

# ① Chapter 7 Abstract FE Analysis

Ex: Recall from FEM 1 :

$$\left\{ \begin{array}{l} -\nabla \cdot a \nabla u + b \cdot \nabla u + cu = f \\ u = 0 \end{array} \right.$$

On weak form: find  $u \in V$  s.t

$$\int_a \nabla u \cdot \nabla v \, dx + \int_b \nabla u \cdot v \, dx + \int_c u \cdot v \, dx = \int_f v \, dx \quad \forall v \in V.$$

- \* Function spaces, What is  $V$ ?
  - \* Existence of weak solutions? Uniqueness?
  - \* Approximation with discrete functions, Interpolation
  - \* Finite Element Method
  - \* Convergence and error analysis.
- 

## Function Spaces

Let  $V$  be a vector space fulfilling basic properties, see pg 179.

We say that the mapping  $\| \cdot \| : V \rightarrow \mathbb{R}$  is a norm if

- (i)  $\|u+v\| \leq \|u\| + \|v\|$ , (ii)  $\|ku\| = |k| \cdot \|u\|$ , (iii)  $\|u\| \geq 0$  with equality iff  $u=0$

② We say that the map  $(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$   
 is an inner (scalar) product if

$$(i) (u, v) = (v, u) \quad (ii) (u + \alpha v, w) = \alpha(u, w) + (v, w)$$

$$(iii) (u, u) \geq 0 \text{ with equality iff } u = 0.$$

A Vector space with an inner product is  
 called an inner product space. An inner  
 product directly gives a norm namely  $\|v\|^2 = (v, v)$ .

In an inner product space it holds

Cauchy-Schwarz inequality:  $(u, v) \leq \|u\| \cdot \|v\| \quad \forall u, v \in V$   
 with  $\|w\| = (w, w)^{1/2}$

The parallelogram law:  $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$

### Complete Spaces

A Cauchy Sequence  $\{v_i\}_{i=1}^{\infty} \in V$  fulfills  $\forall \epsilon > 0$

$\exists n \text{ s.t. } \|v_i - v_j\| \leq \epsilon, i, j \geq n$

A Convergent Sequence  $\{v_i\}_{i=1}^{\infty} \in V$  fulfills  $\forall \epsilon > 0$

$\exists n \text{ s.t. } \|v - v_i\| \leq \epsilon, i \geq n$  for some  $v \in V$ .

If all Cauchy sequences are convergent  
 the space is complete. A complete normed  
 vector space is a Banach space. A complete  
 inner product space is a Hilbert space.

### ③ $L^p$ Spaces

Let  $\|v\|_{L^p(\Omega)} = \left( \int_{\Omega} |v|^p dx \right)^{1/p}$ ,  $1 \leq p < \infty$

$\|v\|_{L^p(\Omega)} = \sup_{x \in \Omega} |v(x)|$ ,  $p = \infty$

and  $L^p(\Omega) = \{ v : \Omega \rightarrow \mathbb{R} : \|v\|_{L^p(\Omega)} < \infty \}$ .

$L$  stands for Lebesgue who defined the integral.

It can be shown that  $L^p(\Omega)$  is a Banach Space and  $L^2(\Omega)$  is a Hilbert Space with inner product  $(u, v)_{L^2(\Omega)} = \int_{\Omega} u \cdot v dx$ .

Weak derivatives

Let  $D(\Omega) = \{ \varphi \in C_c^\infty(\Omega) : \text{supp } \varphi \subset \subset \Omega \}$ ,

where  $w \subset \subset \Omega$  means that  $w$  is contained in a closed bounded subset of  $\Omega$ .

Let  $\alpha = (\alpha_1, \dots, \alpha_d)$  in  $d$  spatial dimensions be a multi-index with  $|\alpha| = \sum_{i=1}^d \alpha_i$ ,  $\alpha_i \geq 0$ .

$$D^\alpha \varphi = \prod_{i=1}^d \left( \frac{\partial}{\partial x_i} \right)^{\alpha_i} \varphi, \quad \varphi \in D(\Omega).$$

Using partial integration and  $\varphi \in D(\Omega)$  we have

$$\int_{\Omega} \frac{\partial u}{\partial x_i} \varphi dx = - \int_{\Omega} u \cdot \frac{\partial \varphi}{\partial x_i} dx \quad \forall \varphi \in D(\Omega)$$

④ In general on a bounded domain sufficiently regular boundary

$$\int_{\Omega} (D^\alpha u) \varphi dx = (-1)^{|\alpha|} \int_{\Omega} u D^\alpha \varphi dx, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

We note that only LHS requires derivatives of  $u$ .

