to 0$
6. lecture content in week 14

abstract bilinear notation

- we write

$$a : V \times V \to \mathbb{R}$$

for the bilinear function on the left side of the weak form

$$\begin{aligned} a(cu + dv, w) &= c\, a(u, w) + d\, a(v, w) \\ a(u, cv + dw) &= c\, a(u, v) + d\, a(u, w) \end{aligned}$$

- thus $a(u, v) \in \mathbb{R}$ if $u, v \in V$
- for now we do *not* assume it is symmetric (i.e. $a(u, v) = a(v, u)$)

examples: bilinear forms

- Poisson equation: $a(u, v) = \int_\Omega \nabla u \cdot \nabla v\, dx$
- Helmholtz equation: $a(u, v) = \int_\Omega \nabla u \cdot \nabla v\, dx + \gamma \int_\Omega uv\, dx$

abstract linear notation

- we write
$$\ell : V \to \mathbb{R}$$
for the linear function on the right side of the weak form
$$\ell(cv + dw) = c\,\ell(v) + d\,\ell(w)$$
- thus $\ell(v) \in \mathbb{R}$ if $v \in V$
- we will assume ℓ is continuous: $\ell \in V'$ is in the dual space

examples: linear right-hand sides

- Poisson equation w homogeneous Dirichlet: $\ell(v) = \int_\Omega fv\,dx$
- Helmholtz eqn w general Neumann: $\ell(v) = \int_\Omega fv\,dx + \int_{\partial\Omega} g_N v\,ds$

recall: minimization formulation leads to well-posedness

- this notation was already used in a theorem

Theorem (well-posedness)

For V a Hilbert space, suppose $a : V \times V \to \mathbb{R}$ is a symmetric bilinear function. Assume it is coercive: there is $\alpha > 0$ so that

$$a(v,v) \geq \alpha \|v\|^2 \quad \forall v \in V.$$

Assume it is continuous: there is $C \geq 0$ so that

$$|a(u,v)| \leq C\|u\|\|v\| \quad \forall u,v \in V.$$

Suppose $\ell \in V'$. Then there exists $u \in V$ which is the unique minimizer of J:

$$u \quad \overset{\min}{\leftarrow} \quad J(v) = \frac{1}{2} a(v,v) - \ell(v).$$

Furthermore $u \in V$ solves

$$a(u,v) = \ell(v) \quad \forall v \in V.$$

basic idea of the finite element method (again)

- now we have very clean notation for the basic FE idea

weak form

Find $u \in V$ so that

$$a(u, v) = \ell(v) \quad \forall v \in V$$

- FE method solves same weak form, but over a finite-dim'l subspace V_h

finite element method

Find $u_h \in V_h$ so that

$$a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h$$

Outline

1. recall: Poisson problems and their finite element (FE) approximations
2. abstract bilinear form notation
3. Céa's lemma: the FE approximation is (essentially) the best
4. polynomial and piecewise-linear interpolation on $\mathbb{R}^1$
5. convergence of the FE approximation as $h \to 0$
6. lecture content in week 14

conforming FE methods

Definition

we say that an FE method is *conforming* if

- $V_h \subset V$, and
- $a(\cdot, \cdot)$ and $\ell(\cdot)$ used in the FE method are the same as those in the original PDE weak form

- recall the problems together:

$$\text{find } u \in V \text{ so that} \quad a(u, v) = \ell(v) \quad \forall v \in V$$
$$\text{find } u_h \in V_h \text{ so that} \quad a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h$$

Definition

the function $u - u_h$ is the FE *error*

- in a conforming FE method we also know:

$$u - u_h \in V$$
$$a(u, v_h) = \ell(v_h) \quad \forall v_h \in V_h$$

Galerkin orthogonality

- in a conforming FE method we have:

$$a(u, v_h) = \ell(v_h) \quad \forall v \in V_h$$
$$a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h$$

- subtract and use bilinear:

$$a(u - u_h, v_h) = a(u, v_h) - a(u_h, v_h) = 0 \quad \forall v_h \in V_h$$

Galerkin orthogonality

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in V_h,$$

that is, the error $u - u_h$ is orthogonal to the entire FE space V_h

picture:

assumptions about the weak form PDE

assumptions for next slide

- $\ell \in V'$ is linear and continuous
- $a : V \times V \to \mathbb{R}$ is a bilinear function
- a is coercive (elliptic): there is $\alpha > 0$ so that

$$a(v,v) \geq \alpha \|v\|^2 \qquad \forall v \in V$$

- a is continuous (bounded): there is $M \geq 0$ so that

$$|a(u,v)| \leq M \|u\| \|v\| \qquad \forall u, v \in V$$

- same assumptions as in the minimization approach
- note $\alpha \leq M$ … *why?*
- we'll check these assumptions in examples

quasi-optimality, often called Cea's lemma

Lemma (quasi-optimality)

