Dobbiaco, July 1-5, 2005
Mathematical Models in Life Science: Theory and Simulation

An Introduction to Finite Elements Mathematical Aspects

Alfio Quarteroni & Alessandro Veneziani

CMCS
Chaire de Modélisation et Calcul Scienti↓que
Ecole Polytechnique Fédérale de Lausanne (EPFL)
iacs.epfl.ch/cmcs

MOX
Dipartimento di Matematica
Politecnico di Milano
mox.polimi.it

Outline

- Introduction to the basic properties of the Finite Elements Method
- Basic reference: A. Quarteroni, A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer-Verlag, Heidelberg, 1996, Chapters 3,5.

Integral Formulation Of Boundary Value Problems

Consider the following **boundary value problem**:

$$-(\alpha u')'(x) + (\beta u')(x) + (\gamma u)(x) = f(x) \quad 0 < x < 1, \tag{1}$$

$$u(0) = u(1) = 0. \tag{2}$$

where α, β and γ are **continuous functions** on $[0,1]$ with $\alpha(x) \geq \alpha_0 > 0$ for any $x \in [0,1]$.

Multiply (1) by *a function* $v \in C^1([0,1])$, hereafter called a **"test function"**, and *integrate* over the interval $[0,1]$

$$\int_0^1 \alpha u'v' \, dx + \int_0^1 \beta u'v \, dx + \int_0^1 \gamma uv \, dx = \int_0^1 fv \, dx + [\alpha u'v]_0^1,$$

(note the **integration by parts** on the first integral.)

Integral Formulation Of Boundary Value Problems

If we suppose $v = 0$ at $x = 0$ and $x = 1$:

$$\boxed{\int_0^1 \alpha u'v' \, dx + \int_0^1 \beta u'v \, dx + \int_0^1 \gamma uv \, dx = \int_0^1 fv \, dx.}$$

Denote by V the **test function space**.

V is given by all functions v that are continuous, vanish at $x = 0$ and $x = 1$ and whose first derivative is *piecewise continuous*,
i.e., continuous everywhere except at a finite number of points in $[0,1]$ where the left and right limits v'_- and v'_+ exist but do not necessarily coincide.

V is a vector space usually denoted by $\mathrm{H}_0^1(0,1)$:

$$\mathrm{H}_0^1(0,1) = \left\{ v \in \mathrm{L}^2(0,1) : \; v' \in \mathrm{L}^2(0,1), \; v(0) = v(1) = 0 \right\} \tag{3}$$

Integral Formulation Of Boundary Value Problems

A **solution** $u \in C^2([0,1])$ of (1), is also a solution of **the following problem**

$$\text{find } u \in V: \ a(u,v) = (f,v) \ \text{ for all } v \in V, \tag{4}$$

where $(f,v) = \int_0^1 fv \, dx$ (scalar product of $\mathrm{L}^2(0,1)$) and

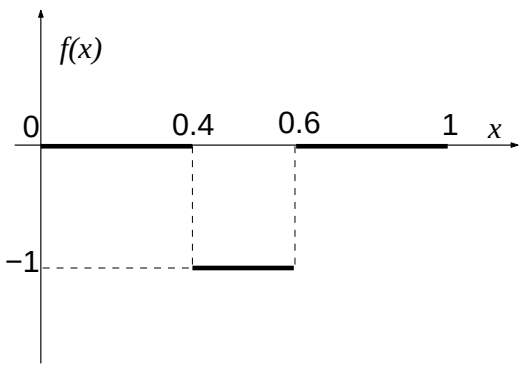
$$\boxed{a(u,v) = \int_0^1 \alpha u'v' \, dx + \int_0^1 \beta u'v \, dx + \int_0^1 \gamma uv \, dx} \tag{5}$$

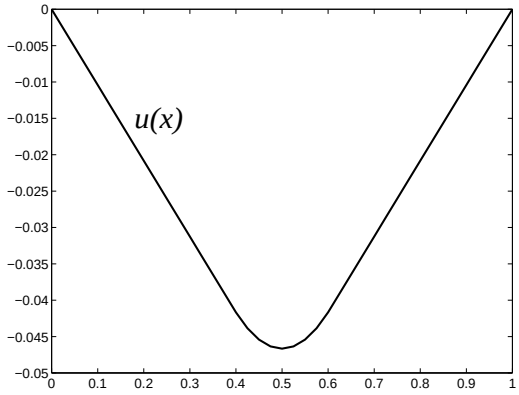
$a(\cdot,\cdot)$ is a **bilinear form**, linear with respect to both arguments.

➪ **Problem (4) is called the** *weak formulation* **of problem (1).**

Integral Formulation Of Boundary Value Problems

➪ (4) contains only the **first derivative of** u: it might cover cases in which a classical solution $u \in C^2([0,1])$ of (1) does not exist although the physical problem is well defined:
e.g. $\alpha = 1$, $\beta = \gamma = 0$:



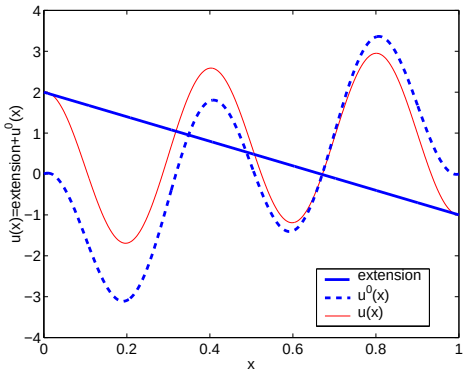


Non homogeneous boundary conditions

Equation (1) can be supplied with **non homogeneous boundary conditions**: $u(0) = u_0$, $u(1) = u_1$.

A formulation like (4) is obtained by proceeding as follows.

$\bar{u}(x) = xu_1 + (1-x)u_0$ is the straight line that interpolates the data at the endpoints (usually called **extension of the boundary data**) and set $\overset{0}{u} = u(x) - \bar{u}(x)$.

