

Introduction to the Finite Element Method (FEM)

A Model Problem and the Variational Framework

Kyoungman Kim

Sungkyunkwan University
Department of Mathematics

October 6, 2026

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

1D Poisson Equation: Variational Formulation

For $f \in L^2(0, 1)$, consider

$$-u''(x) = f(x), \quad 0 < x < 1, \quad u(0) = u(1) = 0. \quad (1)$$

Multiply by a test function v with $v(0) = v(1) = 0$ and integrate by parts:

$$\int_0^1 f v \, dx = \underbrace{[-u'v]_0^1}_{=0} + \int_0^1 u' v' \, dx. \quad (2)$$

Variational formulation

Find $u \in H_0^1(0, 1)$ such that

$$\int_0^1 u' v' \, dx = \int_0^1 f v \, dx \quad \forall v \in H_0^1(0, 1). \quad (3)$$

The space $H_0^1(0, 1)$ imposes zero endpoint values. The weak form requires only first derivatives, which suits piecewise linear FEM.

Overview of FEM

The finite element method (FEM) approximates the solution using **piecewise polynomial basis functions** on a mesh.

- Decompose the domain Ω into discrete elements K that form a mesh $\mathcal{T}_h$, such as intervals in 1D or triangles in 2D.
- Choose basis functions ϕ_i whose restriction to each element K is a polynomial.

Finite element representation of the solution

$$u(x) \approx u_h(x) = \sum_{i=1}^M u_i \phi_i(x). \quad (1)$$

The coefficients $u_1, \dots, u_M$ are the unknowns to be computed.

A One-Dimensional Boundary Value Problem

We begin with a simple boundary value problem to illustrate FEM.

Model problem

Given $f \in L^2(0, 1)$, consider

$$\begin{cases} -u''(x) = f(x), & 0 < x < 1, \\ u(0) = u(1) = 0. \end{cases} \quad (1)$$

- How can we solve this problem numerically?
- We first derive a variational formulation and then restrict it to a finite-dimensional space.
- If f is available only as discrete data, we also need a reconstruction of f or quadrature for the load integrals.

Test Functions and Integration by Parts

An elementary test space

$$V_{\text{pc}} = \left\{ v \in C([0, 1]) : \begin{array}{l} v \text{ is } C^1 \text{ on each subinterval of a finite partition,} \\ v' \text{ is bounded on each subinterval, } v(0) = v(1) = 0 \end{array} \right\}. \quad (1)$$

For a sufficiently regular solution u , multiply the equation by $v \in V_{\text{pc}}$:

$$\begin{aligned} \int_0^1 f v \, dx &= \int_0^1 -u'' v \, dx \\ &= \underbrace{[-u'v]_0^1}_{=0} + \int_0^1 u' v' \, dx. \end{aligned} \quad (2)$$

Variational identity

$$\langle u', v' \rangle = \langle f, v \rangle, \quad \langle p, q \rangle := \int_0^1 p(x) q(x) \, dx. \quad (3)$$

The Weak Formulation

The energy space

The natural Hilbert space is

$$V = H_0^1(0, 1), \quad V_{\text{pc}} \subset V. \quad (1)$$

We will define weak derivatives and Sobolev spaces later.

Weak problem

Find $u \in V$ such that

$$\int_0^1 u'v' \, dx = \int_0^1 fv \, dx \quad \text{for every } v \in V. \quad (2)$$

- This identity involves only first derivatives.
- The boundary conditions are encoded in V .
- In one dimension, H^1 functions have continuous representatives, so their endpoint values are well defined.

A Finite-Dimensional Subspace

Choose a partition

$$0 = x_0 < x_1 < \cdots < x_N = 1, \quad I_j = (x_{j-1}, x_j), \quad h_j = x_j - x_{j-1}. \quad (1)$$

Here $j = 1, \dots, N$, and $h := \max_j h_j$ is the mesh size.

Continuous piecewise linear functions

$$V_h = \left\{ v_h \in C([0, 1]) : \begin{array}{l} v_h|_{I_j} \in \mathbb{P}_1(I_j), \quad j = 1, \dots, N, \\ v_h(0) = v_h(1) = 0 \end{array} \right\}. \quad (2)$$

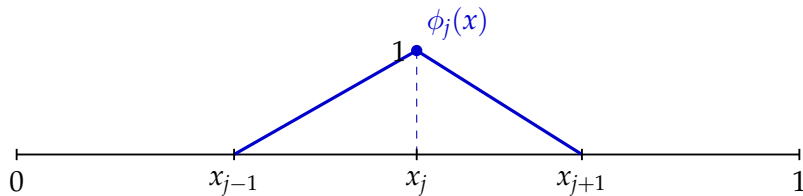
Thus $V_h \subset H_0^1(0, 1)$ and $\dim V_h = N - 1$.

Nodal basis

For $i = 1, \dots, N - 1$, let $\phi_i \in V_h$ satisfy

$$\phi_i(x_j) = \delta_{ij}, \quad v_h(x) = \sum_{i=1}^{N-1} v_h(x_i) \phi_i(x). \quad (3)$$

Piecewise Linear Basis Functions



Hat function, $j = 1, \dots, N - 1$

$$\phi_j(x) = \begin{cases} \frac{x - x_{j-1}}{h_j}, & x_{j-1} \leq x \leq x_j, \\ \frac{x_{j+1} - x}{h_{j+1}}, & x_j < x \leq x_{j+1}, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Its support is $[x_{j-1}, x_{j+1}]$.

The Discrete Variational Formulation

Galerkin problem

Find the numerical solution $u_h \in V_h$ such that

$$\langle u_h', v_h' \rangle = \langle f, v_h \rangle \quad \forall v_h \in V_h. \quad (1)$$

Write

$$u_h(x) = \sum_{i=1}^{N-1} u_i \phi_i(x), \quad u_i = u_h(x_i), \quad (2)$$

and choose $v_h = \phi_j, j = 1, \dots, N-1$. Then

$$\begin{aligned} \left\langle \sum_{i=1}^{N-1} u_i \phi_i', \phi_j' \right\rangle &= \langle f, \phi_j \rangle, \\ \sum_{i=1}^{N-1} u_i \langle \phi_i', \phi_j' \rangle &= \langle f, \phi_j \rangle. \end{aligned} \quad (3)$$

The Stiffness Matrix

$$A_{ji} := \langle \phi'_i, \phi'_j \rangle, \quad b_j := \langle f, \phi_j \rangle, \quad i, j = 1, \dots, N-1. \quad (1)$$

A linear system for the nodal coefficients

$$\sum_{i=1}^{N-1} A_{ji} u_i = b_j, \quad A\mathbf{u} = \mathbf{b}. \quad (2)$$

The matrix A is called the **stiffness matrix**.

Symmetry and Positive Definiteness

Symmetry follows from $A_{ji} = \langle \phi'_i, \phi'_j \rangle = \langle \phi'_j, \phi'_i \rangle = A_{ij}$.

The energy identity from the matrix product

For $\mathbf{z} \in \mathbb{R}^{N-1}$, set $w_h = \sum_{i=1}^{N-1} z_i \phi_i$. Then

$$\begin{aligned} \mathbf{z}^T A \mathbf{z} &= \sum_{j=1}^{N-1} z_j (A \mathbf{z})_j = \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} z_j A_{ji} z_i \\ &= \sum_{i,j=1}^{N-1} z_i z_j \int_0^1 \phi'_i(x) \phi'_j(x) \, dx \\ &= \int_0^1 \left(\sum_{i=1}^{N-1} z_i \phi'_i(x) \right) \left(\sum_{j=1}^{N-1} z_j \phi'_j(x) \right) \, dx \\ &= \int_0^1 |w'_h(x)|^2 \, dx \geq 0. \end{aligned} \tag{1}$$

We interchange the finite sums and the integral. Both factors equal $w'_h(x)$.

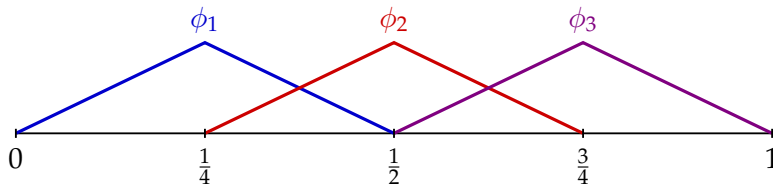
If the energy is zero, $w'_h = 0$ almost everywhere, so w_h is constant.