Make the assumptions on the previous slide. Then

$$\|u - u_h\| \le \frac{M}{\alpha}\|u - v_h\| \quad \forall v_h \in V_h.$$

Thus, up to a factor of M/α, u_h is closer to the exact solution than any other element of V_h.

Proof. Let $v_h \in V_h$ be arbitrary. By coercivity, bilinearity, Galerkin orthogonality, and continuity:

$$\begin{aligned}
\alpha\|u - u_h\|^2 &\le a(u - u_h, u - u_h) \\
&= a(u - u_h, u - v_h + v_h - u_h) = a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\
&= a(u - u_h, u - v_h) + 0 \\
&= M\|u - u_h\|\|u - v_h\|
\end{aligned}$$

Dividing by $\|u - u_h\|$ gives $\alpha\|u - u_h\| \le M\|u - v_h\|$, as claimed. □

example: homogeneous Dirichlet problem for Poisson equation

- here $V = H_0^1(\Omega)$, with norm $\|v\|_{H^1} = \left(\|v\|_{L^2} + \|\nabla v\|_{L^2}\right)^{1/2}$
- ℓ is clearly linear, and bounded via CS, and a is clearly bilinear
- coercivity is from Poincaré-Friedrichs: $\|\phi\|_{L^2} \le C\|\nabla\phi\|_{L^2}$
- specifically, we may take $\alpha = \frac{1}{2}\min\{1, 1/C^2\}$:

$$\begin{aligned} a(v,v) = \int_\Omega |\nabla v|^2\,dx = \|\nabla v\|_{L^2}^2 &\ge \frac{1}{2}\|\nabla v\|_{L^2}^2 + \frac{1}{2}\|\nabla v\|_{L^2}^2 \\ &\ge \frac{1}{2}\|\nabla v\|_{L^2}^2 + \frac{1}{2C^2}\|v\|_{L^2}^2 \ge \alpha\|v\|_{H^1}^2 \end{aligned}$$

- a is continuous via Cauchy-Schwarz, with $M = 1$:

$$|a(u,v)| \le \int_\Omega |\nabla u||\nabla v|\,dx \le \|\nabla u\|_{L^2}\|\nabla v\|_{L^2} \le \|u\|_{H^1}\|v\|_{H^1}$$

Corollary (quasi-optimality in H^1 norm)

$$\|u - u_h\|_{H^1} \le \frac{1}{\alpha}\|u - v_h\|_{H^1} \quad \forall v_h \in V_h$$

example: homogeneous Dirichlet problem for Poisson equation 2

- we can do better in this case!
- switch the norm on $V = H_0^1(\Omega)$ to $|\!|\!| v |\!|\!| = \|\nabla v\|_{L^2}$
  - Poincaré-Friedrichs says this is a norm
  - called the *energy norm*, by poorly-motivated traditional reasons
- now a is coercive with $\alpha = 1$,

$$a(v, v) = \int_\Omega |\nabla v|^2\, dx = |\!|\!| v |\!|\!|^2 \geq 1\, |\!|\!| v |\!|\!|^2,$$

- it is easy to see that a is continuous with $M = 1$
- the quasi-optimality lemma becomes ...

Corollary (optimality in energy norm)

$$|\!|\!| u - u_h |\!|\!| \leq |\!|\!| u - v_h |\!|\!| \quad \forall v_h \in V_h$$

- this is actual optimality in this norm

example: general Neumann problem for Helmholtz equation

- recall Helmholtz: $a(u,v) = \int_\Omega \nabla u \cdot \nabla v\, dx + \gamma \int_\Omega uv\, dx$
- here $V = H^1(\Omega)$, with usual norm $\|v\|_{H^1} = \left(\|v\|_{L^2} + \|\nabla v\|_{L^2}\right)^{1/2}$
- coercivity does not need PF
- in fact, $\alpha = \min\{1,\gamma\}$:

$$\begin{aligned} a(v,v) &= \int_\Omega |\nabla v|^2\, dx + \gamma \int_\Omega |v|^2\, dx \\ &\geq \alpha \left(\int_\Omega |\nabla v|^2\, dx + \int_\Omega |v|^2\, dx\right) = \alpha \|v\|_{H^1}^2 \end{aligned}$$

