Numerical Methods for Partial Differential Equations (PDE) (6)

Finite Element Methods (FEM)

Brief Comparison with Other Methods

Finite Difference (FD) Method:

FD approximates an *operator* (e.g., the derivative) and solves a problem on a set of *points* (the grid)

Finite Element Method (FEM):

FE uses exact operators but approximates the *solution basis functions.* Also, FE solves a problem on the *interiors* of grid cells (and optionally on the gridpoints as well).

Brief Comparison with Other Methods

Spectral Methods:

Spectral methods use global basis functions to approximate a solution across the entire domain.

V_n space of algebraic or trigonometric polynomials on Ω.

Finite Element Method (FEM):

FE methods use compact basis functions to approximate a solution on individual elements.

V_n space of the piecewise polynomials

Overview of the Finite Element Method

$$(S) \Leftrightarrow (W) \approx (G) \Leftrightarrow (M)$$

Strong form	Weak form	Galerkin approx.	Matrix form

The set of governing PDE's, with boundary conditions, is called the "strong form" of the problem.

The weak form is a *variational statement* of the problem in which we integrate against a *test function*. This has the effect of relaxing the problem; instead of finding an exact solution everywhere, we are finding a solution that satisfies the strong form on average over the domain.

Methods:
- Spectral
- FEM

The Galerkin's method for elliptic problems

Weak form of a generic elliptic problem on the domain $\Omega \subset \Re^d$ (variational form)

$$(*)\quad \boxed{\begin{array}{l}\text{Find} \quad u \in V: \\ a(u,v) = F(v) \qquad\qquad \forall v \in V\end{array}}$$

Let V be a suitable Hilbert space, subspace of $H^1(\Omega)$

$\boldsymbol{a : V \times V \to \Re,}$ **Coercive, continuos and bilinear FORM**

$\boldsymbol{F : V \to \Re,}$ **Linear, continuous FUNCTIONAL**

Under these assumptions the Lax-Milgram theorem ensures existence and uniqueness of the solution for problem (*).

The Galerkin's method for elliptic problems

- The Galerkin method for the numerical approximation of the problem (*) consists in computing an approximate solution

$$\boldsymbol{u_h \in V_h}$$

where V_h is a family of spaces depending on a positive parameter h, such that

$$V_h \subset V, \qquad \dim V_h = N_h < \infty$$

- The approximate problem takes the form

(**)

$$\boxed{\begin{array}{ll} \text{Find} \quad u_h \in V_h : & \\ a(u_h, v_h) = F(v_h) & \forall v_h \in V_h \end{array}}$$

Galerkin Form

The Galerkin's Method for elliptic problems

Let $\{\varphi_i, \quad i=1,2,...,N_h\}$ be a base for the space V_h , then to satisfy (**) for every function $\boxed{\forall v_h \in V_h}$ it is sufficient to prove that every function of the base satisfy:

$$\boxed{a(u_h,\varphi_i) = F(\varphi_i) \qquad i=1,2,...,N_h}$$

Since $u_h \in V_h$ we have $u_h(x) = \sum_{j=1}^{N_h} u_j \varphi_j(x),$

where u_j , $j=1,2,\dots,N_h$ are unknown coefficients,

$$\sum_{j=1}^{N_h} u_j a(\varphi_j,\varphi_i) = F(\varphi_i) \qquad i=1,2,...,N_h$$

for the linearity of a(u,v). .

The Galerkin Method: algebraic formulation

$$\sum_{j=1}^{N_h} u_j a\left(\varphi_j, \varphi_i\right) = F\left(\varphi_i\right) \qquad i = 1,2,\ldots,N_h$$

Let A be the matrix of elements $\quad a_{ij} \equiv a\left(\phi_j, \phi_i\right)$

Stiffness Matrix

and f be a vector of elements $\quad f_i \equiv F\left(\phi_i\right)$

Solve the linear system

$$\boxed{\mathbf{A\,u = f}}$$

for the unknown coefficients u_i

The Galerkin Method: algebraic formulation

- The characteristics of the stiffness matrix are independent on the chosen basis for V_h and depend exclusively on the weak problem that is approaching.

Theorem The matrix A associated with the discretization of an elliptic problem with the Galerkin method is **positive definite.**

Theorem The matrix A is symmetric if and only if the bilinear form a (.,.) is **symmetric**.

- The **condition number**, the **structure** and the **sparsity** of the stiffness matrix depend on the chosen basis for V_h.

Analysis of the Galerkin's Method for elliptic problems

- The solution u_h exists and is unique (Lax-Milgram th.);

- Stability of the discrete solution u_h

  The stability of the method guarantees that the norm $||u_h||_V$ of the solution is bounded from above for h that tends uniformly to zero.

- Convergence of the discrete solution u_h to the exact solution u of the problem.

  As long as you take h sufficiently small, we can approximate how accurate you want the exact solution u with the Galerkin solution u_h

Finite Element Method 1D

- Goal: approximate the space $H^1(a,b)$ with a finite dimensional space whose basis depend on a parameter h
- Discretize the domain: $\Omega \equiv (a,b)$

  by subdividing Ω into a set T_h of subintervals $K_j \equiv (x_{j-1}, x_j)$

  of width $h_j = x_j - x_{j-1}$

$$a = x_0 < x_1 < \ldots < x_N < x_{N+1} = b$$

$$h = \max_j \quad h_j$$

- The functions in $H^1(a,b)$ are continuous functions on [a,b],
- A possible choice for the space V_h : piecewise polynomials of degree r

$$\boxed{X_h^r \equiv \left\{ v_h \in C^0\left(\bar{\Omega}\right): \quad v_h|_{K_j} \in \mathbf{P}_r, \quad \forall K_j \in T_h \right\}}$$

Finite Element Method 1D

$$X_h^r \equiv \left\{ v_h \in C^0\left(\bar{\Omega}\right): \quad v_h\big|_{K_j} \in \mathbf{P}_r, \quad \forall K_j \in \mathrm{T}_h \right\}$$

$\mathbf{P_r}$ Space of polynomials of degree less or equal to r.

- Spaces $\boldsymbol{X_h^r}$ are subspaces of $H^1(a,b)$, as they are made from derivable functions except in a finite number of points (nodes x_i).
- Basis Functions φ_i : the support of the generic basic function has non-empty intersection with a small number of other basic functions (many elements of the matrix A are zero). **Lagrangian basis** (the coefficients of the expansion of the generic function itself are the values of $\boldsymbol{v_h \in X_h^r}$ at appropriate points – nodes-).

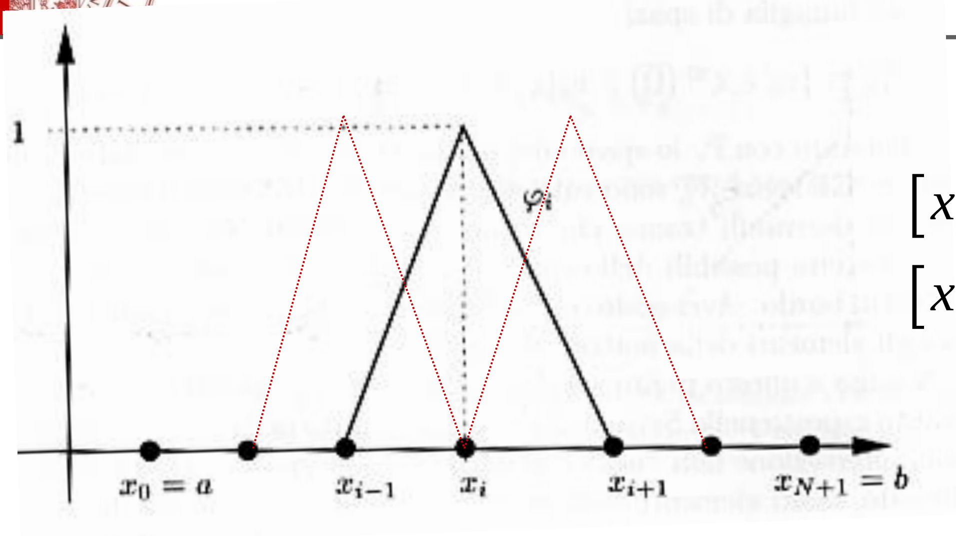
Finite Element Method 1D

Space X_h^1

- Space of piecewise linear continuous functions on the partition T_h of [a,b]
- *Degree of freedom*: number N+2 of nodes in the partition
- Base functions φ_i, i=0,1,…,N+1.

$$\varphi_i \in X_h^1 \quad \text{such that} \quad \varphi_i\left(x_j\right) = \delta_{ij} \quad i, j = 0, 1, \ldots, N+1$$

Finite Element Method 1D



Local support

$$[x_{i-1}, x_i] \quad \text{and} \quad [x_i, x_{i+1}] \quad \text{if i} \neq 0 \text{ or i} \neq \text{N}+1$$

$$[x_0, x_1] \quad \text{or} \quad [x_N, x_{N+1}] \quad \text{if } \ i = 0 \ \text{ or } \ i = \text{N}+1$$

$$\varphi_i(x) = \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & \text{in} \quad x_{i-1} \le x \le x_i \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} & \text{in} \quad x_i \le x \le x_{i+1} \\ 0 & \text{elesewhere} \end{cases}$$

The Stiffness matrix is tridiagonal:

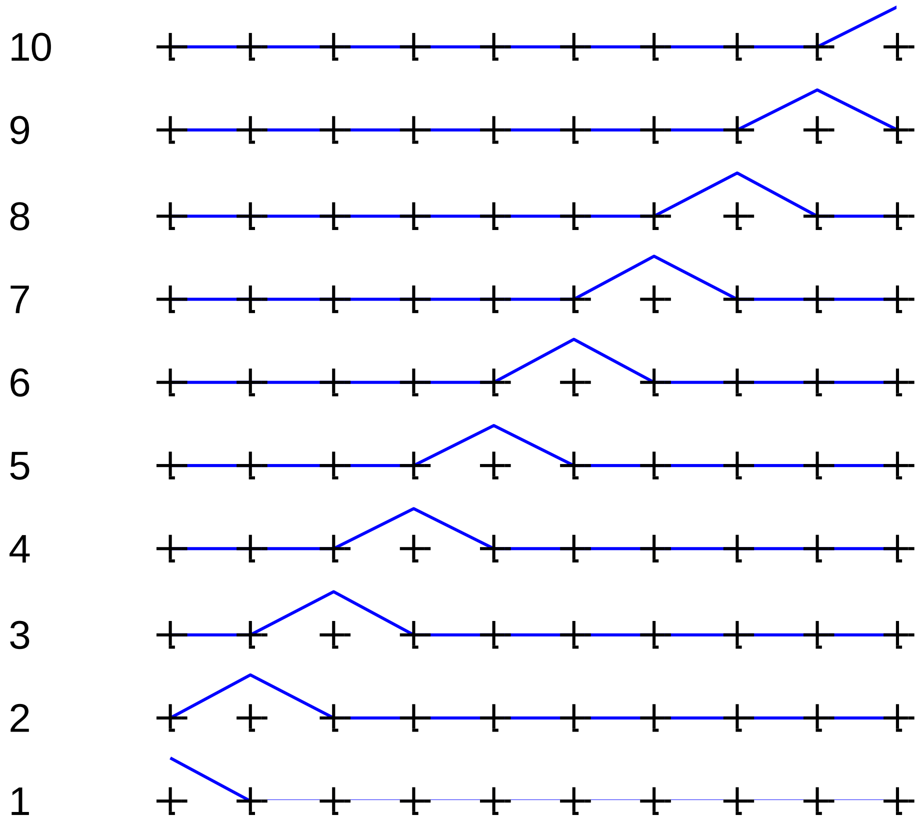
$$\boldsymbol{a_{ij} = 0 \quad se \quad j \notin \{i-1, i, i+1\}}$$

Only φ_{i-1} and φ_{i+1} have support overlapping the support of φ_i.