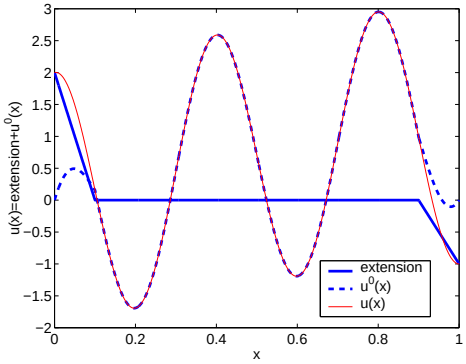


$\overset{0}{u} \in V$ satisfi es the following problem

$$\text{fi nd } \overset{0}{u} \in V : \; a(\overset{0}{u}, v) = (f, v) - a(\bar{u}, v) \;\; \text{for all } v \in V.$$

Non homogeneous boundary conditions

Different extensions of the boundary data can be considered as well.



Different b.c. can be considered as well: e.g. Neumann boundary conditions, $u'(0) = u'(1) = 0$.

➪ **We still obtain a weak form (4), provided the space V is now $\mathrm{H}^1(0,1)$.**

Formulation and Properties of the Galerkin Method

The Galerkin method **is based on the weak formulation** (4).
V_h denotes **a finite dimensional vector subspace of** V

- The Galerkin method reads

$$\text{find } u_h \in V_h : \ a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h. \tag{6}$$

This is a finite dimensional problem.
Let $\{\varphi_1, \ldots, \varphi_N\}$ denote a set of N linearly independent functions of V_h: **a BASIS** of V_h.

- We can write

$$u_h(x) = \sum_{j=1}^{N} u_j \varphi_j(x).$$

The integer N denotes the **dimension** of the vector space V_h.

Formulation and Properties of the Galerkin Method

Taking $v_h = \varphi_i$ as *test functions*, the Galerkin problem (6) is equivalent to seeking N unknown coefficients $\{u_1, \ldots, u_N\}$ such that

$$\sum_{j=1}^{N} u_j a(\varphi_j, \varphi_i) = (f, \varphi_i) \qquad \forall i = 1, \ldots, N. \tag{7}$$

Set $\mathrm{A_G} = [a_{ij}]$, with

$$a_{ij} = a(\varphi_j, \varphi_i)$$

the stiffness matrix, the unknown vector $\mathbf{u} = (u_1, \ldots, u_N)$ and the right-hand side vector $\mathbf{f_G} = (f_1, \ldots, f_N)$, with $f_i = (f, \varphi_i)$.

Problem (7) is equivalent to **the linear system**

$$\mathrm{A_G}\mathbf{u} = \mathbf{f_G}. \tag{8}$$

Formulation and Properties of the Galerkin Method

The structure of $\mathrm{A_G}$, as well as the degree of accuracy of u_h, depends on the form of the basis functions $\{\varphi_i\}$, and therefore **on the choice of V_h**.

Two remarkable instances:

- the *finite element method*, where V_h is a space of piecewise polynomials over subintervals of $[0,1]$ of length not greater than h which are **continuous and vanish at the endpoints** $x = 0$ and 1
- the *spectral method* in which V_h is a space of algebraic polynomials (typically featuring high degree) still vanishing at the endpoints $x = 0, 1$.

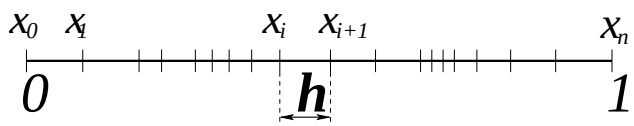
The Finite Element Method

The finite element method (FEM) is a special technique for constructing a subspace V_h based **on the piecewise polynomial interpolation**.

We introduce a partition $\mathcal{T}_h$ of [0,1] into n subintervals $I_j = [x_j, x_{j+1}]$, $n \geq 2$, of width $h_j = x_{j+1} - x_j$, $j = 0, \ldots, n-1$, with

$$0 = x_0 < x_1 < \ldots < x_{n-1} < x_n = 1$$

and let $h = \max_{\mathcal{T}_h}(h_j)$.



The Finite Element Method

Consider for $k \geq 1$ **the family of piecewise polynomials** X_h^k:

$$X_h^k = \left\{ v \in C^0([0,1]) : v|_{I_j} \in \mathbb{P}_k(I_j) \, \forall I_j \in \mathcal{T}_h \right\} \tag{9}$$

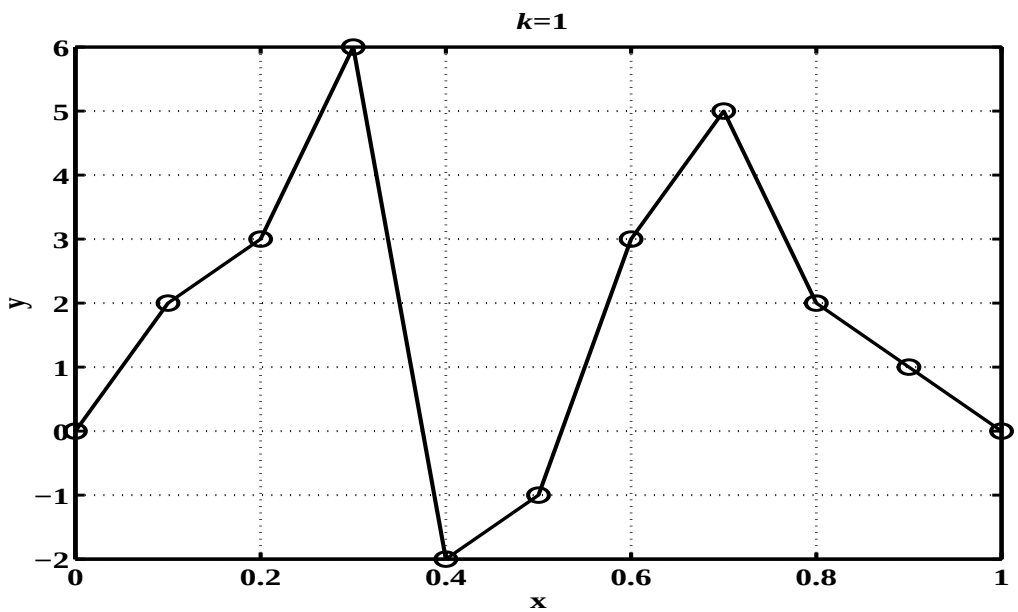
(Remember that functions in H_0^1 in 1D are continuous)