Since $w_h(0) = w_h(1) = 0$, we have $w_h = 0$. Linear independence of the basis gives $\mathbf{z} = 0$.

Thus A is **symmetric positive definite** and invertible.

Example: A Uniform Mesh with $h = 1/4$

The nodes are $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$, with one basis function per interior node.



The first basis function

$$\phi_1(x) = \begin{cases} 4x, & 0 \leq x \leq \frac{1}{4}, \\ 2 - 4x, & \frac{1}{4} < x \leq \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The derivative almost everywhere

$$\phi_1'(x) = \begin{cases} 4, & 0 < x < \frac{1}{4}, \\ -4, & \frac{1}{4} < x < \frac{1}{2}, \\ 0, & \frac{1}{2} < x < 1. \end{cases} \quad (2)$$

Example: Computing the Matrix Entries

$$\langle \phi'_1, \phi'_1 \rangle = \int_0^{1/4} 4^2 \, dx + \int_{1/4}^{1/2} (-4)^2 \, dx = 4 + 4 = 8,$$

$$\langle \phi'_1, \phi'_2 \rangle = \int_{1/4}^{1/2} (-4)(4) \, dx = -4, \tag{1}$$

$$\langle \phi'_1, \phi'_3 \rangle = 0.$$

Local support and matrix sparsity

Only neighboring hat functions have derivatives whose supports overlap on an interval of positive length.

The same calculation gives 8 on the diagonal and -4 on the two adjacent diagonals.

Example: The Discrete System

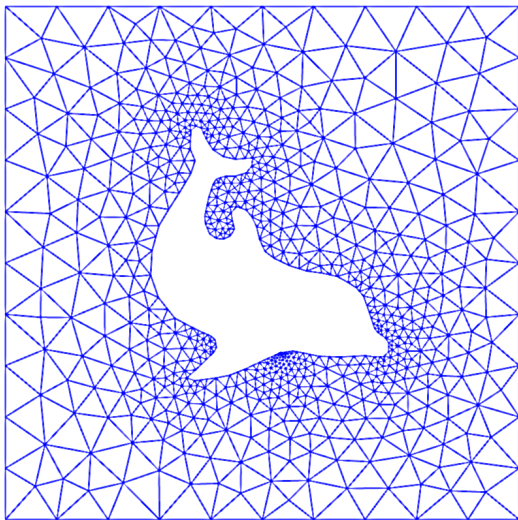
The resulting linear system

$$\begin{pmatrix} 8 & -4 & 0 \\ -4 & 8 & -4 \\ 0 & -4 & 8 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \quad b_j = \int_0^1 f(x)\phi_j(x) \, dx. \quad (1)$$

Once we solve for the coefficients, we obtain the finite element solution

$$u_h(x) = u_1\phi_1(x) + u_2\phi_2(x) + u_3\phi_3(x). \quad (2)$$

Beyond the One-Dimensional Example



- 1 How do we differentiate basis functions at mesh nodes?
We use weak derivatives and Sobolev spaces.
- 2 Does discretization always produce a symmetric positive definite matrix?
No. The operator and boundary conditions matter.
- 3 Is the bilinear form in a variational problem always an inner product?
No. We use bilinear forms and the Lax–Milgram theorem.

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

Definition of a Weak Derivative

Let $U \subset \mathbb{R}^n$ be open and $u, v \in L^1_{\text{loc}}(U)$. A multi-index is

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n, \quad |\alpha| = \alpha_1 + \dots + \alpha_n, \quad (1)$$

and $D^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}$ on smooth functions.

Weak derivative

We say that $D^\alpha u = v$ in the weak sense if

$$\int_U u D^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_U v \varphi \, dx \quad \text{for every } \varphi \in C_c^\infty(U). \quad (2)$$

- The test functions are smooth and compactly supported in U .
- When a weak derivative exists, it is unique almost everywhere.

The Weak Derivative of a Hat Function

For the hat function ϕ_j , define g_j almost everywhere by

$$g_j(x) = \begin{cases} 1/h_j, & x \in (x_{j-1}, x_j), \\ -1/h_{j+1}, & x \in (x_j, x_{j+1}), \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

For $\varphi \in C_c^\infty(0, 1)$, integration by parts on the two subintervals gives

$$\begin{aligned} \int_0^1 \phi_j \varphi' \, dx &= [\phi_j \varphi]_{x_{j-1}}^{x_j} + [\phi_j \varphi]_{x_j}^{x_{j+1}} - \int_0^1 g_j \varphi \, dx \\ &= - \int_0^1 g_j \varphi \, dx. \end{aligned} \quad (2)$$

Continuity ensures cancellation at the interface

g_j is therefore the weak derivative of ϕ_j . Its values at the finitely many nodes do not affect its L^2 equivalence class.

Definition of a Sobolev Space

Let $1 \leq p \leq \infty$ and $k \in \mathbb{N}_0$.

Sobolev space

$$W^{k,p}(U) = \{u \in L^p(U) : D^\alpha u \in L^p(U) \text{ for every } |\alpha| \leq k\}. \quad (1)$$

Here $D^\alpha u$ denotes a weak derivative.

We use the norm

$$\|u\|_{W^{k,p}(U)} = \begin{cases} \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(U)}^p \right)^{1/p}, & 1 \leq p < \infty, \\ \sum_{|\alpha| \leq k} \operatorname{ess\,sup}_{x \in U} |D^\alpha u(x)|, & p = \infty. \end{cases} \quad (2)$$

The term with $\alpha = 0$ is u itself.

Basic Properties of Sobolev Spaces

- 1 The Sobolev norm is well defined on equivalence classes modulo equality almost everywhere.
- 2 $W^{k,p}(U)$ is a **Banach space** for $1 \leq p \leq \infty$.
- 3 For $p = 2$, write

$$H^k(U) := W^{k,2}(U). \quad (1)$$

This is a **Hilbert space** with inner product

$$(u, v)_{H^k(U)} = \sum_{|\alpha| \leq k} \int_U D^\alpha u D^\alpha v \, dx. \quad (2)$$

The case $k = 1$

$$\|u\|_{H^1(U)}^2 = \|u\|_{L^2(U)}^2 + \|\nabla u\|_{L^2(U)}^2. \quad (3)$$

All function spaces in this presentation are over $\mathbb{R}$.

Seminorms and Sobolev Norms

A seminorm s satisfies

$$s(u) \geq 0, \quad s(cu) = |c|s(u), \quad s(u + v) \leq s(u) + s(v). \quad (1)$$

It may satisfy $s(u) = 0$ even when $u \neq 0$.

Sobolev seminorm

For $k \geq 1$ and $1 \leq p < \infty$, define

$$|u|_{W^{k,p}(U)} := \left(\sum_{|\alpha|=k} \|D^\alpha u\|_{L^p(U)}^p \right)^{1/p}. \quad (2)$$

The norm includes orders $0, \dots, k$. The seminorm uses order k only.

$$|u|_{H^1(U)} = \|\nabla u\|_{L^2(U)}, \quad \|u\|_{H^1(U)}^2 = \|u\|_{L^2(U)}^2 + |u|_{H^1(U)}^2. \quad (3)$$

A nonzero constant on bounded U has zero seminorm but positive norm.

Sobolev Spaces with Zero Boundary Values

Definition by closure

$$W_0^{k,p}(U) := \overline{C_c^\infty(U)}^{W^{k,p}(U)}. \quad (1)$$

Equivalently, $u \in W_0^{k,p}(U)$ if and only if there is a sequence $(u_m) \subset C_c^\infty(U)$ such that

$$\|u_m - u\|_{W^{k,p}(U)} \longrightarrow 0 \quad \text{as } m \rightarrow \infty. \quad (2)$$

The space used for homogeneous Dirichlet problems

$$H_0^1(U) := W_0^{1,2}(U). \quad (3)$$

This is a closed subspace of $H^1(U)$, hence a Hilbert space. On a bounded Lipschitz domain, it consists exactly of the H^1 functions whose trace on ∂U is zero.