FEM 1D Basis functions

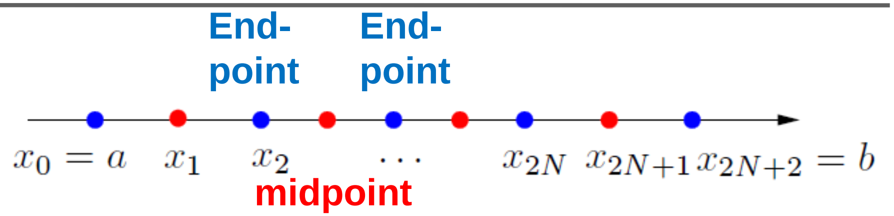
$$\varphi_i(x) = \begin{cases} 1 & for \quad x = x_i \\ 0 & for \quad x = x_j,\ j \neq i \end{cases}$$

blue lines – basis functions φ_i



Finite Element Method 1D

Space X_h^2



- Functions in X_h^2 are piecewise polynomials of degree 2 with local support K defined by three distinct points in T_h (two end-points and the midpoint).
- Continuity is guaranteed if we consider the two end-points of K.
- Degrees of freedom of X_h^2 :

  end-points of the intervals and midpoints

The Lagrangian base consists of the functions:

$$\varphi_i \in X_h^2 \quad \text{such that} \quad \varphi_i\left(x_j\right) = \delta_{ij} \qquad i, j = 0, 1, \ldots, 2N+2$$

Finite Element Method 1D

$$\varphi_i(x) = \begin{cases} \dfrac{(x_{i+1}-x)(x-x_{i-1})}{(x_{i+1}-x_i)(x_i-x_{i-1})} & \text{if} \quad x_{i-1} \le x \le x_{i+1} \\ 0 & \text{elsewhere} \end{cases}$$

(i odd)

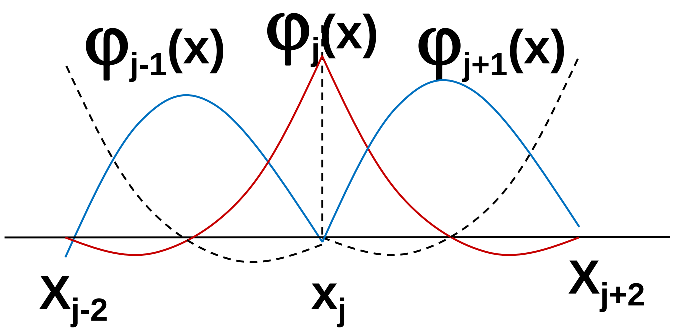
midpoint

$$\varphi_i(x) = \begin{cases} \dfrac{(x-x_{i-1})(x-x_{i-2})}{(x_i-x_{i-1})(x_i-x_{i-2})} & \textit{for} \quad x_{i-2} \le x \le x_i \\ \dfrac{(x_{i+1}-x)(x_{i+2}-x)}{(x_{i+1}-x_i)(x_{i+2}-x_i)} & \textit{for} \quad x_i \le x \le x_{i+2} \\ 0 & \textit{elsewhere} \end{cases}$$

(i even)

end-points

Consider three consecutive nodes



Stiffness matrix A has 5 nonzero diagonals

The choice of the space V_h

$$V_h = \left\{ v_h \in X_h^r : \ v_h(a) = v_h(b) = 0 \right\} \quad N_h = \dim(V_h) = (N+1)*r-1$$

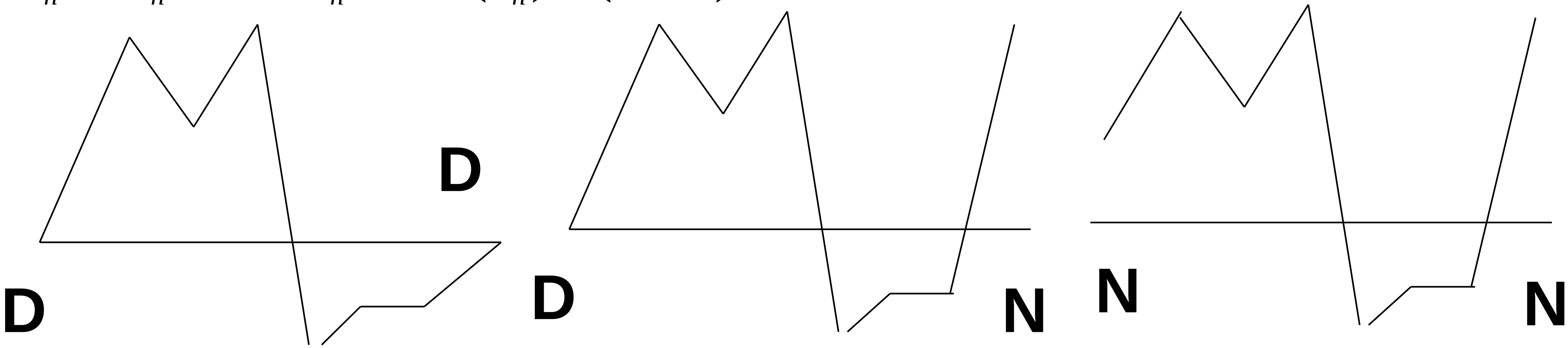
for Dirichlet BC in a and b

For r=1 is the number of internal nodes

$$V_h = \left\{ v_h \in X_h^r : \ v_h(a) = 0 \right\} \quad N_h = \dim(V_h) = (N+1)*r$$

for Dirichlet BC in a and Neumann BC in b

$$V_h = X_h^r \qquad N_h = \dim(V_h) = (N+1)*r+1$$ for Neumann BC in a and b



The Galerkin Method: FEM

$$\sum_{j=1}^{N_h} u_j a\left(\varphi_j, \varphi_i\right) = F\left(\varphi_i\right) \qquad i = 1,2,\ldots,N_h$$

The Galerkin-FEM problem is solved by the linear system

$$\boxed{\mathbf{A\ u = f}}$$

$$A \in \mathbb{R}^{N_h \times N_h} \qquad A_{ij} = a(\varphi_i, \varphi_j)$$

Stiffness Matrix

$$\mathrm{u} = (u_1, \ldots, u_{N_h})^T \in \mathbb{R}^{N_h}, \mathrm{f} = (f_1, \ldots, f_{N_h})^T \in \mathbb{R}^{N_h}, f_i \equiv F\left(\varphi_i\right)$$

A is sparse with structure:

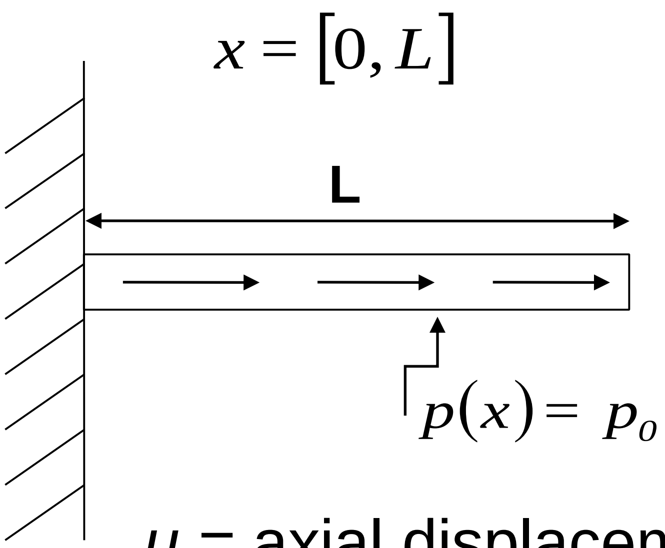
r = 1 A is tridiagonal,

r = 2 A is pentadiagonal,

r = 3 A is eptadiagonal

Example Problem 1: 1D Poisson equation

Axial deformation of a bar subjected to a uniform load



Strong Form:

$$EA\frac{d^2u}{dx^2} = p_0$$

$$BC: \begin{cases} u(0) = 0 \\ EA\left.\frac{du}{dx}\right|_{x=L} = 0 \end{cases}$$

u = axial displacement

E=Young's modulus = 1

A=Cross-sectional area = 1

Weak Form

We now reformulate the problem into the weak form.

$$\frac{d^2u}{dx^2} = p_0$$ **Strong Form**

$$\frac{d^2u}{dx^2} - p_0 = 0$$ **Residual** ***R=0***

$$\int_0^L \left(\frac{d^2u}{dx^2} - p_0 \right) v dx = 0$$ **Weak Form**

v is our test function

We will choose the test function later. A solution of the strong form will also satisfy the weak form, but not vice versa.

Weak Form

Returning to the weak form:

$$\int_0^L \left(\frac{d^2u}{dx^2} - p_0 \right) v dx = 0$$

$$\int_0^L \frac{d^2u}{dx^2} v dx = \int_0^L p_0 v dx$$

Integrate LHS by parts:

$$= -\int_0^L \frac{du}{dx}\frac{dv}{dx} dx + \left[v(x) \frac{du}{dx} \right]_{x=0}^{x=L}$$

$$= -\int_0^L \frac{du}{dx}\frac{dv}{dx} dx + v(L) \left. \frac{du}{dx} \right|_{x=L} - v(0) \left. \frac{du}{dx} \right|_{x=0}$$

Weak Form

Recall the boundary conditions on *u* and *v*:

$$u(0)=0$$

$$\left.\frac{du}{dx}\right|_{x=L}=0$$

$$v(0)=0$$

Choosing the test function:

We can choose any *v* we want, so let's choose *v* such that it satisfies *homogeneous* boundary conditions wherever the actual solution satisfies *Dirichlet* boundary conditions
So in our example, *u*(0)=0 so let *v*(0)=0.