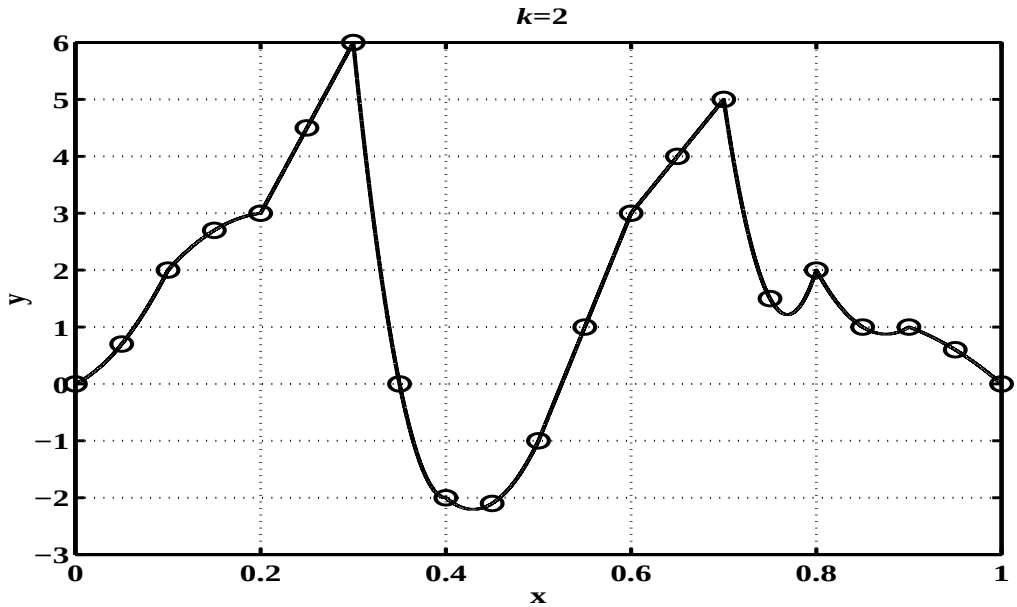
- Functions $v_h \in X_h^k$ are continuous piecewise polynomials over $[0,1]$
- Their **restriction** over each interval $I_j \in \mathcal{T}_h$ is a **polynomial of degree** $\leq k$.

Now, we set

$$V_h = X_h^{k,0} = \left\{ v_h \in X_h^k \, : \, v_h(0) = v_h(1) = 0 \right\}. \tag{10}$$

The dimension N of the finite element space V_h is equal to $nk - 1$.





Accuracy of the Galerkin FEM

Since:

$$\min_{w_h \in V_h} \|u - w_h\|_{\mathrm{H}_0^1(0,1)} \leq \|u - \Pi_h^k u\|_{\mathrm{H}_0^1(0,1)} \tag{11}$$

where $\Pi_h^k u$ is the interpolant of the exact solution $u \in V$, from the Céa Lemma, the Galerkin ***approximation error*** $\|u - u_h\|_{\mathrm{H}_0^1(0,1)}$ can be estimated by the *interpolation error* $\|u - \Pi_h^k u\|_{\mathrm{H}_0^1(0,1)}$.

Theorem

Let $u \in \mathrm{H}_0^1(0,1)$ be the exact solution and $u_h \in V_h$ its finite element approximation *using continuous piecewise polynomials of degree $k \geq 1$. Assume also that $u \in \mathrm{H}^s(0,1)$ for some $s \geq 2$. Then the following error estimate holds*

$$h^{-1}|u - u_h| + |D(u - u_h)| \leq \frac{M}{\alpha_0} C h^l |D^{l+1} u| \tag{12}$$

where $l = \min(k, s-1)$. M and α_0 are respect. the continuity and coercivity constants of the bilinear form $a(\cdot,\cdot)$. C is a positive constant independent of h. Here, $|w| := \sqrt{\int_\Omega w^2 dx}$ and $D^r w := d^r w / dx^r$.

Consequences

Convergence The estimate (12) shows that the Galerkin method is ***convergent***, i.e. the approximation error tends to zero as $h \to 0$ and the order of convergence is k.

Regularity threshold There is **no convenience** to increase the degree k of the finite element approximation **if the solution u is not sufficiently smooth**.
l is called a *regularity threshold*.

To **gain accuracy**: **reduce the stepsize h**.
The following table summarizes the orders of convergence of the FEM for $k = 1, \ldots, 4$ and $s = 1, \ldots, 5$.

k	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$
1	only convergence	$\boxed{h^1}$	h^1	h^1	h^1
2	only convergence	h^1	$\boxed{h^2}$	h^2	h^2
3	only convergence	h^1	h^2	$\boxed{h^3}$	h^3
4	only convergence	h^1	h^2	h^3	$\boxed{h^4}$