- a is continuous via Cauchy-Schwarz, with $M = \max\{1,\gamma\}$:

$$|a(u,v)| \leq \int_\Omega |\nabla u||\nabla v|\, dx + \gamma \int_\Omega |u||v|\, dx \leq M \|u\|_{H^1} \|v\|_{H^1}$$

Corollary (quasi-optimality in H^1 norm)

$$\|u - u_h\|_{H^1} \leq \frac{\max\{1,\gamma\}}{\min\{1,\gamma\}} \|u - v_h\|_{H^1} \quad \forall v_h \in V_h$$

how will this show that the FE method converges?

Lemma (quasi-optimality)

$$\|u - u_h\| \le \frac{M}{\alpha}\|u - v_h\| \quad \forall v_h \in V_h.$$

- up to a factor, u_h is closer to the exact solution than any other $v_h \in V_h$
- how does this show that the FE method converges?
- it is not obvious!
- it requires notation, explanation, and additional theory

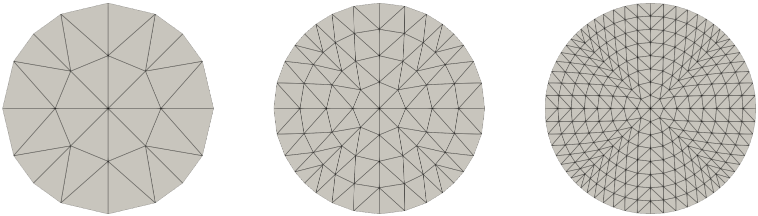
h = mesh resolution

- from now on, $h > 0$ denotes the mesh resolution, for example,

$$h = \max\{\text{diam}(T_k)/2 \,:\, T_k \in \mathcal{T}\}$$

- picture of T_k and its diameter:

- we consider a sequence of refined meshes, for which $h \to 0$



how will quasi-optimality show that the FE method converges?

- let us assume about the original, continuous PDE problem:
  1. the weak form is well-posed and determines a unique solution u
  2. the exact solution u is a smooth function on Ω
- in the problems so far we know $u \in H^1(\Omega)$, but it is often true that $u \in H^2(\Omega)$ or $u \in C^2(\Omega)$ or even $u \in C^\infty(\Omega)$
  - a *regularity* theorem would supply such knowledge
- now we make a definition related to the FE problem

Definition

on a given mesh $\mathcal{T}$ over Ω, with a given value $h > 0$, there is an interpolation map

$$\mathbb{I}_h : C(\overline{\Omega}) \to P_1$$

which takes the nodal values of the input $w \in C(\overline{\Omega})$ and constructs the element of P_1 with those nodal values

how will quasi-optimality show that the FE method converges?

- we need an *interpolation theorem* to say how close $\mathbb{I}_h w$ is to w

generic interpolation theorem

$$\|w - \mathbb{I}_h w\| = O(h^k)$$

  - actual theorems later in these slides, for 1D intervals and 2D triangles
  - such theorems provide a rate $O(h^k)$ for interpolation accuracy, given $w \in H^\ell(\Omega)$, for $\Omega \subset \mathbb{R}^d$, with relationships between k, ℓ, d
- then we substitute $v_h = \mathbb{I}_h u$ into quasi-optimality and use the rate:

$$\|u - u_h\| \le \frac{M}{\alpha} \|u - \mathbb{I}_h u\| = O(h^k)$$

- thus $u \to u_h$ as $h \to 0$, so the method converges

before moving on . . . Cea's actual lemma is better!