Hence,

$$-\int_0^L \frac{du}{dx}\frac{dv}{dx}dx + \cancel{v(L)\left.\frac{du}{dx}\right|_{x=L}} - \cancel{v(0)\left.\frac{du}{dx}\right|_{x=0}}$$

$$\boxed{\int_0^L \frac{du}{dx}\frac{dv}{dx}dx = \int_0^L p_0 v\,dx}$$

The weak form satisfies Neumann conditions automatically!

Remarks on Variational Problem

$$\int_0^L \frac{du}{dx}\frac{dv}{dx}dx = \int_0^L p_0 v dx$$

u's first derivative is integrable $\Rightarrow u \in H^1$

v also in H^1, with added requirement that it vanishes on boundaries

$$\Rightarrow v \in H_0^1$$

Find $u \in H^1$ such that $a(u,v) = F(v) \quad \forall v \in H_0^1$

H^1 is a Sobolev space, which is a subspace of a Hilbert space.

H^1 is a space of functions that can include discontinuous functions and functions with singularities.

$\Rightarrow$ singularities and discontinuities in our solution OK!

Galerkin's Method

We still haven't done the "finite element method" yet, we have just restated the problem in the weak formulation.

So what makes it "finite elements"?

Solving the problem locally on elements

Finite-dimensional approximation to an infinite-dimensional space → Galerkin's Method

Galerkin's Method

Choose finite basis $\{\phi_i\}_{i=1}^{N}$

Then,

$$u(x) = \sum_{j=1}^{N} c_j \phi_j(x), \qquad c_j \text{ unkowns to solve for}$$

$$v(x) = \sum_{j=1}^{N} b_j \phi_j(x), \qquad b_j \text{ arbitrarily chosen}$$

Insert these into our weak form:

$$\int_0^L \frac{du}{dx}\frac{dv}{dx}dx = \int_0^L p_0 v dx$$

$$\int_0^L \sum_{j=1}^{N} c_j \frac{d\phi_j}{dx}(x) \sum_{i=1}^{N} b_i \frac{d\phi_i}{dx}(x) dx = \int_0^L p_0 \sum_{i=1}^{N} b_i \phi_i(x) dx$$

Galerkin's Method

$$\int_0^L \sum_{j=1}^N c_j \frac{d\phi_j}{dx}(x) \sum_{i=1}^N b_i \frac{d\phi_i}{dx}(x) dx = \int_0^L p_0 \sum_{i=1}^N b_i \phi_i(x) dx$$

Rearranging:

$$\sum_{i=1}^N b_i \sum_{j=1}^N c_j \int_0^L \frac{d\phi_j}{dx} \frac{d\phi_i}{dx} dx = \sum_{i=1}^N b_i \int_0^L p_0 \phi_i dx$$

Cancelling :

$$\sum_{j=1}^N c_j \int_0^L \frac{d\phi_j}{dx} \frac{d\phi_i}{dx} dx = \int_0^L p_0 \phi_i dx \qquad \forall i$$

Galerkin's Method

$$\sum_{j=1}^{N} c_j \int_0^L \frac{d\varphi_j}{dx}\frac{d\varphi_i}{dx}\,dx = \int_0^L p_0 \varphi_i \,dx$$

We now have a matrix problem $\mathbf{Kc} = \mathbf{F}$, where c_j

is a vector of unknowns,

$$K_{ij} = \int_0^L \frac{d\varphi_j}{dx}\frac{d\varphi_i}{dx}\,dx,$$

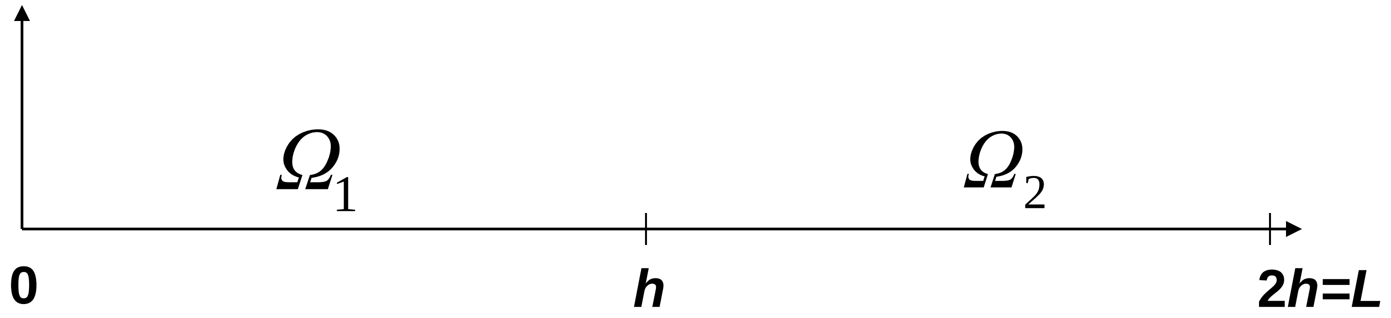
and $F_i = \int_0^L p_0 \varphi_i \,dx$

We can already see K_{ij} will be symmetric since we can

interchange i, j without effect.

Discretization and Basis Functions

Now we discretize our domain. For this example, we will discretize x=[0, L] into 2 “elements”.

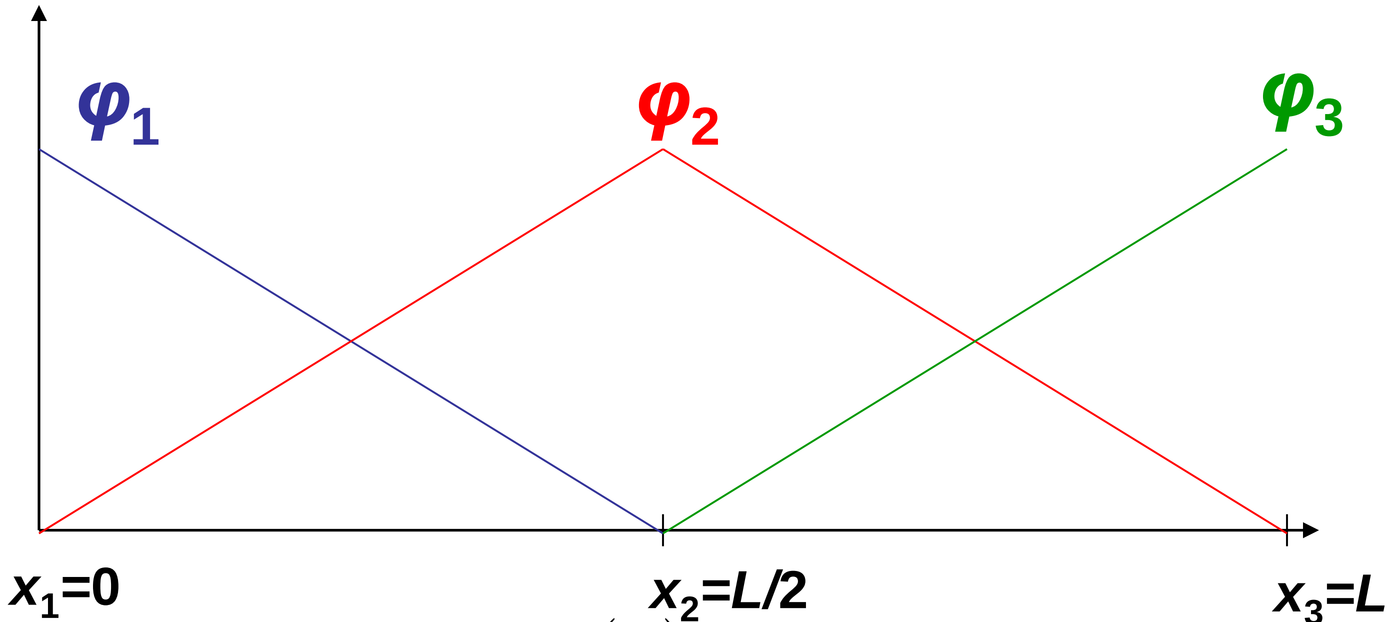


In 1-D, elements are segments. In 2-D, they are triangles, tetrads, etc. In 3-D, they are solids, such as tetrahedra. We will solve the Galerkin problem on each element.

Discretization and Basis Functions

For a set of basis functions, we can choose anything. For simplicity here, we choose piecewise linear "hat functions".

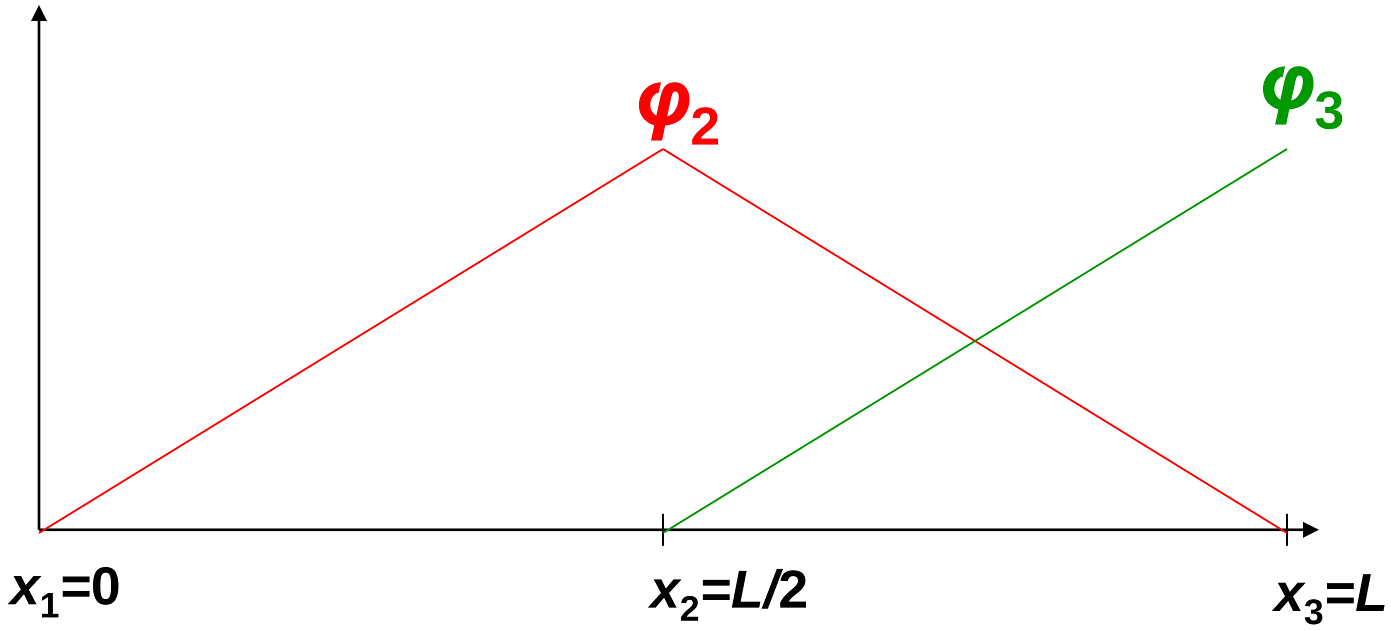
Our solution $u_h(x)$ will be a linear combination of these functions.