The Finite Element Basis

How to generate a suitable basis $\{\varphi_j\}$ for the finite element space X_h^k ($k = 1$ and $k = 2$) ?
The basic point is to choose appropriately a set of *degrees of freedom* for each element I_j of the partition $\mathcal{T}_h$, i.e., **the parameters** which permit to **uniquely identify a function in X_h^k**.
The generic function v_h in X_h^k can therefore be written as

$$v_h(x) = \sum_{i=0}^{nk} v_i \varphi_i(x)$$

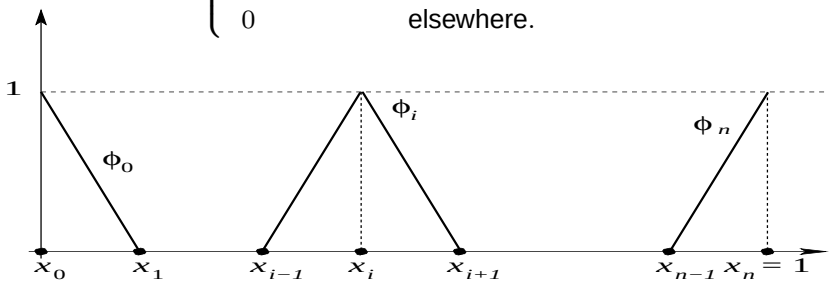
where $\{v_i\}$ denote the set of the degrees of freedom of v_h
The basis functions φ_i (which are also called *shape functions*) satisfy **the Lagrange interpolation property**:

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

The space X_h^1 - Continuous piecewise linear functions

The number of degrees of freedom for v_h is equal to the number $n+1$ of nodes in the partition.
The most natural choice for φ_i, $i = 1, \ldots, n-1$, is

$$\varphi_i(x) \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & \text{for } x_{i-1} \le x \le x_i, \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} & \text{for } x_i \le x \le x_{i+1}, \\ 0 & \text{elsewhere.} \end{cases} \tag{13}$$

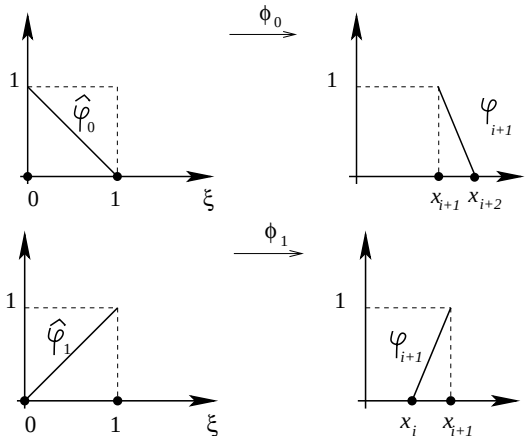


The support (i.e., the subset of $[0,1]$ where φ_i is non-vanishing) consists of the union of the intervals I_{i-1} and I_i if $1 \le i \le n-1$; it coincides with the interval I_0 (respectively I_{n-1}) if $i = 0$ (resp., $i = n$).

For any interval $I_i = [x_i, x_{i+1}]$, $i = 0, \ldots, n-1$, the two basis functions φ_i and φ_{i+1} can be regarded as **the images of two "reference" shape functions $\widehat{\varphi}_0$ and $\widehat{\varphi}_1$**, defined over the *reference interval* $[0,1]$ through the linear affine mapping $\phi : [0,1] \to I_i$

$$x = \phi(\xi) = x_i + \xi(x_{i+1} - x_i), \qquad i = 0, \ldots, n-1. \tag{14}$$

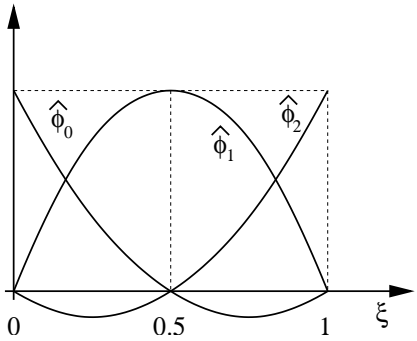
Defining $\widehat{\varphi}_0(\xi) = 1 - \xi$, $\widehat{\varphi}_1(\xi) = \xi$ we have $\varphi_i(x) = \widehat{\varphi}_0(\xi(x))$ and $\varphi_{i+1}(x) = \widehat{\varphi}_1(\xi(x))$, where $\xi(x) = (x - x_i)/(x_{i+1} - x_i)$



The space X_h^2 - Piecewise polynomial of degree 2

Every function can be uniquely determined once **three values of it at three distinct points of I_i are assigned**. To ensure continuity of v_h over $[0,1]$ the degrees of freedom are chosen as the function values at the nodes x_i of $\mathcal{T}_h$, $i = 0, \ldots, n$, and at the midpoints of each interval I_i, $i = 0, \ldots, n-1$, for a total number equal to $2n+1$. The shape functions for X_h^2 on the **reference interval** $[0,1]$ are

$$\widehat{\phi}_0(\xi) = (1-\xi)(1-2\xi), \quad \widehat{\phi}_1(\xi) = 4(1-\xi)\xi, \quad \widehat{\phi}_2(\xi) = \xi(2\xi-1)$$



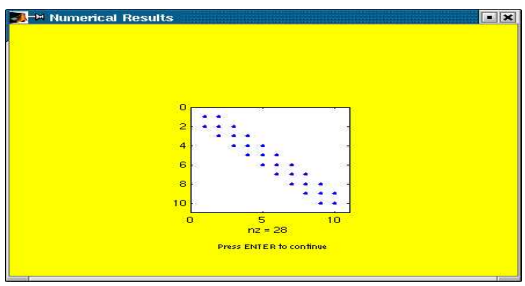
The shape functions and are still the images of the reference functions through the affi ne mapping (14).