Lemma (Cea's lemma . . . from his 1964 PhD thesis)

Make the same assumptions about our weak form, and also assume that the bilinear form is symmetric $a(u, v) = a(v, u)$. *Then*

$$\|u - u_h\| \le \sqrt{\frac{M}{\alpha}}\|u - v_h\|$$

for all $v_h \in V_h$. *(The constant* $\sqrt{M/\alpha}$ *is smaller than* M/α*, thus improved relative to the original lemma.)*

Proof. Let $v_h \in V_h$ be arbitrary. By bilinearity, then Galerkin orthogonality, and then $a(u_h - v_h, u_h - v_h) \ge 0$ from coercivity:

$$\begin{aligned} a(u - v_h, u - v_h) &= a(u - u_h + u_h - v_h, u - u_h + u_h - v_h) \\ &= a(u - u_h, u - u_h) + 2a(u - u_h, u_h - v_h) + a(u_h - v_h, u_h - v_h) \\ &= a(u - u_h, u - u_h) + 0 + a(u_h - v_h, u_h - v_h) \\ &\ge a(u - u_h, u - u_h) \end{aligned}$$

Thus by coercivity (again), and continuity, we get the claim:

$$\alpha\|u - u_h\|^2 \le a(u - u_h, u - u_h) \le a(u - v_h, u - v_h) \le M\|u - v_h\|^2 \qquad \square$$

Outline

1. recall: Poisson problems and their finite element (FE) approximations
2. abstract bilinear form notation
3. Céa's lemma: the FE approximation is (essentially) the best
4. polynomial and piecewise-linear interpolation on $\mathbb{R}^1$
5. convergence of the FE approximation as $h \to 0$
6. lecture content in week 14

Lagrange interpolation

- methods for polynomial interpolation in $\mathbb{R}^1$ are old (Vandermonde 1771, Waring 1779, Lagrange 1795)

polynomial interpolation in 1D (Lagrange's formula)

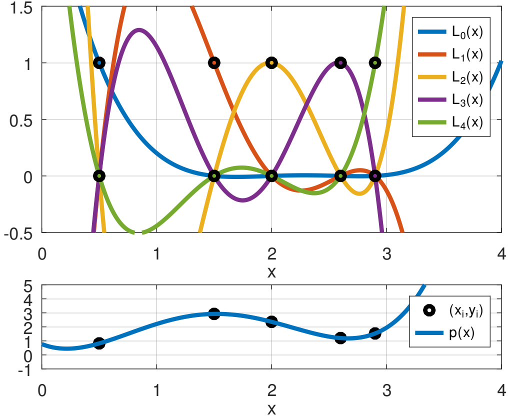
Given $n+1$ points (x_i, y_i) for $i = 0, 1, \dots, n$, with distinct x_i. There exists a unique polynomial $p(x)$ of degree n such that $p(x_i) = y_i$ for all i, and in fact

$$p(x) = \sum_{i=0}^{n} y_i L_i(x)$$

where $L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^{n} \frac{x - x_k}{x_i - x_k}$ is a polynomial of degree n satisfying $L_i(x_j) = \delta_{ij}$.

Lagrange interpolation

- $n = 4$ example:



polynomial interpolation in 1D

$$p(x) = \sum_{i=0}^{n} y_i L_i(x) \quad \text{where} \quad L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^{n} \frac{x - x_k}{x_i - x_k}$$

polynomial interpolation remainder theorem

- a bound for the error in approximating a function by polynomial interpolation is comparably old (Lagrange ∼1800?)
- the polynomial interpolant $p(x)$ is built as before, but now $y_i = f(x_i)$

Theorem (polynomial interpolation remainder)

Suppose $f \in C^{n+1}([a,b])$ and x_i, $i = 0, 1, \dots, n$, are distinct points in $[a,b]$. Let $p(x)$ be the unique polynomial $p(x)$ of degree n such that $p(x_i) = f(x_i)$. Then

$$f(x) = p(x) + R(x)$$

where

$$R(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)(x - x_1)\dots(x - x_n)$$

and $\xi = \xi(x)$ is some point in $[a,b]$.

- obviously: $R(x) = f(x) - p(x)$

the remainder formula: similar to Taylor's

- this *should* sound familiar!

Theorem (polynomial interpolation remainder)

Let $p(x)$ be the unique polynomial of degree n such that $p(x_i) = f(x_i)$. Then $f(x) = p(x) + R(x)$ where

$$R(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)(x - x_1)\dots(x - x_n)$$

for some point $\xi = \xi(x) \in [a, b]$.

Theorem (Taylor's theorem with remainder)

Let $p(x)$ be the Taylor polynomial of degree n at basepoint x_0. Then $f(x) = p(x) + R(x)$ where

$$R(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

for some point $c = c(x)$ between x_0 and x.

proof in the $n = 1$ (linear) case

Theorem (linear interpolation remainder)

Suppose $f \in C^2([a,b])$. For nodes $a = x_0 < x_1 = b$, let $p(x)$ be the unique linear polynomial such that $p(a) = f(a)$ and $p(b) = f(b)$. Then for each $x \in [a,b]$ we have $f(x) = p(x) + R(x)$ where

$$R(x) = \frac{f''(\xi)}{2}(x-a)(x-b),$$

for some point $\xi \in (a,b)$ which depends on x.