Basis functions satisfy: $\varphi_i\left(x_j\right)=\delta_j^i \Rightarrow$ Our solution will be interpolatory. Also, they satisfy the partition of unity.

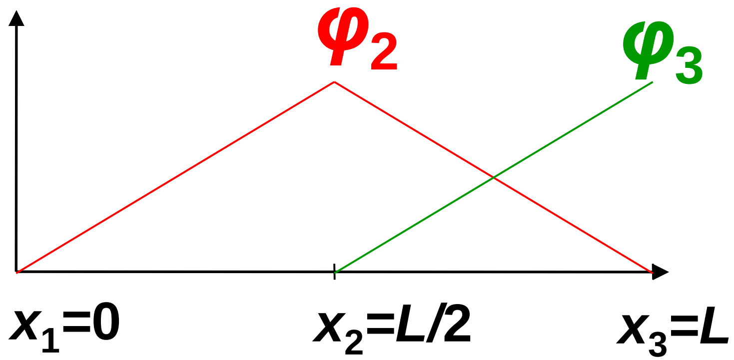
Discretization and Basis Functions

We can throw out φ_1 a priori because, since in this example $u(0)=0$, we know that the coefficient c_1 must be 0.



Basis Functions

$$\varphi_2 = \begin{cases} \dfrac{2x}{L} & \text{if } x \in \left[0, \frac{L}{2}\right] \\ 2 - \dfrac{2x}{L} & \text{if } x \in \left[\frac{L}{2}, L\right] \\ 0 & \text{otherwise} \end{cases}$$



$$\varphi_3 = \begin{cases} \dfrac{2x}{L} - 1 & \text{if } x \in \left[\frac{L}{2}, L\right] \\ 0 & \text{otherwise} \end{cases}$$

Matrix Formulation

Given our matrix problem $\mathbf{Kc} = \mathbf{F}$:

$$\underbrace{\sum_{j=1}^{N} c_j}_{\mathbf{c}} \underbrace{\int_0^L \frac{d\phi_j}{dx}\frac{d\phi_i}{dx}dx}_{\mathbf{K}} = \underbrace{\int_0^L p_0 \phi_i dx}_{\mathbf{F}} \Rightarrow \mathbf{Kc} = \mathbf{F}$$

We can insert the ϕ_i chosen on the previous slide and arrive at a linear algebra problem. Differentiating the basis functions, then evaluating the integrals, we have:

$$\mathbf{K} = \frac{1}{L}\begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix}, \quad \mathbf{F} = \frac{p_0}{L}\begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$$

Differentiating the basis functions can be done in advance, (analytically) since the basis functions are known, and integration would be performed numerically by quadrature.

(by Gaussian quadrature, since it is exact for polynomials.)

Notice $\mathbf{K}$ is symmetric as expected.

Solution

Solving the Gaussian elimination problem on the previous slide, we obtain our coefficients c_i:

$$\mathbf{c} = \begin{bmatrix} \frac{3 p_0 L^2}{8} \\ \frac{p_0 L^2}{2} \end{bmatrix},$$ which when multiplied by basis functions ϕ_i gives

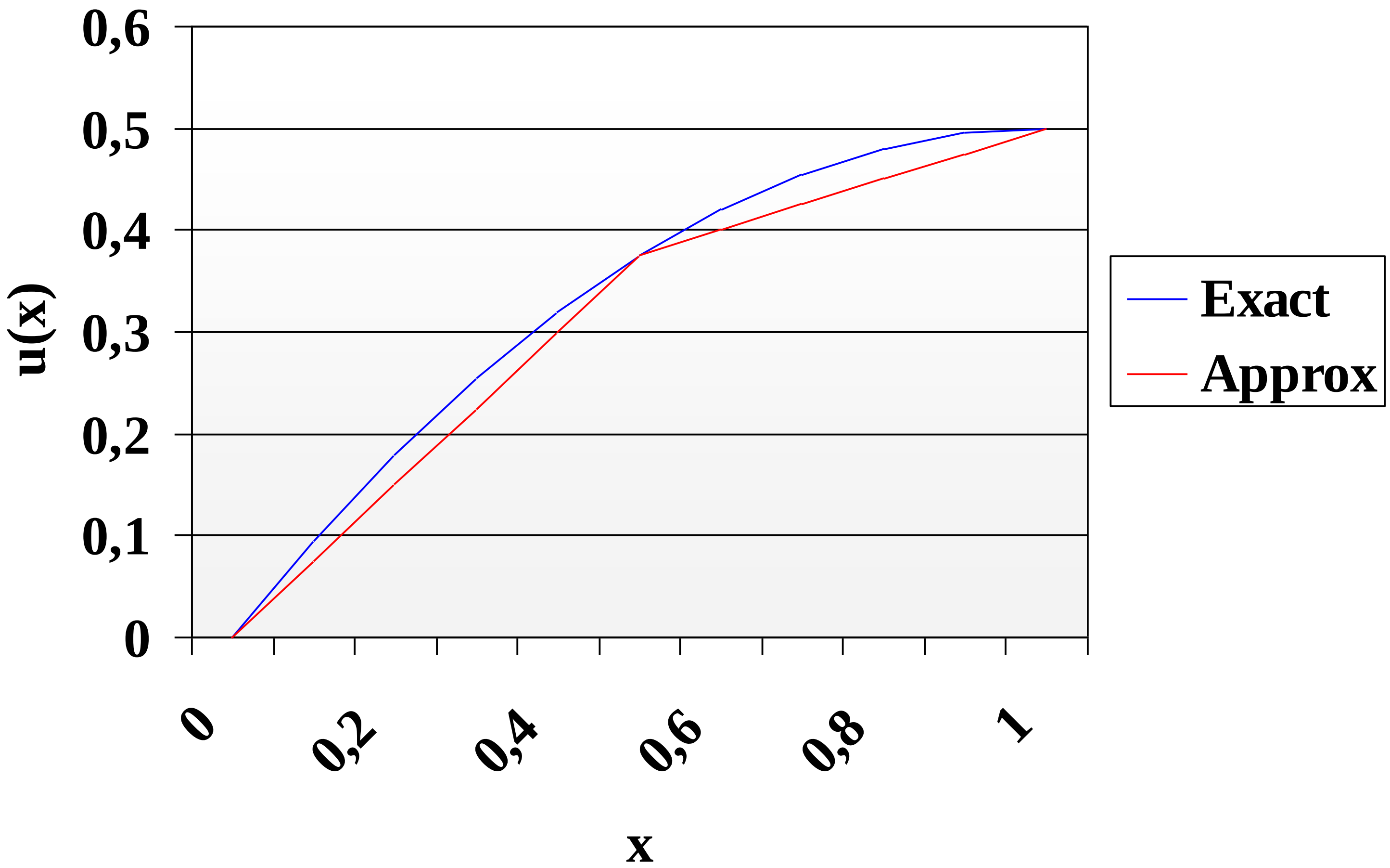
our final numerical solution:

$$u_h(x) = \begin{cases} \frac{3}{4} p_0 L x & \text{when } x \in \left[0, \frac{L}{2}\right] \\ \frac{1}{4} p_0 (L^2 + Lx) & \text{when } x \in \left[\frac{L}{2}, L\right] \end{cases}$$

The exact analytical solution for this problem is:

$$u(x) = p_0 L x - \frac{p_0 x^2}{2}$$

Solution



Notice that the numerical solution is “interpolatory”, or nodally exact.

Example Problem 2: 1D Diffusion-reaction problem

Strong Formulation

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

Weak Formulation

$$\text{Find} \quad u \in H_0^1(0,1):$$

$$\int_0^1 u'v'dx + \int_0^1 \sigma uv dx = \int_0^1 fv dx \qquad \forall v \in V \text{ with } V = H_0^1(0,1)$$

Galerkin - FEM

$$\text{Find} \quad u_h \in V_h:$$

(*) $$\int_0^1 u_h' v_h' dx + \int_0^1 \sigma\, u_h v_h dx = \int_0^1 f v_h dx \qquad \forall v_h \in V_h$$

Example Problem 2: 1D Diffusion-reaction problem

Partition $\mathcal{T}_h$ of $\Omega=[0,1]$ into subintervals K_j

Space V_h: choice of linear elements

$$V_h = \left\{ v_h \in X_h^1 : \; v_h(0) = v_h(1) = 0 \right\}$$

Bases Functions $\left\{ \varphi_i(x) \right\}_{i=1}^{N}$ for X_h^1 is the set of hat functions, one for each internal node (internal vertices of T_h = degrees of freedom N_h)

$$u_h(x) = \sum_{i=1}^{N} \xi_i \varphi_i(x)$$

$$\forall v_h \in V_h$$

Imposing that the Galerkin-FEM (*) is satisfied for every bases functions, we get a linear system

Example Problem 2: 1D Diffusion-reaction problem

We get a tridiagonal linear system of N eqs. and N unknowns

$$Au = f$$

$$A = \left[a_{ij} \right] \qquad a_{ij} = \int_0^1 \varphi_j' \varphi_i' dx + \int_0^1 \sigma \varphi_j \varphi_i \, dx$$

Stiffness matrix

$$u = \left[u_i \right]$$

$$f = \left[f_i \right] \qquad f_i = \int_0^1 f \varphi_i dx$$

Remark

$$\boxed{u_i = u_h\left(x_i\right) \qquad 1 \le i \le N}$$

$$a_{ij} = \int_0^1 \left(\varphi_j' \varphi_i' + \sigma \varphi_j \varphi_i\right) dx = \int_{x_{i-1}}^{x_i} \left(\varphi_j' \varphi_i' + \sigma \varphi_j \varphi_i\right) dx + \int_{x_i}^{x_{i+1}} \left(\varphi_j' \varphi_i' + \sigma \varphi_j \varphi_i\right) dx$$

A priori estimate of the error

Theorem Let $u \in V$ be the exact solutionof the problem:

$$\boxed{\begin{array}{ll} \text{Find} \quad u \in V: & \\ a(u,v) = F(v) & \forall v \in V \end{array}}$$

and u_h **the approximate solution** by FEM of degree **r**, that is the solution of the Galerkin-FEM problem

$$\text{Find} \quad u_h \in V_h: \qquad V_h = \{ v_h \in X_h^r : \quad v_h(0) = v_h(1) = 0 \}$$

$$\int_0^1 u_h' v_h' dx + \int_0^1 \sigma\, u_h v_h dx = \int_0^1 f v_h dx \qquad \forall v_h \in V_h$$

Let $u \in H^{p+1}(I)$ for a given p such that $r \le p$. Then the following holds **(a priori estimate of the error):**

$$\|u - u_h\|_V \le \frac{M}{\alpha} C h^r |u|_{H^{r+1}(I)}$$ **Energy norm**

where C is a constant which does not depend on h.