Let  $u \in L^1(\Omega)$ . If there is a  $g \in L^1(\Omega)$  s.t

$$\int_{\Omega} g \cdot \varphi dx = (-1)^{|\alpha|} \int_{\Omega} u \cdot D^\alpha \varphi dx \quad \forall \varphi \in \mathcal{D}(\Omega)$$

We say  $g$  is the weak derivative  $D^\alpha u$  of  $u$ .

Exe 7.2

### Sobolev spaces

Let  $u \in L^1(\Omega)$  and all weak derivatives  $D^\alpha u$ ,  $|\alpha| \leq k$  exist. Then

$$\|u\|_{W_k^p(\Omega)} = \left( \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p}, \quad 1 \leq p < \infty$$

$$\|u\|_{W_k^\infty(\Omega)} = \max_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)} \quad p = \infty$$

Further let  $W_k^p(\Omega) = \{u \in L^1(\Omega) : \|u\|_{W_k^p(\Omega)} < \infty\}$

Sobolev spaces are Banach spaces.

When  $p=2$  they are Hilbert spaces with inner product  $(u, v)_{W_k^2(\Omega)} = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}$

and norm  $\|u\|_{W_k^2(\Omega)} = (u, u)_{W_k^2(\Omega)}^{1/2}$

⑤ For  $k=1$  we let  $H^1(\Omega) = W_+^2(\Omega)$  with  
 $(u, v)_{H^1(\Omega)} = (u, v)_{L^2(\Omega)} + (\nabla u, \nabla v)_{L^2(\Omega)}$  and  
 $\|u\|_{H^1(\Omega)} = (u, u)_{H^1(\Omega)}^{1/2} = (\|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2)^{1/2}$ .

### Traces

Functions in  $H^1(\Omega)$  are in 1D continuous, in 2D they may have singular points, and in 3D they may have singularities along curves. Therefore it is unclear what  $v|_{\partial\Omega}$  means for a general  $v \in H^1(\Omega)$  in 2,3D.

The trace operator  $\gamma: W_+^1(\Omega) \rightarrow L^p(\partial\Omega)$  is well defined and fulfills

$$\|\gamma v\|_{L^p(\partial\Omega)} \leq C \|v\|_{W_+^1(\Omega)}^{1-\frac{1}{p}} \cdot \|v\|_{W_+^1(\Omega)}^{\frac{1}{p}}$$

Using  $\gamma$  we can define Hilbert spaces fulfilling boundary conditions in the sense of traces,  $H_0^1(\Omega) = \{v \in H^1(\Omega) : \gamma v|_{\partial\Omega} = 0\}$

In practice we often neglect  $\gamma v|_{\partial\Omega} = v|_{\partial\Omega}$ .

## ⑥ Interpolation

Since functions in  $H^1(\Omega)$  are not continuous in 2D & 3D we can not apply the nodal interpolant.

### The Element Interpolant

Let  $\{N_i\}_{i=1}^n$  be the nodes in a mesh  $K$  with corresponding hat functions  $\{\varphi_i\}_{i=1}^n$

Let  $w(N_i)$  be the patch of all elements

that share the node  $N_i$



Let  $P_i v \in P_1(w(N_i))$  be defined as solution to

$$\int_{w(N_i)} P_i v \cdot w dx = \int_{w(N_i)} v \cdot w dx \quad \forall w \in P_1(w(N_i))$$

i.e. the  $L^2$ -projection onto piecewise linear,

then  $\Pi_C v = \sum_{i=1}^n P_i v(N_i) \varphi_i$ .

It holds  $\|v - \Pi_C v\|_{H^k(\Omega)} \leq h^{1-n} \|v\|_{H^k(w(\Omega))}$

$m=0,1, n \leq k \leq 2, \forall k \in K$ .

Here  $\|v\|_{H^k(\Omega)} = \left( \sum_{|\alpha|=k} \|D^\alpha v\|_{L^2(\Omega)}^2 \right)^{1/2}$  and

$$\|v\|_{H^k(w)} = \left( \sum_{|\alpha| \leq k} \|D^\alpha v\|_{L^2(w)}^2 \right)^{1/2}$$

Note that  $\Pi_C$  is well defined for functions in  $L^2(\Omega)$ .

## 7 Weak form and existence of Solution

Let  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$  be a bilinear form  
i.e linear in both arguments and let  
 $(\cdot) : V \rightarrow \mathbb{R}$  be a linear functional.