Proof. Interpolation is exact at ends ($R(a) = R(b) = 0$), so fix $x \in (a,b)$. For $t \in [a,b]$ define

$$g(t) = f(t) - p(t) - \frac{f(x) - p(x)}{(x-a)(x-b)}(t-a)(t-b).$$

Note that g is C^2, and that it has three zeros: $g(a) = g(x) = g(b) = 0$. By Rolle's theorem there are points $a < c < x < d < b$ so that $g'(c) = 0$ and $g'(d) = 0$. Applying Rolle's theorem again there is $c < \xi < d$ so that $g''(\xi) = 0$. However,

$$0 = g''(\xi) = f''(\xi) - 0 - \frac{f(x) - p(x)}{(x-a)(x-b)}2,$$

as claimed. □

piecewise-linear interpolation in $\mathbb{R}^1$

- we are primarily interested in the piecewise-linear FE space P_1
- suppose $\Omega = (a, b) \subset \mathbb{R}^1$ and $f \in C^2(\overline{\Omega})$
- consider a mesh $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$; picture:

- let $p(x)$ be the continuous, piecewise-linear interpolant of $f(x)$
  - on $[x_{i-1}, x_i]$: $p(x) = f(x_{i-1}) + \dfrac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}(x - x_{i-1})$
- the previous theorem gives a bound on $[x_{i-1}, x_i]$:

$$|f(x) - p(x)| \le \frac{M}{2} |(x - x_{i-1})(x - x_i)|$$

where $M = \max_{\xi \in [x_{i-1}, x_i]} |f''(\xi)|$

piecewise-linear interpolation error bound

- let $h = \max_i |x_i - x_{i-1}|$, and compute:

$$\max_{x \in [x_{i-1}, x_i]} |(x - x_{i-1})(x - x_i)| = \frac{1}{4}|x_{i+1} - x_i|^2 \le \frac{1}{4} h^2$$

- thus on every subinterval (= cell or element) $x \in [x_{i-1}, x_i]$,

$$|f(x) - p(x)| \le \frac{\|f''\|_\infty}{8} h^2$$

Theorem (uniform interpolation bound in $\mathbb{R}^1$)

Let $f \in C^2([a,b])$. For a mesh $\{x_i\}$ on $[a,b]$ with $h = \max |x_i - x_{i-1}|$, consider the piecewise-linear interpolant $p = \mathbb{I}_h f \in P_1$. Then

$$\|f - \mathbb{I}_h f\|_\infty \le \frac{\|f''\|_\infty}{8} h^2.$$

how about the H^1 norm?

- we also want an error bound in the H^1 norm, so we need a lemma

Lemma

Suppose $f \in C^2([a,b])$ *and suppose* $[x_{i-1}, x_i] \subset [a,b]$. *Let* $p(x)$ *be the unique linear polynomial such that* $p(x_j) = f(x_j)$. *Then*

$$\int_{x_{i-1}}^{x_i} |f'(x) - p'(x)|^2\, dx \le h^3 \|f''\|_\infty^2$$

where $h = \max_i |x_i - x_{i-1}|$.

Proof. Note $p'(x)$ is constant in $[x_{i-1}, x_i]$. By the mean value theorem there is $c \in (x_{i-1}, x_i)$ so that $f'(c)$ has this value. By FTC and triangle inequality:

$$\begin{aligned}\int_{x_{i-1}}^{x_i} |f'(x) - p'(x)|^2\, dx &= \int_{x_{i-1}}^{x_i} |f'(x) - f'(c)|^2\, dx = \int_{x_{i-1}}^{x_i} \left| \int_x^c f''(t)\, dt \right|^2 dx \\ &\le \int_{x_{i-1}}^{x_i} \left(\int_{x_{i-1}}^{x_i} |f''(t)|\, dt \right)^2 dx \le (x_i - x_{i-1})^3 \|f''\|_\infty^2. \qquad \square\end{aligned}$$

Theorem (interpolation error theorem for P_1 in $\mathbb{R}^1$)

Let $\Omega = (a, b)$ and $f \in C^2(\overline{\Omega})$. Consider a sequence of meshes $\{x_i\}$ on Ω with $h = \max_i |x_i - x_{i-1}|$. Suppose the meshes are quasi-uniform in the sense that there is $c > 0$ so that $ch \le |x_i - x_{i-1}| \le h$ for all i. The piecewise-linear interpolants $\mathbb{I}_h f \in P_1$ satisfy

$$\|f - \mathbb{I}_h f\|_{L^2} = O(h^2)\|f''\|_\infty$$
$$\|f - \mathbb{I}_h f\|_{H^1} = O(h^1)\|f''\|_\infty$$

as $h \to 0$.