Improve the accuracy

You can follow two different strategies:

1. Decrease h, i.e. refine the grid;
2. Increasing r, i.e. use of higher degree finite element.

The second way is meaningful only if the solution is sufficiently regular.

$$\boldsymbol{u} \in \boldsymbol{V} \cap \boldsymbol{H}^{p+1}(\boldsymbol{I}) \Longrightarrow$$ Max r value is r = p

r	$u \in \mathrm{H}^1(\Omega)$	$u \in \mathrm{H}^2(\Omega)$	$u \in \mathrm{H}^3(\Omega)$	$u \in \mathrm{H}^4(\Omega)$	$u \in \mathrm{H}^5(\Omega)$
1	*converge*	$\boxed{h^1}$	h^1	h^1	h^1
2	*converge*	h^1	$\boxed{h^2}$	h^2	h^2
3	*converge*	h^1	h^2	$\boxed{h^3}$	h^3
4	*converge*	h^1	h^2	h^3	$\boxed{h^4}$

Order of convergence with respect to h for the FEM to vary the regularity of the solution.

It converges but the estimate does not hold

The finite element method in the multidimensional case

2D Finite Element Method

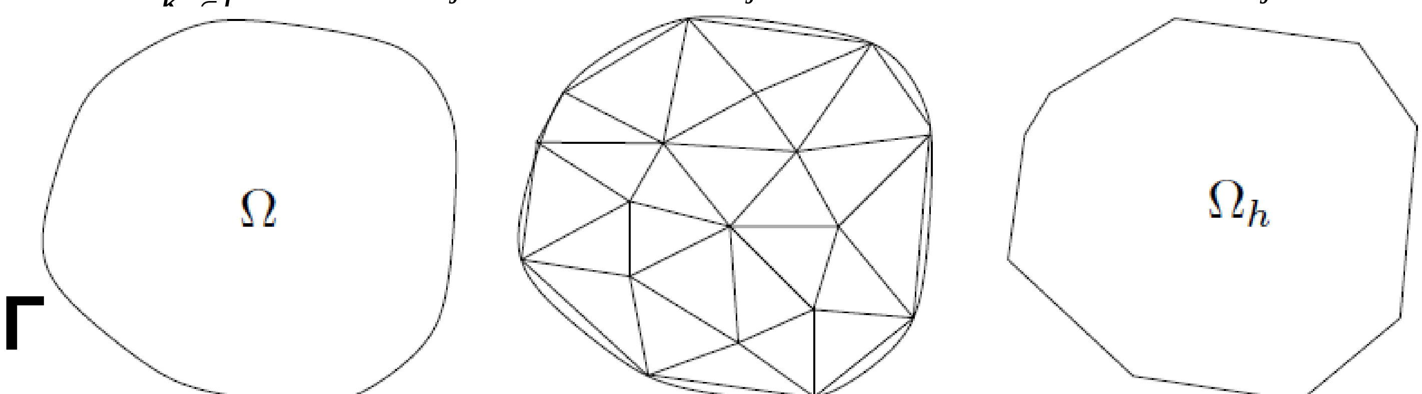
Discretization of the domain $\Omega \subset \Re^2$

Approximate Γ with a polygonal curve

Triangulation of Ω into a set T_h of non-overlapping triangles Ki such that no vertex of one triangle lies on the edge of another triangle,

$$\Omega_h = \text{int}\left(\bigcup_{K_j \in \mathrm{T}_h} K_j \right)$$

$$h = \max_{K_j \in T_h} diam(K_j), \quad diam(K_j) = \text{longest edge of } K_j$$



2D Finite Element Method

P_r polynomial space of degree r

$$\mathbf{P}_1 = \left\{ f\left(x_1, x_2\right) = a + bx_1 + cx_2, \quad \text{con } a,b,c \in \Re \right\} \quad \dim(\mathbf{P}_1) = 3$$

$$\mathbf{P}_2 = \left\{ f\left(x_1, x_2\right) = a + bx_1 + cx_2 + dx_1x_2 + ex_1^2 + gx_2^2, \text{ with } a,b,c,d,e,g \in \mathbb{R} \right\} \dim(\mathbf{P}_2) = 6$$

....

$$\mathbf{P}_r = \left\{ f\left(x_1, x_2\right) = \sum_{i+j\le r} a_{ij} x_1^i x_2^j, \quad \text{with } a_{ij} \in \Re \right\} \quad \dim(\mathbf{P}_r) = \frac{(r+1)(r+2)}{2}$$

Define a finite-dimensional subspace (FEM Space)

$$X_h^r = \left\{ v_h \in C^0\left(\overline{\Omega}\right) : v_h|_K \in \mathbf{P}_r, \quad \forall K \in \mathrm{T}_h \right\} \quad r = 1,2,\ldots$$

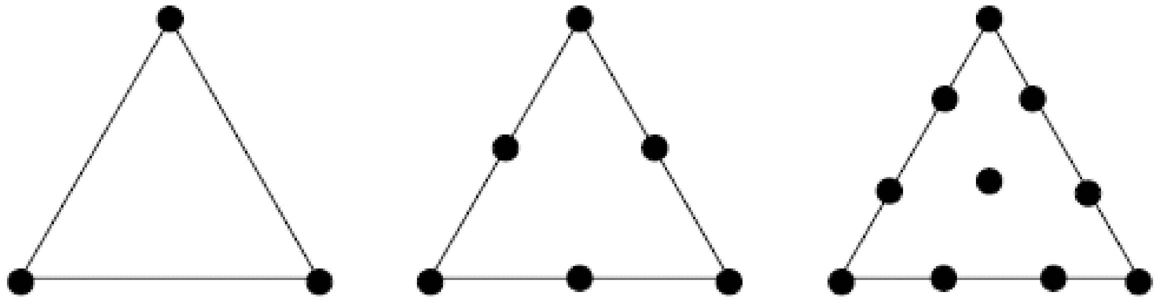
Space of piecewise polynomial functions of degree r, where $v|_K$ denotes the restriction of v to the element (triangle) K (i.e. the function defined on K) of the mesh T_h

2D Finite Element Method

$$\dim \mathrm{P}_r = \frac{(r+1)(r+2)}{2}$$

dim P_1=3 dim P_2=6 dim P_3=10

On every single element of the triangulation T_h, the generic function v_h is well defined if you know its value respectively in 3, 6 and 10 nodes suitably chosen.



Nodes for linear poly (r=1) quadratic (r=2) cubic (r=3)

2D Finite Element Method

We introduce the space generator of the finite elements:

$$X_h^r = \left\{ v_h \in C^0\left(\overline{\Omega}\right) : \; v_h|_K \in \mathbb{P}_r, \quad \forall K \in \mathcal{T}_h \right\} \qquad r = 1,2,\ldots$$

$$\overset{\circ}{X}{}_h^r = \left\{ v_h \in X_h^r : \quad v_h|_{\partial\Omega} = 0 \right\}$$

X_h^r and $\overset{\circ}{X}{}_h^r$

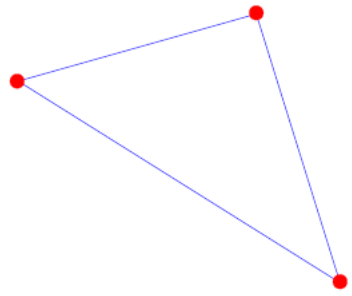
are good approximations of the spaces

$H^1(\Omega)$ and $H_0^1(\Omega)$

Linear Finite Elements

Let r = 1. Choose as degrees of freedom for describing the functions of X_h^1 the values at the vertices of the elements of T_h

It is quite intuitive that a function $v_h \in X_h^1$ is completely determined by the values it takes at the vertices of the elements of the triangulation



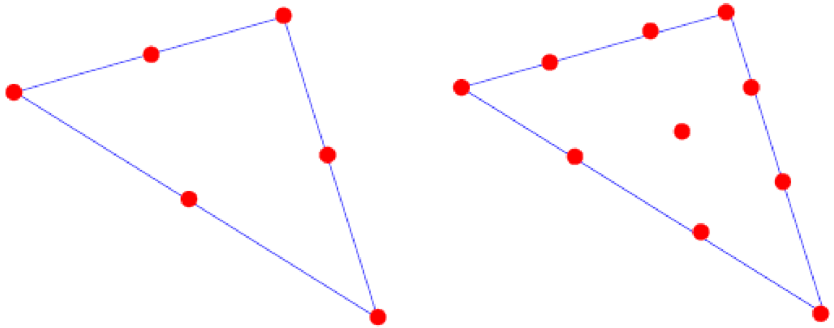
Quadratic and cubic Finite Elements

- Let r = 2. Choose as degrees of freedom for describing the functions of X_h^2

  the values at the vertices of the elements of T_h

  and the midpoints of the edges of the elements of T_h

- For X_h^3 we have 10 degrees of freedom for every element (as in the figure) and similarly for elements of higher degree.