The definition by closure also applies when $p = \infty$. The trace characterization above applies specifically to $H_0^1(U)$.

Definition of a C^k Boundary

Let $U \subset \mathbb{R}^n$ be open and bounded, and let $k \geq 1$.

Local graph representation

We say that ∂U is C^k if, near every $x^0 \in \partial U$, after a translation and an orthogonal change of coordinates, there exist $r > 0$ and a C^k function γ such that

$$U \cap B(x^0, r) = \{x \in B(x^0, r) : x_n > \gamma(x')\}, \quad x' = (x_1, \dots, x_{n-1}). \quad (1)$$

Consequently,

$$\partial U \cap B(x^0, r) = \{x \in B(x^0, r) : x_n = \gamma(x')\}. \quad (2)$$

The graph function need only be defined on a neighborhood of the projection of this boundary patch onto $\mathbb{R}^{n-1}$.

Poincaré's Inequality

Let $U \subset \mathbb{R}^n$ be bounded, connected, and open with C^1 boundary.

Mean-zero form, $1 \leq p \leq \infty$

There is $C = C(U, n, p) > 0$ such that

$$\|u - (u)_U\|_{L^p(U)} \leq C \|\nabla u\|_{L^p(U)}, \quad u \in W^{1,p}(U), \quad (1)$$

where

$$(u)_U := \frac{1}{|U|} \int_U u \, dx. \quad (2)$$

Zero-trace form

For some $C_P > 0$,

$$\|v\|_{L^2(U)} \leq C_P \|\nabla v\|_{L^2(U)} \quad \text{for every } v \in H_0^1(U). \quad (3)$$

The second inequality applies to functions with zero boundary trace.

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem**
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

Definition of a Bilinear Map

Let V, W, X be vector spaces over the same field $\mathbb{F}$.

Bilinearity

A map $B : V \times W \rightarrow X$ is bilinear if it is linear in each argument separately. In particular,

$$\begin{aligned} B(\lambda v, w) &= \lambda B(v, w) = B(v, \lambda w), \\ B(v_1 + v_2, w) &= B(v_1, w) + B(v_2, w), \\ B(v, w_1 + w_2) &= B(v, w_1) + B(v, w_2). \end{aligned} \tag{1}$$

Here $\lambda \in \mathbb{F}$, $v, v_1, v_2 \in V$, and $w, w_1, w_2 \in W$.

- A **bilinear form** is a scalar-valued bilinear map.
- A real inner product is a symmetric, positive definite bilinear form.
- A general bilinear form need not be symmetric or positive definite.

The Lax–Milgram Theorem

Let H be a real Hilbert space and $B : H \times H \rightarrow \mathbb{R}$ a bilinear form.

Continuity and coercivity

Suppose there exist $\alpha, \beta > 0$ such that

$$\begin{aligned} |B(u, v)| &\leq \alpha \|u\|_H \|v\|_H && \text{for all } u, v \in H, \\ B(u, u) &\geq \beta \|u\|_H^2 && \text{for all } u \in H. \end{aligned} \tag{1}$$

Existence and uniqueness

For each bounded linear functional $F \in H^*$, there exists a unique $u \in H$ such that

$$B(u, v) = F(v) \quad \text{for every } v \in H. \tag{2}$$

Moreover, $\|u\|_H \leq \beta^{-1} \|F\|_{H^*}$.

Symmetry of B is not required.

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

A Uniformly Elliptic Boundary Value Problem

Let $U \subset \mathbb{R}^n$ be a bounded domain with C^1 boundary. Consider

$$\begin{cases} Lu = f & \text{in } U, \\ u = 0 & \text{on } \partial U, \end{cases} \quad (1)$$

where $A(x) = (a^{ij}(x))_{i,j=1}^n$ and

$$\begin{aligned} Lu &= - \sum_{i,j=1}^n \partial_{x_i} (a^{ij}(x) \partial_{x_j} u) + \sum_{i=1}^n b^i(x) \partial_{x_i} u + c(x)u \\ &= -\nabla \cdot (A \nabla u) + b \cdot \nabla u + cu. \end{aligned} \quad (2)$$

We impose the boundary condition by requiring $u \in H_0^1(U)$.

Uniform ellipticity

There is $\theta > 0$ such that, for almost every $x \in U$ and every $\xi \in \mathbb{R}^n$,

$$\xi^T A(x) \xi = \sum_{i,j=1}^n a^{ij}(x) \xi_i \xi_j \geq \theta |\xi|^2. \quad (1)$$

- The constant $\theta > 0$ is independent of x and the direction ξ .
- A need not be symmetric. Ellipticity controls its quadratic form.
- In the energy estimate, we apply this inequality with $\xi = \nabla v(x)$.

Assumptions on the Coefficients

Regularity and data

Assume real-valued coefficients with the following regularity and data:

$$A \in L^\infty(U; \mathbb{R}^{n \times n}), \quad b \in W^{1,\infty}(U; \mathbb{R}^n), \quad (1)$$

$$c \in L^\infty(U), \quad f \in L^2(U). \quad (2)$$

Sign condition for coercivity

$$c - \frac{1}{2} \nabla \cdot b \geq 0 \quad \text{almost everywhere in } U. \quad (3)$$

- Without this assumption, we can study solvability using the Fredholm alternative.
- We use $V = H_0^1(U)$ with the full H^1 norm.

Deriving the Weak Formulation

For smooth A, u, v with $v|_{\partial U} = 0$, integration by parts gives

$$\int_U -\nabla \cdot (A \nabla u) v \, dx = \int_U (A \nabla u) \cdot \nabla v \, dx. \quad (1)$$

Bilinear form and load functional

$$\begin{aligned} a(u, v) &:= \int_U [(A \nabla u) \cdot \nabla v + (b \cdot \nabla u)v + cuv] \, dx, \\ F(v) &:= \int_U f v \, dx. \end{aligned} \quad (2)$$

Weak problem using only first derivatives

Find $u \in V = H_0^1(U)$ such that

$$a(u, v) = F(v) \quad \text{for every } v \in V. \quad (3)$$

Continuity of the Bilinear Form

We define $\|A\|_{L^\infty}$ using the matrix operator norm. By Hölder's inequality,

$$\begin{aligned}\left| \int_U (A \nabla u) \cdot \nabla v \, dx \right| &\leq \|A\|_{L^\infty} \|\nabla u\|_{L^2} \|\nabla v\|_{L^2} , \\ \left| \int_U (b \cdot \nabla u) v \, dx \right| &\leq \|b\|_{L^\infty} \|\nabla u\|_{L^2} \|v\|_{L^2} , \\ \left| \int_U c u v \, dx \right| &\leq \|c\|_{L^\infty} \|u\|_{L^2} \|v\|_{L^2} .\end{aligned}\tag{1}$$

A bound in the H^1 norm

$$|a(u, v)| \leq M \|u\|_{H^1(U)} \|v\|_{H^1(U)} ,\tag{2}$$

where $M = \|A\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty}$.

All coefficient and function norms above are taken over U .

Boundedness of the Load Functional

The map $F : V \rightarrow \mathbb{R}$, $F(v) = \int_U f v \, dx$, is linear.

Boundedness

For $f \in L^2(U)$ and $v \in V$,

$$\begin{aligned} |F(v)| &\leq \|f\|_{L^2(U)} \|v\|_{L^2(U)} \\ &\leq \|f\|_{L^2(U)} \|v\|_{H^1(U)}. \end{aligned} \tag{1}$$

Hence $F \in V^*$ and

$$\|F\|_{V^*} \leq \|f\|_{L^2(U)}. \tag{2}$$

To apply the Lax–Milgram theorem, we still need to prove that a is coercive.

Coercivity: The Convection Term

For a smooth v with zero boundary values,

$$\begin{aligned}\int_U (b \cdot \nabla v) v \, dx &= \frac{1}{2} \int_U b \cdot \nabla (v^2) \, dx \\ &= \frac{1}{2} \underbrace{\int_{\partial U} (b \cdot \nu) v^2 \, dS}_{=0} - \frac{1}{2} \int_U (\nabla \cdot b) v^2 \, dx \\ &= -\frac{1}{2} \int_U (\nabla \cdot b) v^2 \, dx,\end{aligned}\tag{1}$$

where ν is the outward unit normal.