We say that  $a$  is coercive if  $m\|v\|_V^2 \leq a(v, v) \forall v \in V$   
continuous if  $a(v, v) \leq C \|v\|_V \|v\|_V \forall v \in V$   
and that  $\ell$  is continuous if  $\ell(v) = C \|v\|_V \forall v \in V$

We want to study existence of solution to  
the weak form if:  $a(u, v) = \ell(v) \quad \forall v \in V$ .

First we prove the Riesz representation theorem.

### Theorem 7.1

Let  $V$  be a Hilbert space with inner product  $\langle \cdot, \cdot \rangle$ . Every cont. linear functional  $(\cdot) : V \rightarrow \mathbb{R}$  can be uniquely represented as  
 $\ell(v) = \langle u, v \rangle$  for some  $u \in V$ .

### Proof.

If  $\ell(\cdot) = 0$  then  $u = 0$  so we assume  $\ell(\cdot) \neq 0$ .

Let  $N = \{v \in V : \ell(v) = 0\}$  and let

$N^\perp = \{v \in V : (v, w) = 0 \quad \forall w \in N\}$ . Then  $V = N + N^\perp$ .

⑧ Pick  $w \in N^+ \Rightarrow ((w) \neq 0$ . For any  $v \in V$

$$((v - \frac{((v))}{((w))} w) = 0 \Rightarrow v - \frac{((v))}{((w))} w \in N$$

$$\Rightarrow (v - \frac{((v))}{((w))} w, w) = 0 \Rightarrow ((v) = \frac{((w))}{\|w\|^2} (v, v)$$

$$\Rightarrow u = \frac{((w))}{\|w\|^2} w. \text{ Uniqueness is given by}$$

contradiction. Assume  $u_1 \neq u_2$  s.t.  $((v) = (u_1, v)$   
 $\& ((v) = (u_2, v) \Rightarrow (u_2 - u_1, v) = 0 \quad \forall v \in V$ . Now  
 $(\text{let } v = u_2 - u_1 \Rightarrow \|u_2 - u_1\|^2 = 0 \Rightarrow u_1 = u_2)$ .

Ex: Consider  $V = H^1(\Omega)$  and the  
problem  $\begin{cases} -\nabla \cdot (\gamma \nabla u) + \gamma u = f & \text{in } \Omega \\ n \cdot \nabla u = 0 & \text{on } \partial \Omega \end{cases}$   
 $\gamma, \gamma > 0$ . On weak form we have  
 $\exists u: a(u, v) = (\gamma \nabla u, \nabla v) + (\gamma u, v) = (f, v) \quad \forall v \in V$   
Now  $a(u, v)$  defines a scalar product and  
 $((v) = (f, v)$  is a linear functional  
 $\Rightarrow \exists u$  s.t.  $a(u, v) = ((v)) \quad \forall v \in V$   
We have proven existence of unique soln.  
in  $H^1(\Omega)$ .

For non-symmetric problems we need the  
Lax-Milgram Lemma.

## ⑨ Lax-Milgram

Assume  $a(\cdot, \cdot)$  is coercive, continuous and that  $C$  is continuous. Then there exists a unique solution to: find  $u \in V$  s.t  $a(u, v) = (f, v)$  for  $v \in V$ ,  $V$  being a Hilbert space.

Ex: General second order elliptic problem

$$\text{Consider } (*) \begin{cases} Lu = f \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

$$\text{with } Lu = \sum_{i,j=1}^d -\frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + \sum_{i=1}^d b_i \frac{\partial u}{\partial x_i} + cu$$

or in compact form  $Lu = -\nabla \cdot (a \nabla u) + b \cdot \nabla u + cu$   
with  $a$  as  $d \times d$  matrix,  $b$  as  $d \times 1$  vector.

$$\text{We assume } a_0 \sum_{i=1}^d x_i^2 \leq \sum_{i,j=1}^d a_{ij} x_i x_j \quad \text{for any } x \in \Omega \quad (x_i, x_j \in \mathbb{R})$$

and that  $a_{ij} = a_{ji}$ . Furthermore we assume

$$\sum_{i=1}^d c - 0.5 \sum_{i=1}^d \frac{\partial b_i}{\partial x_i} > c_0 \quad \text{for any } x \in \Omega.$$

Let  $V = H_0^1(\Omega)$  and multiply  $(*)$  with test function and integrate  $\Rightarrow$  find  $u \in V$  s.t

$$a(u, v) = (a \nabla u, \nabla v) + (b \cdot \nabla u, v) + (cu, v) = (f, v) = ((v)).$$