2D FEM Poisson equation

Strong Formulation Find u

$$\begin{cases} -\Delta u = f & in \quad \Omega \\ u = 0 & \text{on} \quad \Gamma \end{cases} \qquad \Omega \subset \Re^2$$

Galerkin - FEM

$$\text{Find} \quad u_h \in V_h = \overset{\circ}{X}{}_h^r$$

$$\int_\Omega \nabla u_h \bullet \nabla v_h \; d\Omega = \int_\Omega f \; v_h \; d\Omega \qquad \forall v_h \in V_h, \quad V_h = \overset{\circ}{X}{}_h^r$$

Each function $\boldsymbol{v} \in \boldsymbol{V_h}$ is uniquely characterized by the values that it assumes in the **degrees of freedom** Ni, with i = 1,…,N_h, of the triangulation T_h (excluding the degrees of freedom of the **boundary** where $v_h = 0$)**.**

2D FEM Poisson equation

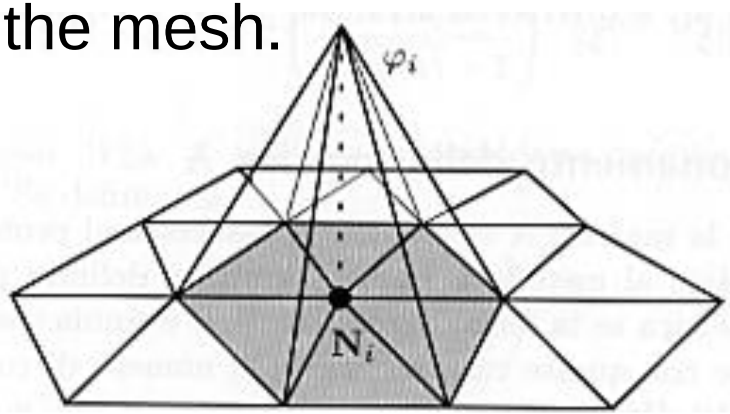
Basis of the Space

$$\varphi_j \in V_h, \quad j = 1,2,\ldots,N_h$$

$$\varphi_j(N_i) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \qquad i, j = 1,2,\ldots,N_h$$

If r = 1, the degrees of freedom are the vertices of the elements, not on the boundaries of Ω , while φ_j is linear on each element (triangle) and and takes the value 1 at node Nj and 0 at all other nodes of the mesh.

N_h internal nodes of T_h



2D FEM Poisson equation

Therefore $$v_h(x) = \sum_{i=1}^{N_h} \eta_i \phi_i(x) \qquad x \in \Omega \quad \text{with} \quad \eta_i = v_h(N_i)$$

Express $$u_h(x) = \sum_{i=1}^{N_h} \xi_i \phi_i(x)$$

and requiring that it satisfies the equation $\int_\Omega \nabla u_h \bullet \nabla v_h \;\; d\Omega = \int_\Omega f \;\, v_h \, d\Omega$
for each function of the base,

$$\sum_{j=1}^{N_h} \eta_j \sum_{i=1}^{N_h} \xi_i \int_\Omega \nabla \varphi_i \bullet \nabla \varphi_j \;\; d\Omega = \sum_{j=1}^{N_h} \eta_j \int_\Omega f \;\; \varphi_j \; d\Omega \qquad j = 1,2,,\ldots,N_h$$

we get a linear system of N_h equations in N_h unknowns ξ_i

2D FEM Poisson equation

$$\sum_{i=1}^{N_h} \xi_i \int_\Omega \nabla \phi_i \bullet \nabla \phi_j d\Omega = \int_\Omega f \phi_j \, d\Omega \qquad j = 1,2,,\ldots,N_h$$

$$A = \left[a_{ij} \right] \qquad a_{ij} = \int_\Omega \nabla \varphi_i \bullet \nabla \varphi_j \; d\Omega$$

Stiffness matrix N_h x N_h

$$\boldsymbol{\xi} = \left[\xi_i \right] \quad \text{with} \quad \xi_i = u_h \left(N_i \right)$$

$$b = \left[b_i \right] \quad \text{with} \quad b_i = \int_\Omega f \varphi_i d\Omega$$

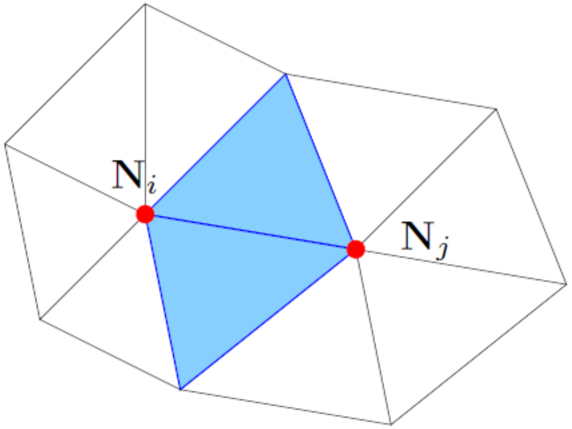
$$\boxed{\mathbf{A} \;\; \boldsymbol{\xi} = \mathbf{b}}$$

Symmetric, def pos., sparse

2D FEM Poisson equation

$$A = \left[a_{ij} \right] \qquad a\left(\varphi_i, \varphi_j\right) = a_{ij} = \sum_{K \in T_h} \int_K \nabla \varphi_i \bullet \nabla \varphi_j \; d\Omega$$

Since the support of the generic function φ_i of the base is formed by the triangles having in common only the node $\mathbf{N_i}$, A is a sparse matrix. In particular $\boldsymbol{a_{ij}}$ is nonzero if and only if $\mathbf{N_i}$ and $\mathbf{N_j}$ are nodes of the same triangle.



2D FEM Poisson equation

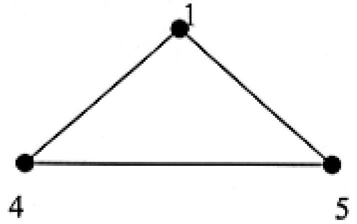
Local Matrix of stiffness related to the element K

$$\begin{bmatrix} a_K(\varphi_i,\varphi_i) & a_K(\varphi_i,\varphi_j) & a_K(\varphi_i,\varphi_k) \\ a_K(\varphi_j,\varphi_i) & a_K(\varphi_j,\varphi_j) & a_K(\varphi_j,\varphi_k) \\ a_K(\varphi_k,\varphi_i) & a_K(\varphi_k,\varphi_j) & a_K(\varphi_k,\varphi_k) \end{bmatrix}$$

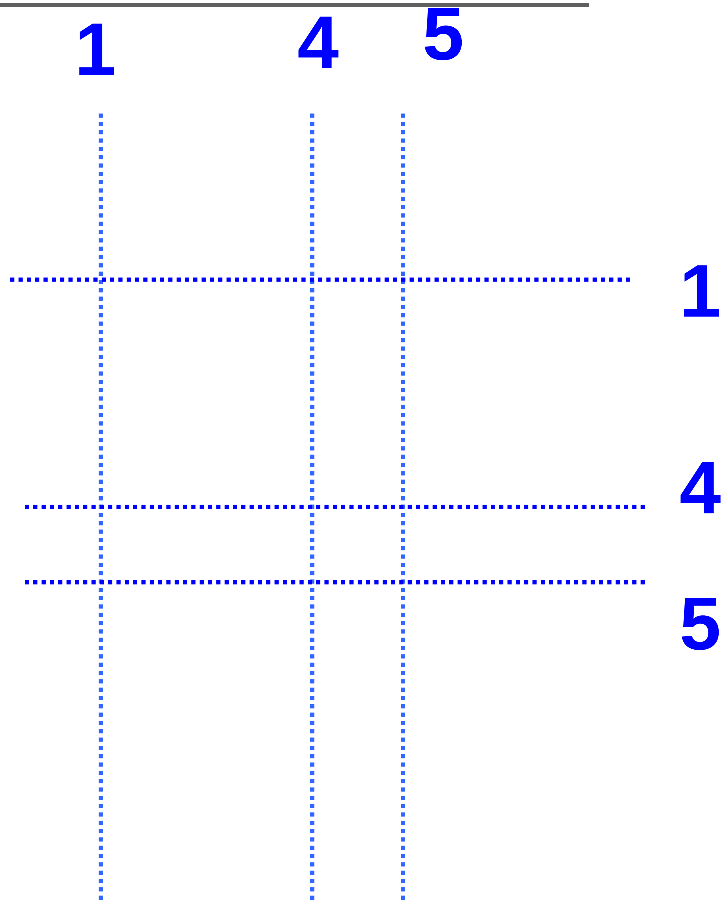
Assembly: the construction of the global stiffness matrix A using the matrices related to each element $K \in \mathrm{T}_h$

2D FEM Poisson equation

$$A^e = \begin{bmatrix} 7 & 3 & 1 \\ 3 & 6 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$



$$A = \begin{bmatrix} 7 & 0 & 0 & 3 & 1 & . & . & . \\ 0 & 0 & 0 & 0 & 0 & . & . & . \\ 0 & 0 & 0 & 0 & 0 & . & . & . \\ 3 & 0 & 0 & 6 & 2 & . & . & . \\ 1 & 0 & 0 & 2 & 5 & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \end{bmatrix}$$



Conditioning of the matrix A

the matrix $A = \left[a_{ij} \right]$ $\quad a\left(\varphi_i, \varphi_j\right) = a_{ij} = \sum_{K \in T_h} \int_K \nabla \varphi_i \bullet \nabla \varphi_j \; d\Omega$

is positive definite; A also appears to be symmetric if the bilinear form a (.,.) is symmetric.

Its condition number is given by

$$K_2(A) = \lambda_{max}(A)/\lambda_{min}(A);$$

where $\lambda_{max}(A)$ and $\lambda_{min}(A)$, are the maximum and minimum, eigenvalues of A, respectively.

It can be shown that $K(A) = C\ h^{-2}$

where C is a constant independent of the mesh parameter h, but dependent on the degree of the finite elements used.

The matrix is therefore ill-conditioned to the decrease of h.

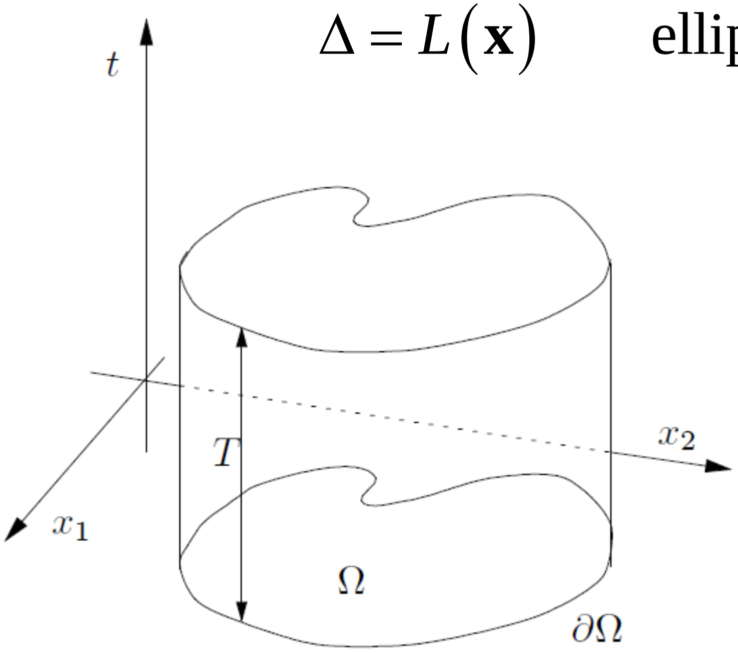
FEM –Parabolic equation 2D

$$\frac{\partial u}{\partial t} + \Delta u = f \qquad x \in \Omega, \quad t > 0$$