Solving the FEM linear system

The fi nite element basis functions of X_h^k have a local support. Therefore

A_{fe} is *sparse*

- $k = 1$: the support of the shape function φ_i is the union of the intervals I_{i-1} and I_i if $1 \le i \le n-1$ For a fixed $i = 1, \ldots, n-1$, only the shape functions φ_{i-1} and φ_{i+1} have a non-vanishing support intersection with that of φ_i
  A_{fe} is ***tridiagonal***

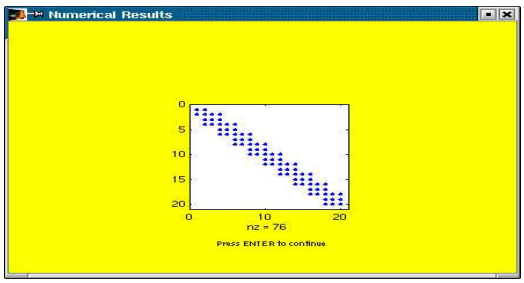


Solving the FEM linear system

The fi nite element basis functions of X_h^k have a local support. Therefore

A_{fe} is *sparse*

- $k = 1$: the support of the shape function φ_i is the union of the intervals I_{i-1} and I_i if $1 \leq i \leq n-1$ For a fixed $i = 1, \ldots, n-1$, only the shape functions φ_{i-1} and φ_{i+1} have a non-vanishing support intersection with that of φ_i
  A_{fe} is ***tridiagonal***
- $k = 2$:
  A_{fe} is a ***pentadiagonal* matrix**

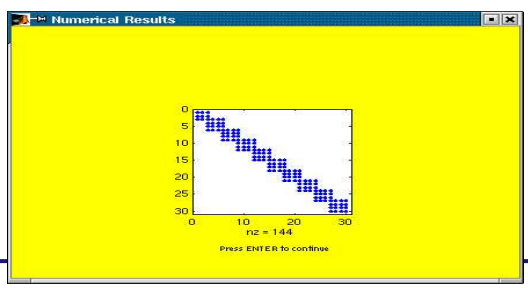


Solving the FEM linear system

The fi nite element basis functions of X_h^k have a local support. Therefore

A_{fe} is *sparse*

- $k = 1$: the support of the shape function φ_i is the union of the intervals I_{i-1} and I_i if $1 \leq i \leq n-1$ For a fixed $i = 1, \ldots, n-1$, only the shape functions φ_{i-1} and φ_{i+1} have a non-vanishing support intersection with that of φ_i
  A_{fe} is ***tridiagonal***
- $k = 2$:
  A_{fe} is a ***pentadiagonal* matrix**
- $k = 3$:
  A_{fe} is an ***eptadiagonal* matrix**

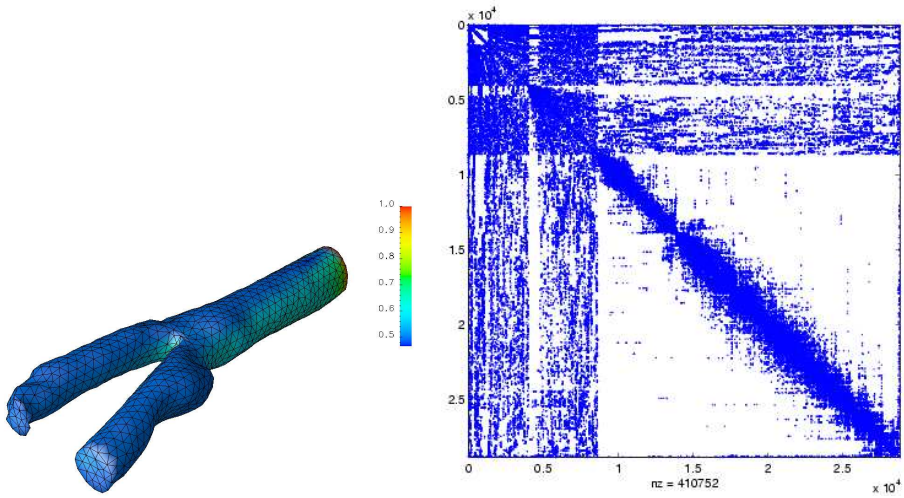


Dobbiaco - July 2005 – p. 22/43

Solving the FEM linear system

In general, for multidimensional domains, the matrix is still sparse, but the pattern does not necessarily have a specific structure.

It depends on **the criteria adopted for numbering the degrees of freedom**



Solving the FEM linear system

In real life applications $\mathrm{A_{fe}}$ features **large dimensions**.

The efficiency of a finite element code is strictly related to **the efficiency of its linear solver**.

Typically, numerical methods for the solution of linear system are subdivided into two classes:

- **Direct Methods**
- **Iterative Methods**

Direct Methods

Based on the factorization:

$$\mathrm{A} = \mathrm{LU} \tag{15}$$

where L and U are lower and upper triangular matrices.
The solution resorts to the solution of the following two triangular systems:

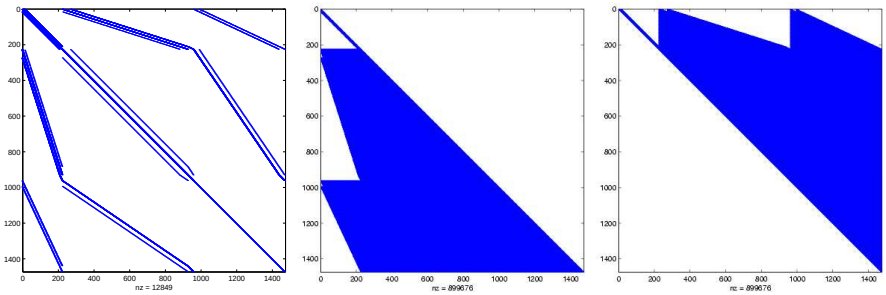
$$\mathrm{L}\mathbf{u} = \mathbf{b}, \quad \mathrm{U}\mathbf{x} = \mathbf{y}. \tag{16}$$

This factorization is possible, e.g., for the matrix A_{fe} in the case of elliptic problems, **thanks to the positivity of the matrix.**

Direct Methods

Main drawback of LU factorization in the context of FEM:
new non-zero elements generated by the factorization (→ *fill in*).

$$\Rightarrow \textbf{High storage resources.}$$



Pattern of a fi nite element (3D) matrix (left) and of its L (middle) and U (right) factors.
The presence of the *fill in* is relevant: the original matrix features 12849 non null entries, while each triangular factor has 899676 non null entries. 1799352 real numbers are stored instead of 12849 when carrying out the factorization.