Extension to $H_0^1(U)$

Choose $v_m \in C_c^\infty(U)$ with $v_m \rightarrow v$ in $H^1(U)$. The assumptions $b, \nabla \cdot b \in L^\infty(U)$ allow passage to the limit, so the same identity holds for every $v \in H_0^1(U)$.

Coercivity: Ellipticity and Poincaré's Inequality

Using the convection identity,

$$\begin{aligned} a(v, v) &= \int_U (A \nabla v) \cdot \nabla v \, dx + \int_U \left(c - \frac{1}{2} \nabla \cdot b \right) v^2 \, dx \\ &\geq \theta \|\nabla v\|_{L^2(U)}^2. \end{aligned} \tag{1}$$

Poincaré's inequality gives

$$\begin{aligned} \|v\|_{H^1(U)}^2 &= \|v\|_{L^2(U)}^2 + \|\nabla v\|_{L^2(U)}^2 \\ &\leq (1 + C_P^2) \|\nabla v\|_{L^2(U)}^2. \end{aligned} \tag{2}$$

Coercivity in the full H^1 norm

$$a(v, v) \geq \frac{\theta}{1 + C_P^2} \|v\|_{H^1(U)}^2, \quad v \in H_0^1(U). \tag{3}$$

Existence and Uniqueness of the Weak Solution

We have verified the hypotheses of the Lax–Milgram theorem on $V = H_0^1(U)$:

$$\begin{aligned} |a(u, v)| &\leq M \|u\|_{H^1(U)} \|v\|_{H^1(U)} , \\ a(v, v) &\geq \beta \|v\|_{H^1(U)}^2 , \quad \beta = \frac{\theta}{1 + C_P^2} > 0, \\ |F(v)| &\leq \|f\|_{L^2(U)} \|v\|_{H^1(U)} . \end{aligned} \tag{1}$$

Conclusion

There is a unique $u \in H_0^1(U)$ satisfying

$$a(u, v) = F(v) \quad \text{for all } v \in H_0^1(U). \tag{2}$$

The solution satisfies the estimate

$$\|u\|_{H^1(U)} \leq \frac{1 + C_P^2}{\theta} \|f\|_{L^2(U)} . \tag{3}$$

Returning to the Finite Element Method

Let $V_h \subset H_0^1(U)$ be a finite-dimensional conforming space. Restricting the trial and test functions to V_h gives

$$a(u_h, v_h) = F(v_h) \quad \text{for every } v_h \in V_h. \quad (1)$$

The discrete problem also has a unique solution

The continuity and coercivity bounds remain valid on V_h . For a basis $\{\phi_1, \dots, \phi_m\}$,

$$K_{ji} = a(\phi_i, \phi_j), \quad b_j = F(\phi_j), \quad K\mathbf{u} = \mathbf{b}. \quad (2)$$

- The model Poisson problem yields a symmetric positive definite matrix.
- For the general operator, K may be nonsymmetric.
- Under the stated coercivity assumptions, K is nevertheless invertible.

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

A Conforming Simplicial Mesh

In the affine simplicial setting, let $\Omega \subset \mathbb{R}^n$ be a bounded polyhedral Lipschitz domain and set $V = H_0^1(\Omega)$.

An exact conforming partition

Let $\mathcal{T}_h$ be a finite collection of nondegenerate closed simplices such that

$$\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K. \quad (1)$$

- Distinct elements have disjoint interiors.
- Two distinct closed elements intersect, if at all, in a common face of some dimension.

The preceding arguments using Poincaré's inequality also apply on bounded Lipschitz domains. We impose the same coefficient assumptions on Ω .

Mesh Size and Shape Regularity

For each element $K \in \mathcal{T}_h$, define

$$h_K := \text{diam } K, \quad h := \max_{K \in \mathcal{T}_h} h_K. \quad (1)$$

Let $\rho_K > 0$ be the inradius of K .

Shape regularity

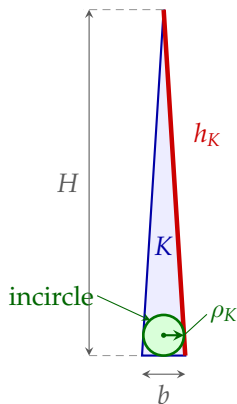
A family of meshes is shape regular if there is a constant $\sigma > 0$, independent of the mesh and its elements, such that

$$\frac{h_K}{\rho_K} \leq \sigma \quad \text{for every } K \in \mathcal{T}_h. \quad (2)$$

- Shape regularity controls the shape of each simplex.
- It permits local refinement.
- It does not require all elements to have comparable sizes.

Why Shape Regularity Matters

A tall, narrow triangle has a large diameter relative to its inradius.



The red edge is one of the longest edges, with length $h_K = \text{diam } K$.

$$h_K = \sqrt{H^2 + (b/2)^2}, \quad (1)$$

$$\rho_K = \frac{bH}{b + 2h_K}.$$

For $H \gg b$, we have $h_K \approx H$ and $\rho_K \approx b/2$.

$$\frac{h_K}{\rho_K} \approx \frac{2H}{b} \gg 1. \quad (2)$$

Here $H = 8b$, so $h_K/\rho_K \approx 17.1$.

Thin elements can produce large basis gradients and large geometric factors in error estimates.

Shape regularity keeps h_K/ρ_K uniformly bounded throughout mesh refinement.

The Conforming Finite Element Space

Fix an integer $p \geq 1$ for the polynomial degree.

Continuous piecewise polynomial functions

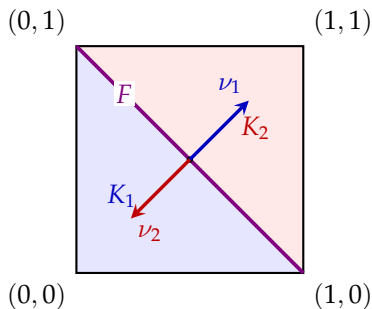
$$V_h = \left\{ v_h \in C^0(\overline{\Omega}) : \begin{array}{ll} v_h|_K \in \mathbb{P}_p(K) & (K \in \mathcal{T}_h), \\ v_h|_{\partial\Omega} = 0 \end{array} \right\}. \quad (1)$$

The space $\mathbb{P}_p(K)$ consists of polynomials of total degree at most p .

- Continuity across element faces is essential.
- Elementwise gradients alone do not imply global H^1 membership.
- Continuity makes the interface terms cancel in integration by parts, yielding $V_h \subset V$.

Conformity: Opposite Outward Normals

Split the unit square into two triangles with common edge F .



Set $s(x, y) = x + y - 1$. Then

$$K_1 : s \leq 0, \quad K_2 : s \geq 0, \quad F : s = 0. \quad (1)$$

Since $\nabla s = (1, 1)$, the outward normals on F are

$$\nu_1 = \frac{\nabla s}{|\nabla s|} = \frac{(1, 1)}{\sqrt{2}}, \quad \nu_2 = -\frac{(1, 1)}{\sqrt{2}}. \quad (2)$$

Each normal points out of its triangle. Both vectors have length 1.

The two outward unit normals sum to zero

$$\nu_1 + \nu_2 = \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \right) = (0, 0). \quad (3)$$

The common edge belongs to both ∂K_1 and ∂K_2 , with opposite outward normals.

Conformity: Integration by Parts

Let $v_h \in V_h$, $\varphi \in C_c^\infty(\Omega)$, and $\ell \in \{1, \dots, n\}$. Define the piecewise derivative

$$g_\ell|_K := \partial_{x_\ell}(v_h|_K), \quad \|g_\ell\|_{L^2(\Omega)}^2 = \sum_{K \in \mathcal{T}_h} \|\partial_{x_\ell}(v_h|_K)\|_{L^2(K)}^2 < \infty. \quad (1)$$

Each restriction is a polynomial. On each simplex, classical integration by parts gives

$$\int_K v_h \partial_{x_\ell} \varphi \, dx = - \int_K g_\ell \varphi \, dx + \int_{\partial K} v_h \varphi (\nu_K)_\ell \, dS. \quad (2)$$

Summing over the elements yields

The remaining boundary terms

$$\int_\Omega v_h \partial_{x_\ell} \varphi \, dx = - \int_\Omega g_\ell \varphi \, dx + S_\ell, \quad S_\ell := \sum_K \int_{\partial K} v_h \varphi (\nu_K)_\ell \, dS. \quad (3)$$

To prove that g_ℓ is the global weak derivative, we must show $S_\ell = 0$ for every such test function.