Proof. Let $p = \mathbb{I}_h f$. From the uniform bound,

$$\|f - p\|_{L^2}^2 = \int_a^b |f(x) - p(x)|^2\,dx \le \int_a^b \|f - p\|_\infty^2\,dx$$
$$\le (b - a)\left(\frac{\|f''\|_\infty}{8} h^2\right)^2 = C_1 h^4 \|f''\|_\infty^2$$

Thus $\|f - p\|_{L^2} \le \sqrt{C_1} h^2 \|f''\|_\infty$, which proves the first big-O claim.

decay rate (proof continued)

Proof cont. Now using the lemma,

$$\|f' - p'\|_{L^2}^2 = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f'(x) - p'(x)|^2 \, dx \le nh^3 \|f''\|_\infty^2$$

Note that by quasi-uniformity of the mesh, $nh \le (b-a)/c$. Thus

$$\|f' - p'\|_{L^2}^2 \le \frac{b-a}{c} h^2 \|f''\|_\infty^2$$

Combining these results,

$$\begin{aligned}\|f - p\|_{H^1}^2 &= \|f - p\|_{L^2}^2 + \|f' - p'\|_{L^2}^2 \le C_1 h^4 \|f''\|_\infty^2 + \frac{b-a}{c} h^2 \|f''\|_\infty^2 \\ &\le (C_1 h^4 + \frac{b-a}{c} h^2) \|f''\|_\infty^2 \le C_2 h^2 \|f''\|_\infty^2.\end{aligned}$$

(Restricting to $0 < h \le 1$, take $C_2 = C_1 + (b-a)/c$.) Thus $\|f - p\|_{H^1} \le \sqrt{C_2} h \|f''\|_\infty$ as claimed. □

Outline

convergence for FE solutions of the Poisson problem in 1D

- we have the following theorem for the 1D Poisson equation $-u'' = f$

Theorem

Suppose $f \in C([a,b])$. Let $u \in H_0^1(\Omega)$ be the unique solution of the weak-form Poisson equation on $\Omega = (a,b)$ with homogeneous Dirichlet conditions:

$$\int_a^b u'(x)v'(x)\,dx = \int_a^b f(x)v(x)\,dx \qquad \forall v \in H_0^1(\Omega).$$

Suppose a conforming P_1 FE method is applied over a sequence of quasi-uniform meshes[a] *$\mathcal{T}_h = \{x_i\}$. Let $u_h \in P_1$ be the solution of*

$$\int_a^b u_h'(x)v_h'(x)\,dx = \int_a^b f(x)v_h(x)\,dx \qquad \forall v_h \in P_1.$$

Then $\|u - u_h\|_{H^1} = O(h)$ as $h \to 0$, and thus $u_h \to u$ in H^1.

[a] There is $c > 0$ independent of h so that $ch \le |x_i - x_{i-1}| \le h$ for all $x_i \in \mathcal{T}_h$.

proof of convergence (Poisson problem in 1D)

Proof. We finally put all the pieces together.

First, because $f \in C([a,b])$, in 1D it is easy to see by integration that $u \in C^2([a,b])$.

Second, by quasi-optimality (Cea's lemma),

$$\|u - u_h\|_{H^1} \le c\|u - v_h\|_{H^1}$$

for all $v_h \in P_1$. We substitute $v_h = \mathbb{I}_h u$, the P_1 interpolant of the exact solution, into the quasi-optimality inequality.