Ω domain on $\Re^2$

$f = f(\mathbf{x}, t)$ given function

$\Delta = L(\mathbf{x})$ elliptic operator on the unknown $\mathrm{u} = u(\mathbf{x}, t)$



Find the solution in a finite interval

$$Q_T = \Omega \times (0, T)$$

Cylinder in $\Re^2 \times \Re^+$

FEM –Parabolic equation 2D

The problem is well-posed if appropriate IC and BC are given:

$$IC \qquad u(\mathbf{x},0)=u_0(\mathbf{x}) \qquad \forall \mathbf{x}\in\Omega$$

$$BC \qquad \begin{cases} u(\mathbf{x},t)=\phi(\mathbf{x},t) & \forall \mathbf{x}\in\Gamma_D \quad \text{e} \quad \forall t>0 \\ \dfrac{\partial u(\mathbf{x},t)}{\partial n}=\psi(\mathbf{x},t) & \forall \mathbf{x}\in\Gamma_N \quad \text{e} \quad \forall t>0 \end{cases}$$

$u_0,\ \phi,\ \psi$ given functions

$\Gamma_D \cup \Gamma_N = \partial\Omega \quad \Gamma_D \cap \Gamma_N = \varphi$

Γ_D Dirichelet Boundary

Γ_N Neumann boundary

FEM –Parabolic equation 1D

1D case:
$$\begin{cases} \dfrac{\partial u}{\partial t} - \nu \dfrac{\partial^2 u}{\partial x^2} = f & 0 < x < L, \quad t > 0 \\ u(x,0) = u_0(x), & 0 < x < L \\ u(0,t) = u(L,t), & t > 0 \end{cases}$$

Heat Eq.

The PDE model represents the **temperature *u(x,t)*** at the point x and time t of a metal bar of length ***[0,L]***, with

- ν **thermal conductivity**,
- extremes are kept at a constant zero degree temperatures,
- the function u_0 is the **initial temperature**,
- f represents the caloric **production** (per unit length) supplied to the bar.

FEM – weak formulation

Multiply the PDE $\forall t > 0$ for a test function *v=v(x)* and then integrate on Ω.

Let $V = H^1_{\Gamma_D}(\Omega)$ *and* $\forall t > 0$

$$\text{Find } u(t) \in V$$

$$\int_\Omega \frac{\partial u(t)}{\partial t} v \, d\Omega + a(u(t), v) = F(v) \qquad \forall v \in V$$

u(0)=u$_0$, where

- ***a(.,.)*** and ***F(.)*** are the bilinear form and the functional associated to the elliptic operator L and the rhs F, respectively,
- we assume $\boldsymbol{\varphi=0}$ (Dirichlet) and $\boldsymbol{\psi=0}$ (Neumann).

Galerkin's Method

(*)

$$\forall t > 0, \quad \text{Find} \quad u_h(t) \in V_h$$

$$\int_{\Omega} \frac{\partial u_h(t)}{\partial t} v_h \, d\Omega + a\left(u_h(t), v_h\right) = F(v_h) \qquad \forall v_h \in V_h$$

with $u_h(0) = u_{0h}$,

$V_h \subset V$ is a finite dimensional space,

u_{0h} is an approximation of u_0 in V_h.

Semi-discretization: discretization with respect to the space variables, but not wrt the differential operators in time.

Galerkin-FEM

- Introduce the bases $\{\varphi_j\}$ for V_h.
- Requiring that the form (*) holds for every bases function, then (*) holds also for any linear combination of them, that is for any function of V_h.
- For each *t>0,* also the solution is represented as a linear combinationn of the basis functions

$$u_h(\mathbf{x},t)=\sum_{j=1}^{N_h} u_j(t)\varphi_j(\mathbf{x})$$

Problem unknowns

Galerkin-FEM

$$\int_{\Omega} \sum_{j=1}^{N_h} \dot{u}_j(t)\varphi_j\varphi_i \, d\Omega + a\left(\sum_{j=1}^{N_h} u_j(t)\varphi_j, \varphi_i\right) = F(\varphi_i) \qquad i = 1, 2, ..., N_h$$

$$\sum_{j=1}^{N_h} \dot{u}_j(t) \underbrace{\int_{\Omega} \varphi_j\varphi_i \, d\Omega}_{m_{ij}} + \sum_{j=1}^{N_h} u_j(t) \underbrace{a(\varphi_j, \varphi_i)}_{a_{ij}} = \underbrace{F(\varphi_i)}_{f_i(t)} \qquad i = 1,2,..., N_h$$

Mass Matrix **Stiffness Matrix**

$\dot{u}_j(t)$ first derivativela of $u_j(t)$ wrt time

Galerkin-FEM

$\mathrm{u} = \left[u_1(t), u_2(t), \ldots, u_{N_h}(t)\right]^T$ Unknown Vector

$M = \left[m_{ij}\right]$ Mass Matrix

$A = \left[a_{ij}\right]$ Stiffness Matrix

$\mathrm{f} = \left[f_1(t), f_2(t), \ldots, f_{N_h}(t)\right]^T$ RHS

$$\boxed{M\dot{u}(t) + A\,u(t) = f(t)}$$

M is invertible (because it is positive definite)

$$\dot{u}(t) = -M^{-1}A\,u(t) + M^{-1}f(t)$$

System of ODEs
first order

θ-method

Discretize the time derivative using finite difference formula and replace the other terms with a linear combination of the values at the time step $\mathbf{t^k}$ and $\mathbf{t^{k+1}}$, with parameter θ $\left(0 \le \vartheta \le 1\right)$

$$M\frac{u^{k+1}-u^{k}}{\Delta t}+A\left[\vartheta u^{k+1}+\left(1-\vartheta\right)u^{k}\right]=\vartheta f^{k+1}+\left(1-\vartheta\right)f^{k}$$

$\Delta t = t^{k+1} - t^{k}$ $\quad k = 0,1,....$ Time Step (here assumed to be constant)

Galerkin-FEM

$$\boldsymbol{M}\,\frac{\boldsymbol{u}^{k+1}-\boldsymbol{u}^{k}}{\Delta t}+\boldsymbol{A}\left[\vartheta\boldsymbol{u}^{k+1}+(1-\vartheta)\boldsymbol{u}^{k}\right]=\vartheta\boldsymbol{f}^{\,k+1}+(1-\vartheta)\,\boldsymbol{f}^{\,k}\qquad(0\le\vartheta\le 1)$$

$\vartheta = 0$ Forward Euler Method

$$M\,\frac{u^{k+1}-u^{k}}{\Delta t}+Au^{k}=f^{\,k}$$

LTE accurate
first order wrt ***Δt***

$\vartheta = 1$ Backward Euler Method

$$M\,\frac{u^{k+1}-u^{k}}{\Delta t}+Au^{k+1}=f^{\,k+1}$$

LTE accurate
first order wrt ***Δt***

$\vartheta = 1/2$ Crank-Nicolson Method

$$M\,\frac{u^{k+1}-u^{k}}{\Delta t}+\frac{1}{2}A\left(u^{k+1}+u^{k}\right)=\frac{1}{2}\left(f^{\,k+1}+f^{\,k}\right)$$

LTE accurate
second order wrt ***Δt***

Consider the two Euler Methods

$\vartheta = 0$ Forward Euler Method

$$M \frac{u^{k+1} - u^k}{\Delta t} + Au^k = f^k$$

$\vartheta = 1$ Backward Euler Method

$$M \frac{u^{k+1} - u^k}{\Delta t} + Au^{k+1} = f^{k+1}$$

System of Linear equations with coefficient matrix K:

$$K = \frac{M}{\Delta t} \quad \vartheta = 0 \qquad K = \frac{M}{\Delta t} + A \quad \vartheta = 1$$

The explicit scheme is conditional stable:

$$\boldsymbol{\Delta t \leq ch^2, \quad c > 0}$$

lumping of the mass matrix:

Reduce matrix M to be diagonal, the eqs. become uncoupled

Case $\left(0 < \vartheta < 1\right)$

$$M\frac{u^{k+1}-u^{k}}{\Delta t}+A\left[\vartheta u^{k+1}+\left(1-\vartheta\right)u^{k}\right]=\vartheta f^{k+1}+\left(1-\vartheta\right)f^{k}$$

Coefficient Matrix

$$\boxed{K=\frac{M}{\Delta t}+\vartheta A \qquad \vartheta>0}$$

It’s time invariant; if the spatial mesh does not change, it can be factorized once at the beginning of the process.

Given that M is symmetric, if A is symmetric then also K will be symmetric, so that you can use, e.g., Cholesky factorization: K = HH^T . At each time step, the system is solved by two triangular systems of N_h unknowns, with a cost $N_h^2/2$

$$\begin{cases} Hy = g \\ H^T u^{k+1} = y \end{cases}$$

Galerkin-FEM

- **A priori estimates**
- **Convergence**
- **Stability**

If $\vartheta \geq 1/2$ the θ-method is unconditionally stable, that is it is stable for every Δt.
(In particular, the implicit methods)

If $\vartheta < 1/2$ the θ-method is conditionally stable for

$$\Delta t \leq \frac{2}{(1-2\vartheta)\lambda_{N_h}}$$

Max eigenvalue of the bilinear form a

Diffusion-transport-reaction

$$\underbrace{\nabla\cdot\left(-c\nabla u\right)}_{\text{diffusive term}}+\underbrace{\beta\cdot\nabla u}_{\text{advection (transport)}}+\underbrace{\alpha u}_{\text{reaction } term}=f \quad in \quad \Omega$$

$$u=0 \qquad on \quad \partial\Omega$$

Weak Form:

$$\text{Find} \quad u\in V\equiv H_0^1(\Omega): \quad a\left(u,v\right)=\left(f,v\right) \qquad \forall v\in V$$

$$a\left(u,v\right)=\int_\Omega c\nabla u\cdot\nabla v d\Omega+\int_\Omega v\beta\cdot\nabla u d\Omega+\int_\Omega \alpha u v d\Omega$$

$$\left(f,v\right)=\int_\Omega f v d\Omega$$

The solution exists and is unique (Lax-Milgram)

Galerkin-FEM

$$\text{Find} \quad u_h \in V_h : \quad a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h$$