Conformity: Cancellation Across Interior Faces

Fix an interior face $F = K_1 \cap K_2$. Let $v_1 = (v_h|_{K_1})|_F$ and $v_2 = (v_h|_{K_2})|_F$ be its two traces. The two contributions to S_ℓ are

$$\begin{aligned} S_{F,\ell} &= \int_F v_1 \varphi(\nu_1)_\ell \, dS + \int_F v_2 \varphi(\nu_2)_\ell \, dS \\ &= \int_F (v_1 - v_2) \varphi(\nu_1)_\ell \, dS, \quad \nu_2 = -\nu_1. \end{aligned} \tag{1}$$

Continuity ensures matching traces

Since v_h is continuous, $v_1 = v_2 =: v_F$. Thus

$$v_1 \nu_1 + v_2 \nu_2 = v_F (\nu_1 + \nu_2) = 0, \quad S_{F,\ell} = 0. \tag{2}$$

- Without matching traces, opposite normals leave a factor of $v_1 - v_2$.
- On exterior faces, $\varphi = 0$ because its support lies inside Ω .
- Intersections of lower dimension do not contribute to the surface integrals.

Conformity: The Global Result

The interior face contributions cancel, and the boundary contributions vanish. Therefore, for every $\varphi \in C_c^\infty(\Omega)$,

$$\int_{\Omega} v_h \partial_{x_\ell} \varphi \, dx = - \int_{\Omega} g_\ell \varphi \, dx \quad (\ell = 1, \dots, n). \quad (1)$$

Global H^1 membership

This is exactly the definition of $\partial_{x_\ell} v_h = g_\ell$ as a weak derivative. Since the mesh is finite and every piece is polynomial, $v_h, g_1, \dots, g_n \in L^2(\Omega)$. Hence

$$v_h \in H^1(\Omega), \quad \nabla v_h|_K = \nabla(v_h|_K) \quad \text{almost everywhere.} \quad (2)$$

The boundary condition ensures membership in H_0^1

On the Lipschitz domain Ω , the Sobolev trace of the continuous v_h equals its boundary restriction. The definition of V_h imposes $v_h|_{\partial\Omega} = 0$, so $v_h \in H_0^1(\Omega)$.

Together, these facts give $V_h \subset H_0^1(\Omega)$.

The Discrete Variational Problem

Use $\|v\|_V = \|v\|_{H^1(\Omega)}$. Assume $F \in V^*$ and

$$\begin{aligned} |a(w, v)| &\leq M \|w\|_V \|v\|_V, \\ a(v, v) &\geq \alpha \|v\|_V^2, \quad M, \alpha > 0, \quad w, v \in V. \end{aligned} \tag{1}$$

Here α is the coercivity constant, denoted by β earlier.

Galerkin problem on $V_h \subset V$

Find $u_h \in V_h$ such that

$$a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h. \tag{2}$$

The form remains continuous and coercive on V_h , so the discrete solution exists and is unique. These assumptions impose no smallness condition on the mesh size h .

From the Variational Problem to a Linear System

Let $N = \dim V_h$. Choose a basis $\phi_1, \dots, \phi_N$ and write

$$u_h = \sum_{j=1}^N U_j \phi_j, \quad \mathbf{u} = (U_1, \dots, U_N)^\top. \quad (1)$$

Testing with ϕ_i gives

$$\sum_{j=1}^N a(\phi_j, \phi_i) U_j = F(\phi_i), \quad i = 1, \dots, N. \quad (2)$$

The algebraic system

$$\begin{aligned} \mathbf{K} \mathbf{U} &= \mathbf{G}, \\ K_{ij} &= a(\phi_j, \phi_i), \quad G_i = F(\phi_i). \end{aligned} \quad (3)$$

We use $\mathbf{K}$ for the matrix, $A(x)$ for the coefficient field, and $K \in \mathcal{T}_h$ for an element.

Symmetry and Invertibility of the Discrete Matrix

Recall $K_{ij} = a(\phi_j, \phi_i)$, where row i corresponds to the test function. For $\mathbf{U} \in \mathbb{R}^N$, set $u_h = \sum_{j=1}^N U_j \phi_j$.

Expressing the quadratic form as $a(u_h, u_h)$

$$\begin{aligned} \mathbf{U}^T \mathbf{K} \mathbf{U} &= \sum_{i=1}^N U_i (\mathbf{K} \mathbf{U})_i = \sum_{i=1}^N \sum_{j=1}^N U_i K_{ij} U_j \\ &= \sum_{i=1}^N \sum_{j=1}^N U_i U_j a(\phi_j, \phi_i) \\ &= \sum_{i=1}^N U_i a\left(\sum_{j=1}^N U_j \phi_j, \phi_i\right) \\ &= a\left(\sum_{j=1}^N U_j \phi_j, \sum_{i=1}^N U_i \phi_i\right) \\ &= a(u_h, u_h) \geq \alpha \|u_h\|_V^2. \end{aligned} \tag{1}$$

We use linearity in each argument. **This step does not require symmetry.**

If $\mathbf{K} \mathbf{U} = 0$, the left side is 0, so $u_h = 0$. Linear independence of the basis gives $\mathbf{U} = 0$, so $\mathbf{K}$ is invertible.

If a is symmetric, $\mathbf{K}$ is also **symmetric positive definite**; otherwise, $\mathbf{K}$ may be nonsymmetric.

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence

The continuous and discrete solutions satisfy

$$\begin{aligned}a(u, v) &= F(v) \quad \forall v \in V, \\a(u_h, v_h) &= F(v_h) \quad \forall v_h \in V_h.\end{aligned}\tag{1}$$

Since $V_h \subset V$, subtracting the two identities gives, for $e = u - u_h$,

Orthogonality with respect to the bilinear form

$$a(e, v_h) = 0 \quad \text{for every } v_h \in V_h.\tag{2}$$

The error is orthogonal to every discrete test function with respect to a . If a is nonsymmetric, (2) does not imply $a(v_h, e) = 0$.

Assume the continuity and coercivity bounds

$$|a(w, v)| \leq M \|w\|_V \|v\|_V, \quad a(v, v) \geq \alpha \|v\|_V^2, \quad M, \alpha > 0. \quad (1)$$

Céa's lemma

Let $u \in V$ solve the weak problem, and let $u_h \in V_h \subset V$ solve the Galerkin problem. Then

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V. \quad (2)$$

The Galerkin error is bounded by M/α times the best approximation error in V_h . This estimate also applies to nonsymmetric bilinear forms.

Céa's Lemma: Proof (1)

Fix any $v_h \in V_h$ and set $e = u - u_h$. Write

$$e = (u - v_h) + (v_h - u_h). \quad (1)$$

Since $v_h - u_h \in V_h$, Galerkin orthogonality gives

$$a(e, v_h - u_h) = 0. \quad (2)$$

Coercivity, orthogonality, and continuity

$$\begin{aligned} \alpha \|e\|_V^2 &\leq a(e, e) \\ &= a(e, u - v_h) + a(e, v_h - u_h) \\ &= a(e, u - v_h) \\ &\leq |a(e, u - v_h)| \\ &\leq M \|e\|_V \|u - v_h\|_V. \end{aligned} \quad (3)$$

Céa's Lemma: Proof (2)

We have proved, for every $v_h \in V_h$,

$$\alpha \|e\|_V^2 \leq M \|e\|_V \|u - v_h\|_V. \quad (1)$$

- If $e = 0$, the assertion is immediate.
- Otherwise, $\alpha \|e\|_V > 0$. Dividing by this quantity gives

$$\|e\|_V \leq \frac{M}{\alpha} \|u - v_h\|_V. \quad (2)$$

Taking the infimum over $v_h \in V_h$ yields

The best-approximation bound

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V. \quad (3)$$

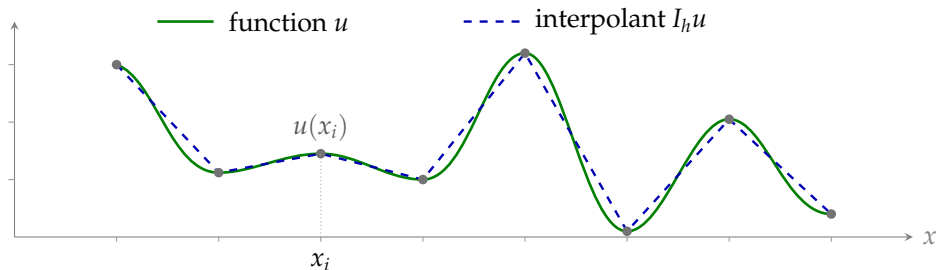
□

Table of Contents

- ① A Model Problem for FEM
- ② Weak Derivatives and Sobolev Spaces
- ③ Bilinear Forms and the Lax–Milgram Theorem
- ④ Application to Uniformly Elliptic Equations
- ⑤ Conforming Discretization and the Linear System
- ⑥ Galerkin Orthogonality and Best Approximation
- ⑦ Interpolation Estimates and Energy-Norm Convergence**

Interpolation: Matching Nodal Values

Interpolation constructs an approximation that agrees with a function at selected nodes.