Finally, use the interpolation error theorem, because we know $u \in C^2([a,b])$:

$$\|u - u_h\|_{H^1} \le c\|u - \mathbb{I}_h u\|_{H^1} = O(h)\|u''\|_\infty = O(h)\|f\|_\infty = O(h).$$

Thus $u_h \to u$ in H^1. □

notes about the proof

- Regarding "easy to see by integration" that $u \in C^2([a, b])$: This says that *regularity theory* is easy in 1D.
  - The logic of regularity in 1D is that you may integrate twice to construct $\tilde{u}$ satisfying $-\tilde{u}'' = f$. This is the calculation I did long ago in the week 2 slides, to solve the problem by an integral operator.
  - Then one uses uniqueness of the weak-form solution to show $\tilde{u} = u$.
- It follows that $-u'' = f$, so u solves the strong form. (But we don't need that fact when proving the FE method convergence.)
- I have written the constant in Cea's lemma simply as "c". We showed that this constant is $c = 1/\alpha = 2/\min\{1, 1/C^2\}$, where C is the constant in Poincaré-Friedrichs. (But this detail is not important here.)
- There is a different argument that shows $\|u - u_h\|_{L^2} = O(h^2)$, but I definitely can't fit it in here. See section II.7 of Braess.[1]

[1] D. Braess, *Finite Elements: Theory, fast solvers, and applications in solid mechanics*, 3rd ed., Cambridge 2007.

FE theory in $\mathbb{R}^2$ and higher

- I have shown all the details in 1D
- 2D has more complications, and I can't show all details
- specifically for the Poisson equation:
  - well-posedness of the weak form is already done in $\mathbb{R}^d$
  - the regularity theory is more difficult in $\mathbb{R}^2$ and higher
    - in $\mathbb{R}^1$ we could just integrate to get a classical solution
    - in $\mathbb{R}^2$ and higher one must use the ellipticity of the Laplacian
    - in $\mathbb{R}^2$ and higher all the straightforward results require Ω to be convex
  - quasi-optimality (Cea's lemma) is already done in $\mathbb{R}^d$
  - the interpolation theory is more difficult in $\mathbb{R}^2$ and higher
    - in $\mathbb{R}^1$ the P_1 interpolation points are the entire boundary of the cells, but this is no longer true
    - in $\mathbb{R}^2$ and higher the interpolation theory is based on H^k norms, not uniform norms as it was in $\mathbb{R}^1$
- I will state some results, from Braess[2] for $\Omega \subset \mathbb{R}^2$

[2] D. Braess, *Finite Elements: Theory, fast solvers, and applications in solid mechanics*, 3rd ed., Cambridge 2007.

solution regularity for the Poisson problem

Definition (H^2 regularity)

Suppose $\Omega \subset \mathbb{R}^2$ is open and bounded and $H_0^1(\Omega) \subset V \subset H^1(\Omega)$. A weak form problem $a(u, v) = \langle f, v\rangle_{L^2}$ for $u \in V$ is *H^2-regular* if there is a constant $c \geq 0$, depending on Ω only, so that the problem has a solution $u \in H^2(\Omega)$ and

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

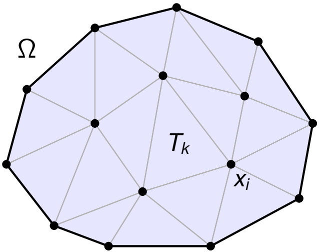
Theorem (H^2 regularity for the Poisson problem; Braess II.7)

Suppose $\Omega \subset \mathbb{R}^2$ is open and bounded. Assume that either Ω is convex or $\partial\Omega$ has a C^2 boundary. Then the homogeneous Dirichlet weak-form problem for the Poisson equation is H^2-regular.

- recall: $$\|f\|_{H^2}^2 = \sum_{|\alpha|\leq 2} \int_\Omega |D^\alpha f(x)|^2\, dx$$
- some requirement that the boundary is "nice" is required
    - if Ω has an interior corner then H^2-regular fails
- the Helmholtz problem is also H^2-regular for nice domains

what precisely is a triangulation, anyway?

- suppose $\Omega \subset \mathbb{R}^2$ is open and bounded
- also suppose that $\partial\Omega$ is a polygon, so $\partial\Omega$ is a union of finitely-many line segments in the plane
- on such a domain, what is a "triangulation"?