Iterative Methods

Conjugate Gradient:
The *conjugate gradient* (CG) method applies to the s.p.d. case.
Setting $\mathbf{r}^{(0)} = \mathbf{b} - \mathrm{A}\mathbf{x}^{(0)}$ and $\mathbf{p}^{(0)} = \mathbf{r}^{(0)}$, **the k-th iteration of CG reads**:

$$\alpha_k = \frac{{\mathbf{p}^{(k)}}^T \mathbf{r}^{(k)}}{{\mathbf{p}^{(k)}}^T \mathrm{A}\mathbf{p}^{(k)}}, \rightarrow \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}$$
$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathrm{A}\mathbf{p}^{(k)}$$
$$\beta_k = \frac{(\mathrm{A}\mathbf{p}^{(k)})^T \mathbf{r}^{(k+1)}}{(\mathrm{A}\mathbf{p}^{(k)})^T \mathbf{p}^{(k)}}, \rightarrow \mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} - \beta_k \mathbf{p}^{(k)}.$$

In this case, the convergence rate is driven by:

$$\|\mathbf{x}^{(k)} - \mathbf{x}_{ex}\|_A^2 \leq 2 \left(\frac{\sqrt{\mathcal{K}(\mathrm{A})} - 1}{\sqrt{\mathcal{K}(\mathrm{A})} + 1} \right)^2 k \|\mathbf{x}^{(0)} - \mathbf{x}_{ex}\|_A^2 \qquad (17)$$

$\Rightarrow$ ***The convergence rate is low when $\mathcal{K}(\mathrm{A})$ is large***.
(e.g. $\mathcal{K}(\mathrm{A}) = \mathcal{O}(h^{-2})$ for FEM on second order problems.)

Preconditioning

The original system can be modifi ed:

$$\mathrm{P}^{-1}\mathrm{A}\mathbf{x} = \mathrm{P}^{-1}\mathbf{b}$$

P is a matrix called *preconditioner* to be suitably set up.
For a *preconditioned conjugate gradient* method convergence is driven by:

$$\frac{\sqrt{\mathcal{K}(\mathrm{P}^{-1}\mathrm{A})} - 1}{\sqrt{\mathcal{K}(\mathrm{P}^{-1}\mathrm{A})} + 1}$$

A good preconditioner:

1. should **reduce the condition number** $\mathcal{K}(\mathrm{P}^{-1}\mathrm{A})$; since the lowest (spectral) condition number attainable is the one of the identity matrix and it is equal to 1, P should identify a good "spectral" approximation of A;

2. ought be **easy to solve**; since at each step a system for the preconditioner P must be solved, the preconditioned scheme is computationally feasible only if the preconditioner can be solved with a (reasonably) low computational effort.

A good preconditioner is the compromise between these requirements.

The Multi-Dimensional Case

The **generalization of the elastic string problem** is

$$\begin{cases} -\triangle u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{18}$$

where $\triangle u = \partial^2 u/\partial x^2 + \partial^2 u/\partial y^2$ is the **Laplace operator**, and Ω is a 2D (or a 3D) bounded domain whose boundary is $\partial\Omega$.
More in general, we could refer to the problem:

$$\begin{cases} -\nabla \cdot (\mu \nabla u) + \beta \cdot \nabla u + \sigma u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{19}$$

μ, β, σ suitable functions of x_1, x_2 (and possibly x_3)

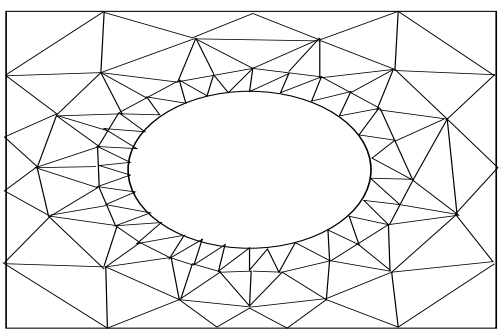
The Galerkin method still leads (formally) to the problem:

$$\text{find } u_h \in V : \ a(u_h, v_h) = (f, v_h) \ \text{ for all } v_h \in V_h,$$

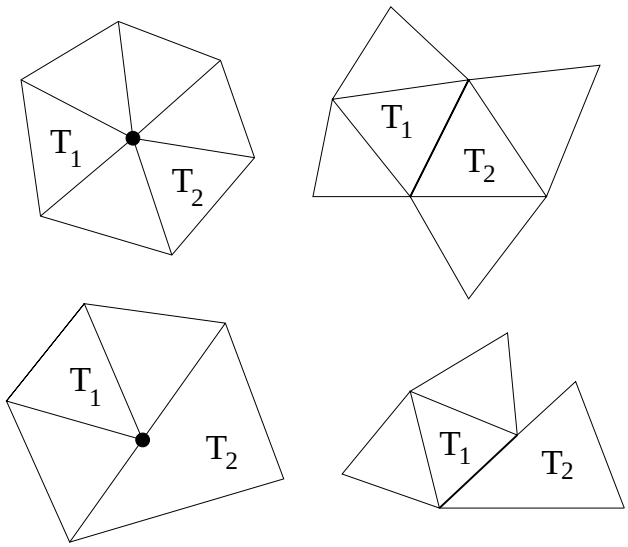
Both the function space V_h and the bilinear form $a(\cdot,\cdot)$ have to be **adapted to the problem at hand**.
Partition the domain Ω into K **non-overlapping triangles** (or tetrahedra in 3D or *elements* in general) T providing a *triangulation* $\mathcal{T}_h$ of the domain such that

$$\overline{\Omega} = \bigcup_{T \in \mathcal{T}_h} T.$$

h is the maximum length of the edges of the triangles.



The **admissible triangulations** are those for which any pair of non disjoint triangles may have either a vertex or an edge in common: no ambiguity in the definition of a point.



Admissible (top) and non-admissible (bottom) triangulations.