Exact at the nodes

$$I_h u(x_i) = u(x_i) \quad \text{at every interpolation node } x_i. \quad (1)$$

Between nodes, $I_h u$ approximates u .

The Nodal Interpolation Operator

Let $a_1, \dots, a_{N_p}$ be degree- p Lagrange nodes on a simplex K , where $N_p = \dim \mathbb{P}_p(K)$. The nodal basis satisfies

$$\phi_i \in \mathbb{P}_p(K), \quad \phi_i(a_j) = \delta_{ij}. \quad (1)$$

Definition

For a continuous function u , define

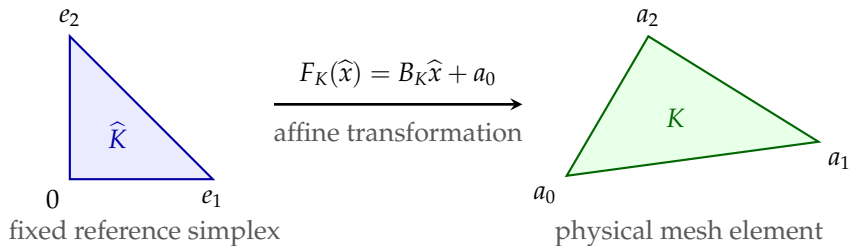
$$I_K : C^0(\bar{K}) \longrightarrow \mathbb{P}_p(K), \quad I_K u := \sum_{i=1}^{N_p} u(a_i) \phi_i. \quad (2)$$

I_K is the **operator**; $I_K u$ is the interpolating polynomial. In particular, $(I_K u)(a_j) = u(a_j)$.

- **Linearity:** $I_K(\alpha u + \beta v) = \alpha I_K u + \beta I_K v$.
- **Polynomial reproduction:** $I_K q = q$ for every $q \in \mathbb{P}_p(K)$.

Our Strategy: Analysis on a Reference Simplex

We prove interpolation estimates on one fixed reference simplex $\hat{K}$, then transfer them to each physical element K .



$$\hat{u} = u \circ F_K, \quad (I_K u) \circ F_K = \hat{I} \hat{u}. \quad (1)$$

- 1 Analyze $\hat{u} - \hat{I} \hat{u}$ on the fixed simplex $\hat{K}$.
- 2 Transfer the estimate to K , tracking how integrals and derivatives scale.

Local Estimates for Céa's Lemma

Use degree- p elements on conforming, shape-regular meshes with $h_K/\rho_K \leq \sigma$ and $h = \max_K h_K$. Assume

$$u \in H_0^1(\Omega) \cap H^m(\Omega), \quad 2 \leq m \leq p+1, \quad m \in \mathbb{N}, \quad m > d/2. \quad (1)$$

Here $\Omega \subset \mathbb{R}^d$ is a bounded polyhedral Lipschitz domain. Let $I_K u \in \mathbb{P}_p(K)$ be the degree- p nodal interpolant.

The required local interpolation estimates

$$\begin{aligned} \|u - I_K u\|_{L^2(K)} &\leq Ch_K^m |u|_{H^m(K)}, \\ |u - I_K u|_{H^1(K)} &\leq Ch_K^{m-1} |u|_{H^m(K)}. \end{aligned} \quad (2)$$

A Polynomial with Vanishing Residual Moments

Work on the fixed reference simplex $\widehat{K} = \text{conv}\{0, e_1, \dots, e_d\}$. For a multi-index $\gamma \in \mathbb{N}_0^d$, write $|\gamma| = \gamma_1 + \dots + \gamma_d$.

Moment lemma

For every $v \in H^m(\widehat{K})$ there is a unique $q \in \mathbb{P}_{m-1}(\widehat{K})$ such that

$$\int_{\widehat{K}} D^\gamma (v - q) \, d\widehat{x} = 0 \quad \text{for every } |\gamma| < m. \quad (1)$$

Set $N_m = \dim \mathbb{P}_{m-1} = \binom{d+m-1}{d}$ and define

$$T : \mathbb{P}_{m-1} \longrightarrow \mathbb{R}^{N_m}, \quad Tq = \left(\int_{\widehat{K}} D^\gamma q \, d\widehat{x} \right)_{|\gamma| < m}. \quad (2)$$

We need $Tq = (\int_{\widehat{K}} D^\gamma v)_{|\gamma| < m}$. It suffices to show that T is injective.

Existence and Uniqueness of the Moment Polynomial

Suppose $Tq = 0$. Expand

$$q(\hat{x}) = \sum_{|\beta| \leq m-1} c_\beta \hat{x}^\beta, \quad \gamma! := \gamma_1! \cdots \gamma_d!. \quad (1)$$

For $|\gamma| = m - 1$, all terms disappear except the matching monomial:

$$D^\gamma q = \gamma! c_\gamma, \quad 0 = \int_{\hat{K}} D^\gamma q = |\hat{K}| \gamma! c_\gamma. \quad (2)$$

Thus every coefficient of degree $m - 1$ vanishes. Repeat the argument for degrees $m - 2, m - 3, \dots, 0$ to conclude that $q = 0$.

Finite-dimensional conclusion

The linear map T is injective between spaces of the same dimension, hence bijective. Every prescribed vector of moments has exactly one preimage $q \in \mathbb{P}_{m-1}$.

Repeated Application of Poincaré's Inequality

Put $w = v - q$. Each $D^\gamma w$, $|\gamma| < m$, belongs to $H^1(\widehat{K})$ and has zero mean. Therefore

$$\|D^\gamma w\|_{L^2(\widehat{K})} \leq C_{P,\widehat{K}} \|\nabla D^\gamma w\|_{L^2(\widehat{K})}. \quad (1)$$

With $|w|_{H^j}^2 := \sum_{|\gamma|=j} \|D^\gamma w\|_{L^2}^2$ and $|w|_{H^0} := \|w\|_{L^2}$, summing gives

$$|w|_{H^j} \leq C_j |w|_{H^{j+1}} \quad (0 \leq j < m). \quad (2)$$

A bound in terms of the highest derivatives

Iterating and using $D^\gamma q = 0$ for $|\gamma| = m$ yields

$$\|v - q\|_{H^m(\widehat{K})} \leq C_Q |v - q|_{H^m(\widehat{K})} = C_Q |v|_{H^m(\widehat{K})}. \quad (3)$$

This argument uses the mean-zero form of Poincaré's inequality and requires no boundary condition on w .

Nodal Interpolation on the Reference Simplex

Let $\hat{a}_1, \dots, \hat{a}_{N_p}$ be the standard degree- p Lagrange nodes, where $N_p = \dim \mathbb{P}_p = \binom{p+d}{d}$. The associated nodal basis satisfies

$$\hat{\phi}_i \in \mathbb{P}_p(\hat{K}), \quad \hat{\phi}_i(\hat{a}_j) = \delta_{ij}. \quad (1)$$

Reference interpolation

$$\hat{I}v := \sum_{i=1}^{N_p} v(\hat{a}_i) \hat{\phi}_i. \quad (2)$$

The nodal basis functions span $\mathbb{P}_p$, so every $q \in \mathbb{P}_p$ has the expansion $q = \sum_i q(\hat{a}_i) \hat{\phi}_i$. Hence

$$\boxed{\hat{I}q = q \quad \text{for every } q \in \mathbb{P}_p.} \quad (3)$$

In particular, this holds for the moment polynomial because $m - 1 \leq p$.