Definition (admissible triangulation; Braess II.5)

A partition $\mathcal{T} = \{T_1, \dots, T_m\}$ of Ω into (closed) triangles T_k if

1. $\overline{\Omega} = \bigcup_{i=1}^m T_k$
2. for $j \neq k$, if $T_j \cap T_k$ is nonempty then it is either a single point or an entire common edge of T_j and T_k

shape-regular triangulations

- $h_T = \text{diam}(T)/2$ is the radius of the smallest circumscribed circle
- ρ_T is the radius of the largest inscribed circle
- to prove convergence under refinement of the mesh we will need for all triangles to get small:
$$h = \max\{h_{T_k})\} \to 0$$
- we must consider a sequence of triangulations, denoted $\mathcal{T}_h$
- however, we need nice sequences of triangulations so that the interpolation errors will get consistently small

Definition (shape-regular triangulation; Braess II.5)

A sequence (family) of admissible triangulation $\mathcal{T}_h$ of Ω is *shape-regular* if there exists $\kappa > 0$ independent of h so that

$$\frac{h_T}{\rho_T} \le \kappa \qquad \forall T \in \mathcal{T}_h$$

interpolation errors in $\mathbb{R}^2$

Theorem (interpolation error theorem for P_1 in $\mathbb{R}^2$)

Suppose $\Omega \subset \mathbb{R}^2$ is an open, connected, and bounded set with a polygonal boundary $\partial\Omega$. Suppose $\mathcal{T}_h$ is a shape-regular triangulation of Ω. Let $\mathbb{I}_h f \in P_1$ be the piecewise-linear interpolant of f. If $f \in H^2(\Omega)$ then

$$\|f - \mathbb{I}_h f\|_{L^2} = O(h^2)|f|_{H^2}$$
$$\|f - \mathbb{I}_h f\|_{H^1} = O(h^1)|f|_{H^2}$$

as $h \to 0$.

- if $f \in H^2(\Omega)$ then $f \in C(\overline{\Omega})$, and so the interpolant is well-defined because the function has point values
- note: $$|f|_{H^2}^2 = \sum_{|\alpha|=2} \int_\Omega |D^\alpha f(x)|^2\,dx = \int_\Omega |f_{xx}|^2 + 2|f_{xy}|^2 + |f_{yy}|^2\,dx$$
- this theorem is a corollary of the Bramble-Hilbert lemma ($\sim$1970)

convergence for FE solutions of the Poisson problem in 2D

- putting together regularity theory, quasi-optimality, and interpolation error theory we get convergence

Theorem

Suppose $\Omega \subset \mathbb{R}^2$ is open, $\partial\Omega$ is a convex polygon. Suppose that the triangulations $\mathcal{T}_h$ are shape-regular. Let $f \in L^2(\Omega)$. Let $u \in H_0^1(\Omega)$ be the unique solution of the weak-form:

$$\int_\Omega \nabla u(x) \cdot \nabla v(x)\, dx = \int_\Omega f(x) v(x)\, dx \qquad \forall v \in H_0^1(\Omega).$$

Suppose the conforming P_1 FE method computes $u_h \in P_1$:

$$\int_\Omega \nabla u_h(x) \cdot \nabla v_h(x)\, dx = \int_\Omega f(x) v_h(x)\, dx \qquad \forall v_h \in P_1.$$

Then $\|u - u_h\|_{H^1} = O(h)$ as $h \to 0$, and thus $u_h \to u$ in H^1.

Outline

1. recall: Poisson problems and their finite element (FE) approximations
2. abstract bilinear form notation
3. Céa's lemma: the FE approximation is (essentially) the best
4. polynomial and piecewise-linear interpolation on $\mathbb{R}^1$
5. convergence of the FE approximation as $h \to 0$
6. lecture content in week 14

lecture content in week 14

- this material is only from slides
  - requires understanding of the week 12 and week 13 slides content!
- Assignment 8 is posted at `bueler.github.io/fa`

to know from these slides:

- abstract bilinear notation
  - weak form is $a(u, v) = \ell(v)$, FE method is $a(u_h, v_h) = \ell(v_h)$
- conforming FE method: $V_h \subset V$ and a, ℓ match
- Galerkin orthogonality: $a(u - u_h, v_h) = 0 \quad \forall v_h \in V_h$
- quasi-optimality (Cea's lemma): $\|u - u_h\| \le \frac{M}{\alpha}\|u - v_h\| \quad \forall v_h \in V_h$
  - u_h is the best (up to a constant)
- interpolation operator: $w \in V \to \mathbb{I}_h w \in V_h$
- polynomial interpolation remainder theorem (1D)
- interpolation error theorem for P_1 (piecewise-linear) in 1D
- proof that $u_h \to u$ as $h \to 0$ for P_1 FE method on Poisson equation (1D)
- a rough idea of why $u_h \to u$ as $h \to 0$ for P_1 FE method in 2D