In order to guarantee a weak solution we need to show that  $a$  is coercive and bounded and that  $(i) \Leftrightarrow$  cont.

$$\textcircled{10} \quad \text{Cont. } a(u, v) \leq \|a\|_{L^\infty(\Omega)} \cdot \|\nabla u\|_{L^2(\Omega)} \cdot \|\nabla v\|_{L^2(\Omega)} + \|b\|_{L^\infty(\Omega)} \cdot \|\nabla u\|_{L^2(\Omega)} \|v\|_V$$

$$+ \|c\|_{L^\infty(\Omega)} \cdot \|u\|_{L^2(\Omega)} \cdot \|v\|_{L^2(\Omega)} \leq C_a \|\nabla u\|_{L^2(\Omega)} \cdot \|\nabla v\|_{L^2(\Omega)}$$

Using the Poincaré-Friedrich inequality

$$\|v\|_{L^2(\Omega)} \leq C \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in V_0(\Omega).$$

Coer. Remember  $u|_{\partial\Omega} = 0$  we have using Gauss.

$$\begin{aligned} 0 &= (\mathbf{n} \cdot b u^2, 1)_{L^2(\Omega)} = (\nabla \cdot (b u^2), 1)_{L^2(\Omega)} = \\ &= ((\nabla \cdot b) u, u) + 2(b \cdot \nabla u, u) \Rightarrow (b \cdot \nabla u, u) = -\frac{1}{2} (\nabla \cdot b) u, u \end{aligned}$$

$$\text{We get } a(v, v) \geq a_0 \|\nabla v\|^2 - \frac{1}{2} (\nabla \cdot b) u, u + (c u, u)$$

$$\geq a_0 \|\nabla v\|_{L^2(\Omega)}^2 + C_0 \|v\|_{L^2(\Omega)}^2 \geq m \|v\|_V^2$$

$$\text{cont. } L(v) \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq C_L \|v\|_V.$$

Existence of unique solution follows by using LM lemma.

From here we assume the conditions of LM to hold.

### The finite element method

Let  $V_h \subset V$  be a discrete subspace spanned by shape functions  $V_h = \text{span}(\{\ell_i\}_{i=1}^n)$

FEM: find  $u_h \in V_h$  s.t.  $a(u_h, v) = L(v) \quad \forall v \in V_h$ .

(11) In order to compute  $u_h$  we expand  $u_h$  in the basis and test with  $v = \varphi_i$ . We get

$$a(u_h, \varphi_i) = ((\varphi_i)), i=1, 2, \dots, n$$

We note that  $u_h = \sum_{j=1}^n \xi_j \varphi_j$  and get

$$b_i = ((\varphi_i)) = \sum_{j=1}^n \xi_j a(\varphi_j, \varphi_i) = \sum_{j=1}^n A_{ij} \xi_j, i=1, 2, \dots, n$$

or in matrix form  $A \xi = b$ .

### Galerkin orthogonality

Let  $e = u - u_h$ .

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

### Cea's Lemma

$$\begin{aligned} m \|e\|_V^2 &\leq a(e, e) = a(e, u - u_h) = a(e, u - v) + a(e, v - u_h) \\ &\leq c_a \|e\|_V \cdot \|u - v\|_V \Rightarrow \|e\|_V \leq \frac{c_a}{m} \|u - v\|_V \quad \forall v \in V_h. \end{aligned}$$

The FE solution is optimal up to a constant.

### Convergence:

$$\text{Assume } \|u - \pi_h u\|_V \leq C h \|u\|_{H^2(\Omega)}$$

Then we immediately get

$$\|e\|_V \leq C h \|u\|_{H^2(\Omega)}$$

i.e. convergence as  $h \rightarrow 0$ .

## ⑫ Chapter 8 The finite element

- \* Definition of a finite element
- \* Lagrangean elements
- \* Isoparametric map
- \* Numerical quadrature
- \* Exotic elements

### Formal definition of a finite element

A finite element consists of the triplet

- (i) A polygon  $K \subset \mathbb{R}^d$
- (ii) A polynomial function space  $P$  on  $K$
- (iii) A set of  $n = \dim(P)$  linear functionals  $L_i(\cdot)$   
 $i=1, 2, \dots, n$  defining the degrees of freedom.

We equip  $P$  with a basis  $\{S_j\}_{j=1}^n$ . The basis functions are generally called shape functions.