Boundedness of the Interpolation Operator

Because $m > d/2$, Sobolev embedding on the fixed simplex gives

$$H^m(\widehat{K}) \hookrightarrow C^0(\overline{\widehat{K}}), \quad |v(\widehat{a}_i)| \leq C_E \|v\|_{H^m(\widehat{K})}. \quad (1)$$

For $j = 0, 1$, the nodal expansion gives

$$\begin{aligned} \|\widehat{I}v\|_{H^j(\widehat{K})} &\leq \sum_{i=1}^{N_p} |v(\widehat{a}_i)| \|\widehat{\phi}_i\|_{H^j(\widehat{K})} \\ &\leq C_{I,j} \|v\|_{H^m(\widehat{K})}. \end{aligned} \quad (2)$$

The regularity needed for point evaluation

$\widehat{I} : H^m(\widehat{K}) \rightarrow H^j(\widehat{K})$ is bounded. The bound uses the H^m norm of the input, even when $j = 0$ or 1 .

We use $H^0 = L^2$. The constants depend on the fixed reference basis.

The Reference-Element Error Estimate

Take the moment polynomial $q \in \mathbb{P}_{m-1} \subset \mathbb{P}_p$ and put $w = v - q$. Since $\widehat{I}q = q$,

$$v - \widehat{I}v = w - \widehat{I}w. \quad (1)$$

For $j = 0, 1$, boundedness and repeated use of Poincaré's inequality give

$$\begin{aligned} \|v - \widehat{I}v\|_{H^j} &\leq \|w\|_{H^j} + \|\widehat{I}w\|_{H^j} \\ &\leq (1 + C_{I,j}) \|w\|_{H^m} \\ &\leq (1 + C_{I,j}) C_Q |v|_{H^m}. \end{aligned} \quad (2)$$

Both estimates on the fixed reference element

$$\|v - \widehat{I}v\|_{L^2(\widehat{K})} \leq C |v|_{H^m(\widehat{K})}, \quad |v - \widehat{I}v|_{H^1(\widehat{K})} \leq C |v|_{H^m(\widehat{K})}. \quad (3)$$

The Affine Map to a Physical Element

Let $a_0, \dots, a_d$ be the vertices of K . Set

$$F_K(\hat{x}) = B_K \hat{x} + a_0, \quad B_K = [a_1 - a_0 \ \cdots \ a_d - a_0], \quad J_K := |\det B_K| > 0. \quad (1)$$

Then $F_K : \hat{K} \rightarrow K$ is bijective. Define the nodes and basis on K by

$$a_{K,i} = F_K(\hat{a}_i), \quad \phi_{K,i} = \hat{\phi}_i \circ F_K^{-1}. \quad (2)$$

Geometric bounds for the scaling argument

In the Euclidean operator norm,

$$\|B_K\| \leq \sqrt{d} h_K, \quad \|B_K^{-1}\| \leq \frac{\sqrt{d}}{\rho_K} \leq \frac{\sqrt{d} \sigma}{h_K}. \quad (3)$$

We first use these bounds to derive the local estimates, then prove them from the geometry of the simplex.

Interpolation Commutes with the Affine Pullback

Write $\hat{u} = u \circ F_K$ and $I_K u(x) = \sum_i u(a_{K,i}) \phi_{K,i}(x)$. Then

$$\begin{aligned}(I_K u)(F_K(\hat{x})) &= \sum_i u(F_K(\hat{a}_i)) \phi_{K,i}(F_K(\hat{x})) \\ &= \sum_i \hat{u}(\hat{a}_i) \hat{\phi}_i(\hat{x}) \\ &= \hat{I}\hat{u}(\hat{x}).\end{aligned}\tag{1}$$

The same error in two coordinate systems

Put $e_K = u - I_K u$ and $\hat{e}_K = e_K \circ F_K$. Thus

$$(I_K u) \circ F_K = \hat{I}\hat{u}, \quad \hat{e}_K = \hat{u} - \hat{I}\hat{u}.\tag{2}$$

Changing Variables in the Local L^2 Error

Since $x = F_K(\hat{x})$ and $dx = J_K d\hat{x}$,

$$\begin{aligned}\|e_K\|_{L^2(K)}^2 &= \int_K |e_K(x)|^2 dx \\ &= J_K \int_{\hat{K}} |e_K(F_K(\hat{x}))|^2 d\hat{x} \\ &= J_K \|\hat{e}_K\|_{L^2(\hat{K})}^2.\end{aligned}\tag{1}$$

Exact L^2 scaling identity

$$\|e_K\|_{L^2(K)} = J_K^{1/2} \|\hat{e}_K\|_{L^2(\hat{K})}.\tag{2}$$

The absolute determinant makes this identity independent of the vertex ordering.

Changing Variables in the Local H^1 Seminorm

The chain rule gives

$$\nabla_{\hat{x}} \hat{e}_K = B_K^T (\nabla_x e_K) \circ F_K, \quad (\nabla_x e_K) \circ F_K = B_K^{-T} \nabla_{\hat{x}} \hat{e}_K. \quad (1)$$

Therefore

$$\begin{aligned} |e_K|_{H^1(K)}^2 &= J_K \int_{\hat{K}} |B_K^{-T} \nabla_{\hat{x}} \hat{e}_K|^2 d\hat{x} \\ &\leq J_K \left\| B_K^{-1} \right\|^2 |\hat{e}_K|_{H^1(\hat{K})}^2. \end{aligned} \quad (2)$$

The inverse matrix in the gradient estimate

$$|e_K|_{H^1(K)} \leq J_K^{1/2} \left\| B_K^{-1} \right\| |\hat{e}_K|_{H^1(\hat{K})}. \quad (3)$$

Here $|\cdot|_{H^1} = \|\nabla(\cdot)\|_{L^2}$ is a seminorm on $H^1(K)$.

Scaling of Higher-Order Derivatives

For $\hat{u}(\hat{x}) = u(B_K \hat{x} + a_0)$,

$$\nabla \hat{u} = B_K^T (\nabla u) \circ F_K, \quad D^2 \hat{u} = B_K^T [(D^2 u) \circ F_K] B_K. \quad (1)$$

Since F_K is affine, each of the m derivatives contributes one factor of B_K , with no derivatives of B_K . Thus

$$\sum_{|\gamma|=m} |D^\gamma \hat{u}(\hat{x})|^2 \leq C \|B_K\|^{2m} \sum_{|\beta|=m} |D^\beta u(F_K(\hat{x}))|^2. \quad (2)$$

Integrate using $d\hat{x} = J_K^{-1} dx$

$$|\hat{u}|_{H^m(\hat{K})} \leq C \|B_K\|^m J_K^{-1/2} |u|_{H^m(K)}. \quad (3)$$

These identities hold for smooth functions and extend to $H^m(K)$ by density. The constant depends only on d and m .

The Local L^2 Interpolation Estimate

Apply the reference estimate to $\widehat{e}_K = \widehat{u} - \widehat{I}\widehat{u}$:

$$\begin{aligned}\|u - I_K u\|_{L^2(K)} &= J_K^{1/2} \|\widehat{e}_K\|_{L^2(\widehat{K})} \\ &\leq C J_K^{1/2} |\widehat{u}|_{H^m(\widehat{K})} \\ &\leq C J_K^{1/2} \|B_K\|^m J_K^{-1/2} |u|_{H^m(K)} \\ &= C \|B_K\|^m |u|_{H^m(K)} \\ &\leq C h_K^m |u|_{H^m(K)}.\end{aligned}\tag{1}$$

The local L^2 error bound

$$\|u - I_K u\|_{L^2(K)} \leq C h_K^m |u|_{H^m(K)}.\tag{2}$$

The factors $J_K^{1/2}$ and $J_K^{-1/2}$ cancel exactly.