Local basis functions:

Introduce on $\mathcal{T}_h$ the set $\mathcal{Z}$ of the *piecewise interpolation nodes*

$$\mathbf{z}_i = (x_i, y_i)^T, \text{ for } i = 1, \ldots, N.$$

Let $\mathbb{P}_k(\Omega)$, $k \geq 0$, the the space of algebraic polynomials of degree $\leq k$ in the space variables x, y:

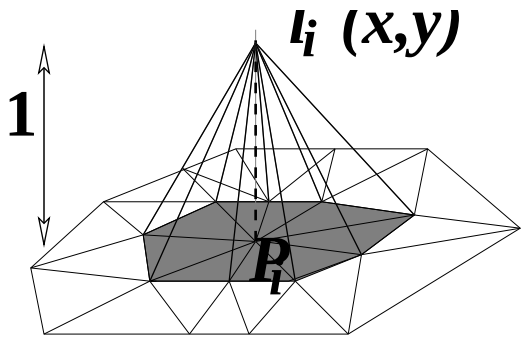
$$\mathbb{P}_k(\Omega) = \{p(x,y) = \sum_{\substack{i,j=0 \\ i+j \leq k}}^{k} a_{ij} x^i y^j,\ x, y \in \Omega\}. \tag{20}$$

For $k \geq 0$, let $\mathbb{P}_k^c(\Omega)$ be the space of piecewise polynomials of degree $\leq k$, such that, for any $p \in \mathbb{P}_k^c(\Omega)$, $p|_T \in \mathbb{P}_k(T)$ for any $T \in \mathcal{T}_h$.
An **elementary basis** for $\mathbb{P}_k^c(\Omega)$ consists of the *Lagrange characteristic polynomials* $l_i = l_i(x,y)$, such that $l_i \in \mathbb{P}_k^c(\Omega)$ and

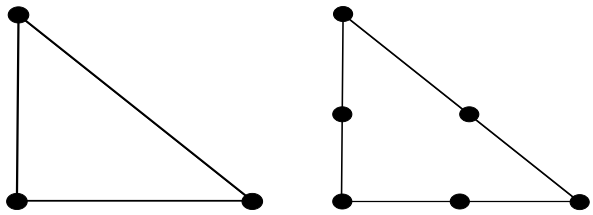
$$l_i(\mathbf{z}_j) = \delta_{ij}, \qquad i, j = 1, \ldots, N, \tag{21}$$

where $\delta_{ij} = 0$ if $i \neq j$ and $= 1$ if $i = j$ (Kronecker symbol).

In the linear case:



If the interpolation nodes coincide with the *vertices* of the triangles, the resulting piecewise **linear** functions will be continuous.



Local interpolation nodes on $\hat{T}$; left, $k = 1$ ($d_1 = 3$), right, $k = 2$ ($d_2 = 6$).

This is not the only possible choice. The midpoints of the edges of the triangles could be used as well, giving rise to a discontinuous piecewise polynomial over Ω.

For $k \geq 0$, the *Lagrange piecewise interpolating polynomial* of f, $\Pi_h^k f \in \mathbb{P}_k^c(\Omega)$, is then given by:

$$\Pi_h^k f(x,y) = \sum_{i=1}^{N} f(\mathbf{z}_i) l_i(x,y). \tag{22}$$

Finite Element Subspace V_h
The finite element method corresponds to taking

$$\boxed{V_h = \left\{ v_h \in C^0(\overline{\Omega}) \ : \ v_{h|_T} \in \mathbb{P}_k(T) \, \forall T \in \mathcal{T}_h, \ v_{h|_{\partial\Omega}} = 0 \right\}.} \tag{23}$$

Bilinear form

As of the bilinear form $a(\cdot,\cdot)$, using the Green's formula that generalizes the formula of integration by parts

$$\int_\Omega -\nabla\cdot(\mu\nabla u)\; v\, dxdy = \int_\Omega \mu\nabla u\cdot\nabla v\, dxdy - \int_{\partial\Omega}\mu\nabla u\cdot\mathbf{n}\, v\, d\gamma, \qquad (24)$$

for any u, v smooth enough and where $\mathbf{n}$ is the outward normal unit vector on $\partial\Omega$, we have:

$$\boxed{a(u_h,v_h) = \int_\Omega \mu\nabla u_h\cdot\nabla v_h\, dxdy + \int_\Omega \beta\cdot\nabla u_h v_h\, dxdy + \int_\Omega \sigma u_h v_h\, dxdy}$$

The FE problem reads:

$$\text{find } u_h\in V_h : a(u_h,v_h) = \mathcal{F}(v_h),\ \forall v_h\in V_h. \qquad (25)$$

Error Analysis

We still combine:

Céa Lemma + Interpolation Error Estimates:

Theorem

Let $u\in \mathrm{H}_0^1(\Omega)$ be the exact solution of (19) and $u_h\in V_h$ be its finite element approximation using continuous piecewise polynomials of degree $k\geq 1$. Assume also that $u\in \mathrm{H}^s(\Omega)$ for some $s\geq 2$. Then the following error estimate holds

$$\boxed{h^{-1}|u-u_h| + |D(u-u_h)| \leq \frac{M}{\alpha_0} C h^l |D^{l+1}u|} \qquad (26)$$

where $l=\min(k,s-1)$. The constants M, α_0 and C are as in (12), while D^l denotes the set of all partial derivatives of order l.

Remarks:

- The finite element solution u_h can be written as:

$$u_h(x,y) = \sum_{j=1}^{N} u_j \varphi_j(x,y).$$

  The Galerkin finite element method leads to the solution of the linear system $\mathrm{A_{fe}}\mathbf{u} = \mathbf{f}$, where $[A_{\mathrm{fe}}]_{ij} = a(\varphi_j, \varphi_i)$ and $\mathbf{u}_i = u_i$.
- $\mathrm{A_{fe}}$ is positive definite (and symmetric provided the bilinear form $a(\cdot,\cdot)$ is symmetric).
- Its **sparsity pattern** depends on the topology of $\mathcal{T}_h$ and the numbering of its nodes.
- The spectral condition number of $\mathrm{A_{fe}}$ is still $\mathcal{O}(h^{-2})$.