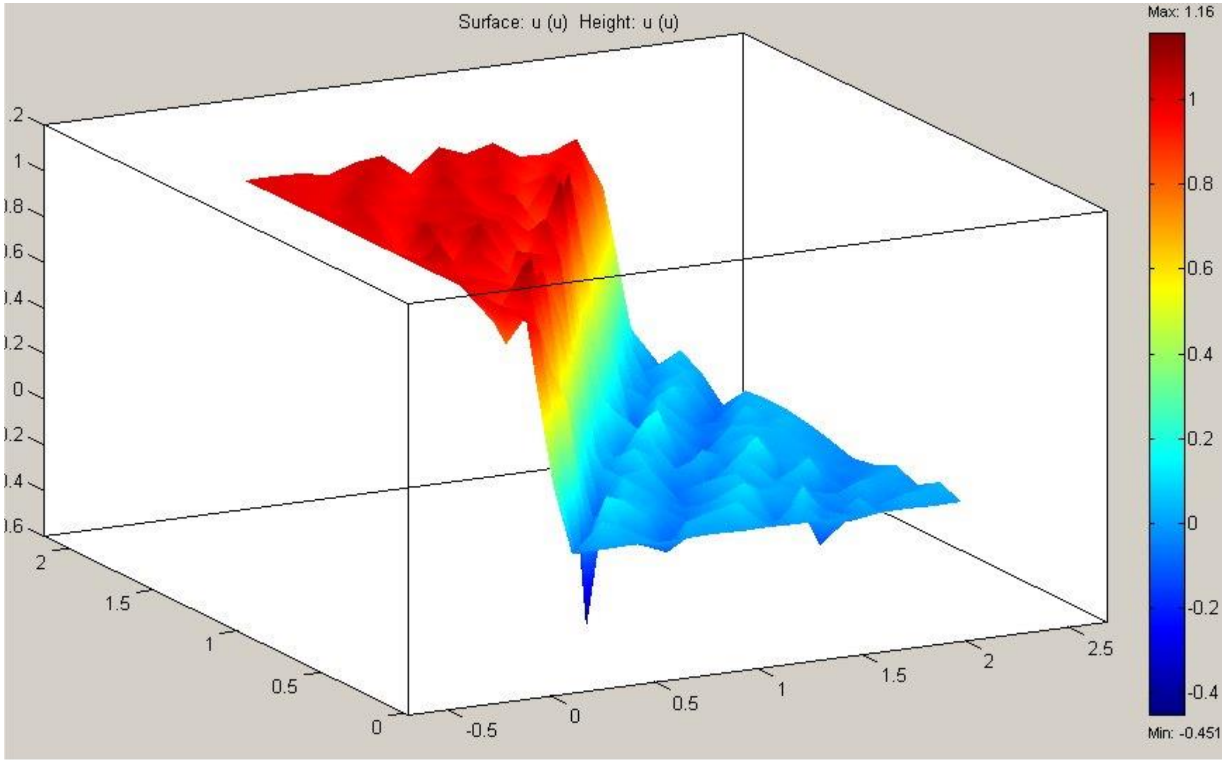
where $\{V_h, h>0\}$ is a family of finite dimensional spaces.

Error estimate:

$$\|u\text{-}u_h\|_V \le \frac{M}{C_\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V$$

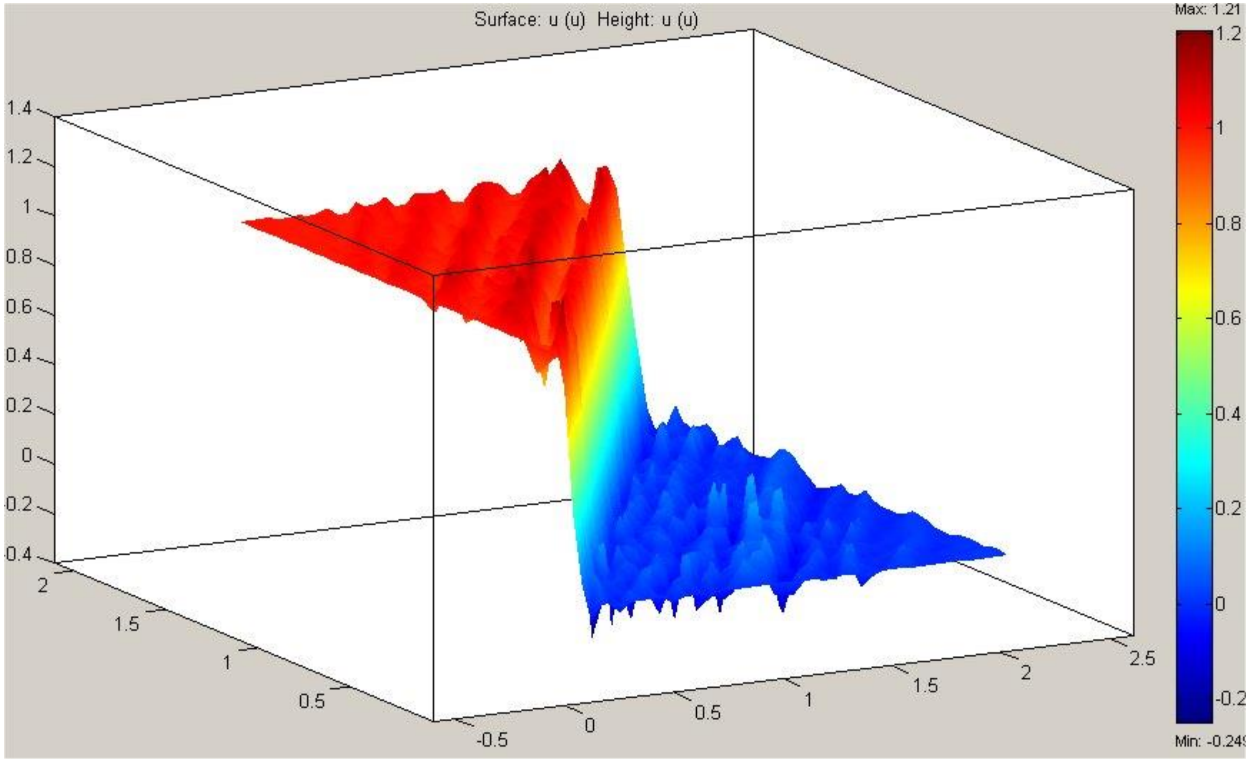
The value M/C_α increases with the ratio β/c then in problems in which the convective term dominates the diffusive the Galerkin method can give rise to inaccurate solutions.

Example



Lagrangian **linear element**: 714 elements, 387 nodes

Example



Lagrangian **quadratic elements**: 714 elements, 387 nodes

Methods of Stabilization (streamline diffusion)

Add a term of artificial diffusion in the direction of the beta field:

$$- hQ\,\mathrm{div}\left(\frac{\partial u}{\partial \beta}\beta\right) \qquad \text{where} \quad Q = \frac{1}{\|\beta\|}$$

In the Galerkin problem you have the additional term:

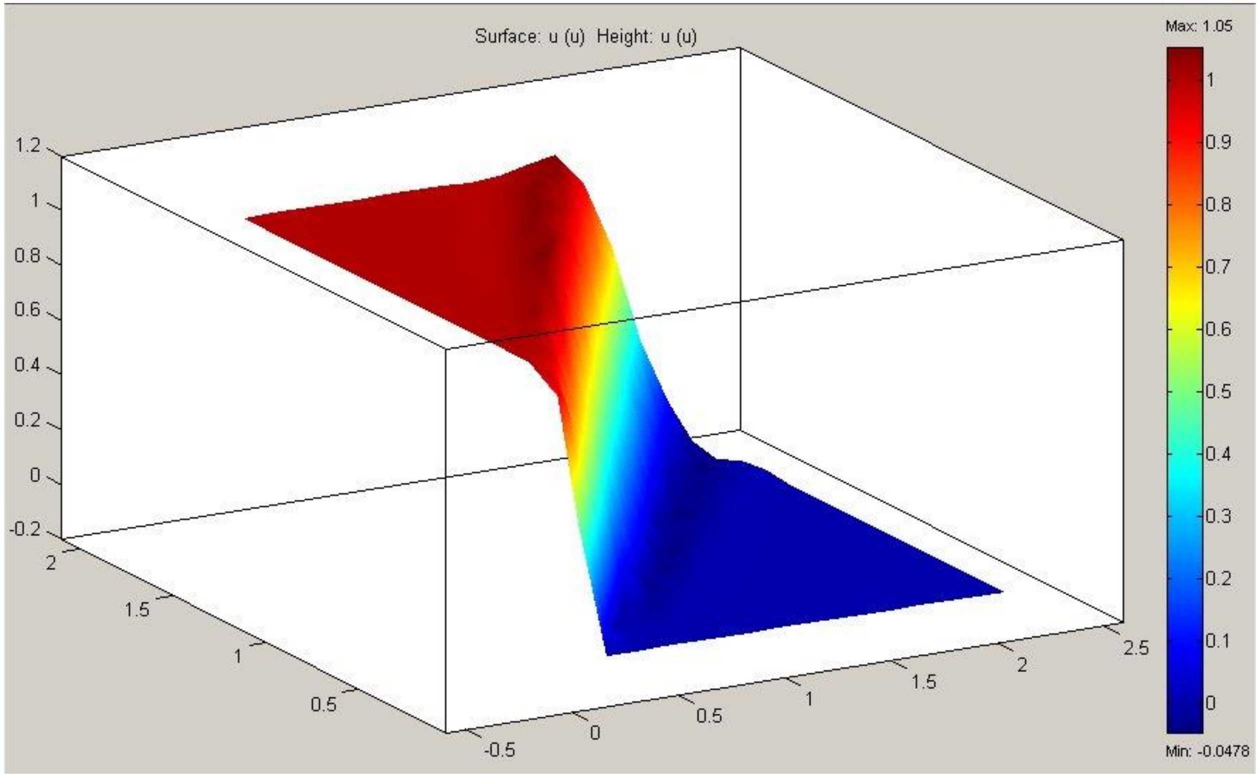
$$b_h(u_h,v_h) = hQ\int_\Omega \big((\beta\cdot\nabla u_h) + (\beta\cdot\nabla v_h)\big)\,\mathrm{d}\Omega$$

The discrete problem becomes:

$$\boxed{\text{Find} \quad u_h \in V_h : \quad a_h(u_h,v_h) = (f,v_h) \qquad \forall v_h \in V_h}$$

$$\text{where } a_h(u_h,v_h) = a(u_h,v_h) + b_h(u_h,v_h)$$

Example



Lagrangian **linear element**: 714 elements, 387 nodes
Streamline Diffusion: fact. 1.0

Bibliografia

1. A.Quarteroni: Modellistica Numerica per Problemi Differenziali, Springer Italia 2006.

2. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method (Cap. 1), Cambridge University Press, Cambridge (1987).

3. T. J. R. Hughes, The Finite Element Method. Linear Static and Dynamic Finite Element Analysis (Cap. 3), Dover Publishers, New York (2000).

4. P. J. Frey e P.-L. George, Mesh Generation. Application to finite elements, Hermes Science, 2000.