The Local H^1 Interpolation Estimate

The same reference estimate, now for the gradient error, gives

$$\begin{aligned} |u - I_K u|_{H^1(K)} &\leq J_K^{1/2} \left\| B_K^{-1} \right\| |\widehat{e}_K|_{H^1(\widehat{K})} \\ &\leq C J_K^{1/2} \left\| B_K^{-1} \right\| |\widehat{u}|_{H^m(\widehat{K})} \\ &\leq C J_K^{1/2} \left\| B_K^{-1} \right\| \|B_K\|^m J_K^{-1/2} |u|_{H^m(K)} \\ &\leq C h_K^{-1} h_K^m |u|_{H^m(K)}. \end{aligned} \tag{1}$$

The local H^1 seminorm bound

$$|u - I_K u|_{H^1(K)} \leq C h_K^{m-1} |u|_{H^m(K)}. \tag{2}$$

Shape regularity makes the constant uniform in K and h . It remains to prove the two geometric bounds used above.

Geometry I: Bounding the Affine Matrix

Write $b_i = a_i - a_0$, so $B_K = [b_1 \ \cdots \ b_d]$ and $|b_i| \leq h_K$. For $z = (z_1, \dots, z_d)^\top \in \mathbb{R}^d$,

$$\begin{aligned} |B_K z| &= \left| \sum_{i=1}^d z_i b_i \right| \leq \sum_{i=1}^d |z_i| |b_i| \\ &\leq h_K \sum_{i=1}^d |z_i| \\ &\leq \sqrt{d} h_K \left(\sum_{i=1}^d |z_i|^2 \right)^{1/2}. \end{aligned} \tag{1}$$

Take the supremum over nonzero z

$$\|B_K\| := \sup_{z \neq 0} \frac{|B_K z|}{|z|} \leq \sqrt{d} h_K. \tag{2}$$

Geometry II: Rows of the Inverse Matrix

Write the rows of B_K^{-1} as $r_i^\top$ and use the inverse map:

$$B_K^{-1} = \begin{pmatrix} r_1^\top \\ \vdots \\ r_d^\top \end{pmatrix}, \quad \hat{x} = B_K^{-1}(x - a_0). \quad (1)$$

The physical barycentric coordinates satisfy

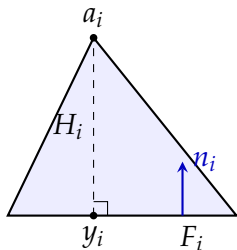
$$x = \left(1 - \sum_{i=1}^d \hat{x}_i\right) a_0 + \sum_{i=1}^d \hat{x}_i a_i, \quad \lambda_i(x) = \hat{x}_i = r_i^\top (x - a_0) \quad (1 \leq i \leq d). \quad (2)$$

Barycentric gradients determine the inverse matrix

$$\nabla \lambda_i = r_i, \quad B_K^{-1} = \begin{pmatrix} (\nabla \lambda_1)^\top \\ \vdots \\ (\nabla \lambda_d)^\top \end{pmatrix}. \quad (3)$$

Geometry III: The Altitude and the Gradient

Let F_i be the face opposite a_i , and let y_i be the orthogonal projection of a_i onto its supporting hyperplane $\text{aff}(F_i)$.



n_i : inward unit normal.

Set $H_i = |a_i - y_i|$ and $a_i - y_i = H_i n_i$. Since $\lambda_i = 0$ on the face hyperplane, $\nabla \lambda_i$ is normal to it. Also,

$$\begin{aligned} 1 &= \lambda_i(a_i) - \lambda_i(y_i) \\ &= \nabla \lambda_i \cdot (a_i - y_i) \\ &= H_i \nabla \lambda_i \cdot n_i. \end{aligned} \tag{1}$$

Therefore

$$\boxed{\nabla \lambda_i = \frac{n_i}{H_i}, \quad |\nabla \lambda_i| = \frac{1}{H_i}.} \tag{2}$$

The projection y_i may lie outside the face. Using its supporting hyperplane also covers obtuse simplices.

Geometry IV: Bounding the Inverse Matrix

The incenter lies a distance ρ_K from the hyperplane containing F_i . The largest distance from this hyperplane to a point in K is H_i , so $\rho_K \leq H_i$. Using the Frobenius norm

$$\|B\|_F^2 = \sum_{i,j} |B_{ij}|^2,$$

$$\left\|B_K^{-1}\right\|_F^2 = \sum_{i=1}^d |\nabla \lambda_i|^2 = \sum_{i=1}^d \frac{1}{H_i^2} \leq \frac{d}{\rho_K^2}. \quad (1)$$

Operator norm and shape regularity

Since $\|B\|_2 \leq \|B\|_F$,

$$\left\|B_K^{-1}\right\| \leq \frac{\sqrt{d}}{\rho_K} \leq \frac{\sqrt{d} \sigma}{h_K}. \quad (2)$$

Together with $\|B_K\| \leq \sqrt{d} h_K$, this completes both local interpolation estimates with constants independent of the mesh size.

Assembling an Admissible Global Interpolant

Define $(I_h u)|_K = I_K u$. Since $m > d/2$, the exact solution has a continuous representative on $\overline{\Omega}$, so each nodal value is well defined.

Continuity across an interior face

Both elements use the same nodes and the same values of u on the face. Unisolvence on the face ensures that the two polynomial traces coincide. Hence $I_h u$ is globally continuous.

The homogeneous boundary condition

For $u \in H_0^1(\Omega) \cap C^0(\overline{\Omega})$, the continuous boundary restriction agrees with its zero trace. All boundary nodal values vanish, so the interpolated polynomial vanishes on every boundary face.

Thus $I_h u \in V_h \subset H_0^1(\Omega)$ and $u - I_h u \in H_0^1(\Omega)$. We can therefore use $I_h u$ as a comparison function in C ea's lemma.

From Local to Global Interpolation Error

Because the element interiors are disjoint, we can sum the squared local errors:

$$\begin{aligned}\|u - I_h u\|_{L^2(\Omega)}^2 &\leq C_{\text{app}}^2 \sum_K h_K^{2m} |u|_{H^m(K)}^2 \\ &\leq C_{\text{app}}^2 h^{2m} |u|_{H^m(\Omega)}^2, \\ |u - I_h u|_{H^1(\Omega)}^2 &\leq C_{\text{app}}^2 \sum_K h_K^{2(m-1)} |u|_{H^m(K)}^2 \\ &\leq C_{\text{app}}^2 h^{2(m-1)} |u|_{H^m(\Omega)}^2.\end{aligned}\tag{1}$$

Take square roots

$$\|u - I_h u\|_{L^2} \leq C_{\text{app}} h^m |u|_{H^m}, \quad |u - I_h u|_{H^1} \leq C_{\text{app}} h^{m-1} |u|_{H^m}.\tag{2}$$

Here $\sum_K |u|_{H^m(K)}^2 = |u|_{H^m(\Omega)}^2$. No factor depends on the number of elements, and quasi-uniformity is not required.

Returning to the Norm in Céa's Lemma

The previous slides use $\|v\|_V = \|v\|_{H^1(\Omega)}$. Define the gradient norm on $V = H_0^1(\Omega)$. Poincaré's inequality gives

$$\|v\|_E := \|\nabla v\|_{L^2(\Omega)}, \quad \|v\|_E \leq \|v\|_V \leq C_\Omega \|v\|_E, \quad C_\Omega = \sqrt{1 + C_{P,\Omega}^2}. \quad (1)$$

Since $u - I_h u \in V$, Poincaré's inequality and Céa's lemma give

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \|u - I_h u\|_V \leq \frac{MC_\Omega C_{\text{app}}}{\alpha} h^{m-1} |u|_{H^m(\Omega)}. \quad (2)$$

$H^1(\Omega)$ error bound

$$\|u - u_h\|_V \leq \frac{MC_\Omega C_{\text{app}}}{\alpha} h^{m-1} |u|_{H^m(\Omega)} \quad (3)$$

One may choose the largest polynomial degree allowed by the solution's regularity. For degree- p elements with $u \in H^{p+1}(\Omega)$, take $m = p + 1$ in (3) to obtain an $O(h^p)$ error bound.

Thank you for your attention!

Acknowledgement

This presentation was prepared using ChatGPT Pro and the Astra model by OpenAI.