Numerical Example: An elliptic problem in 2D

Consider the following problem:

$$-\Delta u = \pi^2 \sin(\pi y) - 6(x-1)H(x-1) \tag{27}$$

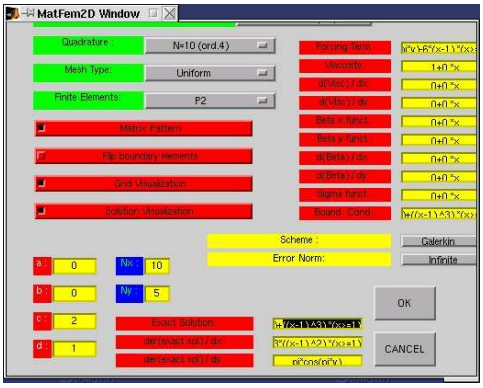
where $H(x)$ is the Heaviside step function, such that:

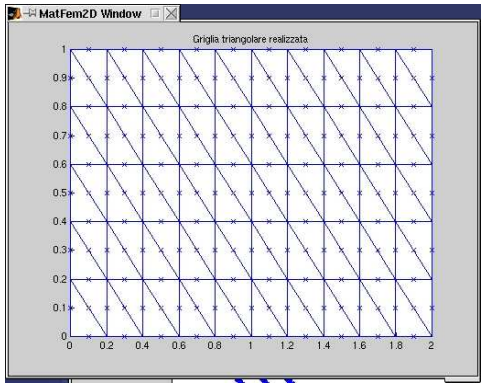
$$H(x) \begin{cases} 0 & \text{for } x < 0; \\ \\ 1 & \text{for } x > 0. \end{cases}$$

The computational domain is the $\Omega = (0,2) \times (0,1)$ with Dirichlet boundary conditions:

$$u = \sin(\pi y) + (x-1)^3 H(x-1). \tag{28}$$

It is possible to prove that $u(x,y) \in H^3(\Omega)$.


MatFem2D Window
Quadrature
N=10 (ord.4)
Mesh Type
Uniform
Finite Elements
P2
Matrix Pattern
Flip boundary elements
Grid Visualization
Solution Visualization
Forcing Term
Viscosity
Beta x funct.
Beta y funct.
sigma funct.
Bound. Cond.
Scheme :
Galerkin
Error Norm:
Infinite
Nx
10
Ny
5
Exact Solution
OK
CANCEL


MatFem2D Window
Griglia triangolare realizzata

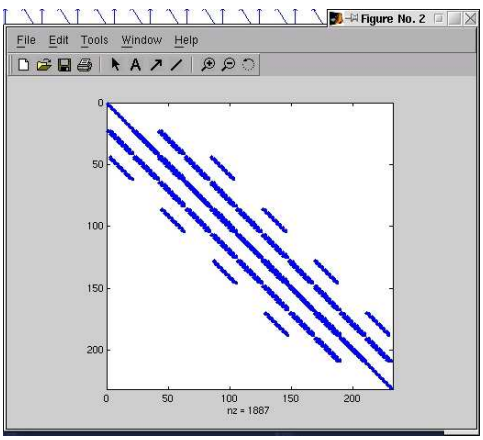


Figure No. 2
File Edit Tools Window Help
nz = 1887

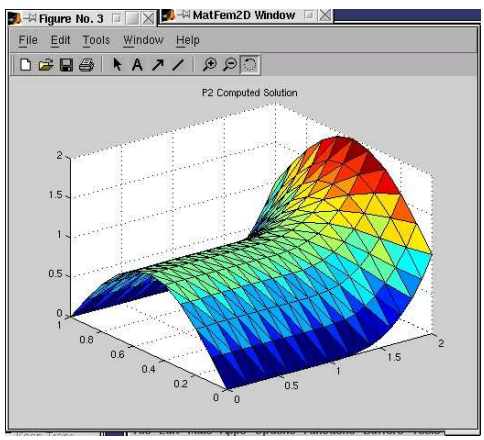
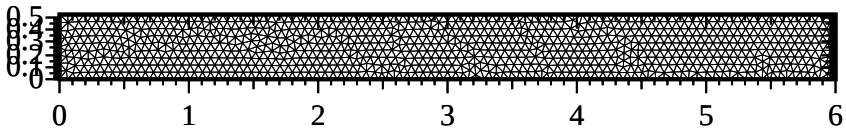
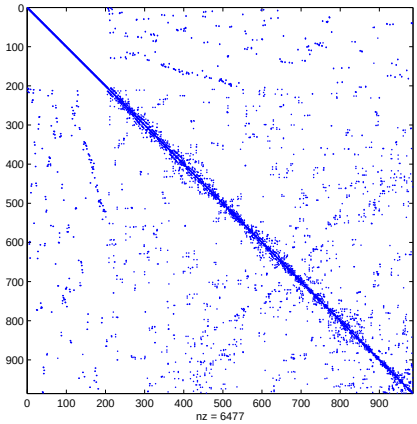


Figure No. 3
MatFem2D Window
File Edit Tools Window Help
P2 Computed Solution

A different grid will give a different matrix pattern:





Error Analysis:

	$h = 0.2$	$h = 0.1$	$h = 0.05$
P1	0.600956	0.301583	0.150929
P2	0.047083	0.0117509	0.00294124

Consider now the following modification:

$$-\Delta u = \pi^2 \sin(\pi y) - 2H(x-1) \tag{29}$$

in the domain $\Omega = (0,2) \times (0,1)$ with the exact solution:

$$u = \sin(\pi y) + (x-1)^2 H(x-1). \tag{30}$$

It is possible to prove that $u(x, y) \in H^2(\Omega)$.

Error Analysis:

	$h = 0.2$	$h = 0.1$	$h = 0.05$
P1	0.57864	0.290369	0.145325
P2	0.0465805	0.0121859	0.00349274