Unisolvency is a necessary compatibility condition for  $L_i(\cdot)$ ,  $P$ , and  $K$ . By definition it is equivalent to  $L_i(v) = 0 \Leftrightarrow v = 0 \quad \forall v \in P \quad \forall i=1, \dots, n$ .

(13) The Shape functions are determined from the  $n$  equations  $L_i(S_j) = \delta_{ij}$ ,  $i,j=1,\dots,n$

The choice of functionals also specify behavior between polygons  $K$ . If we want a continuous finite element space we need to pick  $L_i$  so that the resulting shape functions are continuous.

Ex Lagrangean elements.

The most common element is the Lagrange.

Then  $L_i(v) = v(N_i)$ ,  $i=1,2,\dots,n$ , where  $N_i$  are node points.

For  $d=2$  with  $P=P_1(K)$  on a triangle  $K$

these nodes are triangle vertices.

$\{S_i\}_{i=1}^3$  are the hat functions.

$\Rightarrow$  Continuous.

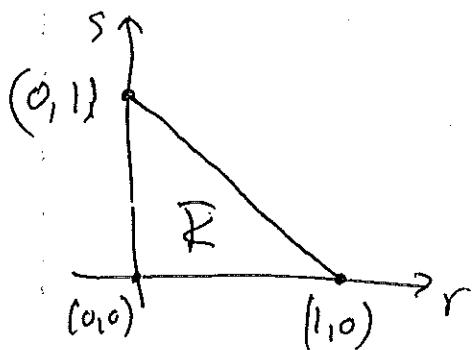
Lagrange elements are  $C^0$  (but not  $C^1$ ).

However, there are  $C^1$  elements which can be used to approximate functions in  $H^2(\Omega)$ .

(14)

## Shape functions for Lagrangean elements

Let  $\bar{K} = \{(r,s) : 0 < r, s < 1, r+s < 1\}$



Linear

We consider  $P = P_1(K)$  (linear polynomials)

We have  $L_1(v) = v(0,0)$ ;  $L_2(v) = v(1,0)$ ,  $L_3(v) = v(0,1)$

We let  $P_1(\mathbb{R}) = \text{span}(\{1, r, s\})$

$\Rightarrow S_1 = c_1 + c_2 r + c_3 s$ ,  $c_i \in \mathbb{R}$  and so on

Since  $L_i(S_1) = \delta_{i,1}$  i.e.

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} L_1(1) & L_1(r) & L_1(s) \\ L_2(1) & L_2(r) & L_2(s) \\ L_3(1) & L_3(r) & L_3(s) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = Vc$$

$$c = V^{-1}e_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \Rightarrow S_1 = 1 - r - s.$$

In the same way  $S_2 = r$ ,  $S_3 = s$ .

## Quadratic

Now we let  $P = P_2(K)$ , quadratic polynomials

We let  $L_1(v) = v(0,0)$ ,  $L_2(v) = v(1,0)$ ,  $L_3(v) = v(0,1)$

$L_4(v) = v(0.5, 0.5)$ ,  $L_5(v) = v(0, 0.5)$ ,  $L_6(v) = v(0.5, 0)$

(15) Let  $P_2(E) = \text{Span}(\{1, r, s, r^2, rs, s^2\})$

Again  $L_i(s_1) = c_{i1} \Rightarrow$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ L_i(1) & L_i(r) & L_i(s) & L_i(r^2) & L_i(rs) & L_i(s^2) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = Vc$$

$$\Rightarrow c = [1, -3, -4, 2, 4, 2]$$

$$S_1 = 1 - 3r - 3s + 2r^2 + 4rs + 2s^2$$

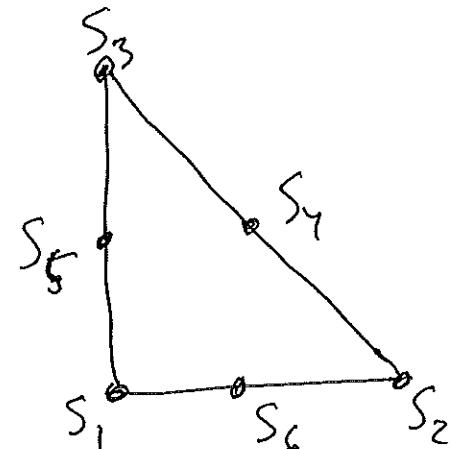
$$S_2 = 2r^2 - r$$

$$S_3 = 2s^2 - s$$

$$S_4 = 4rs$$

$$S_5 = 4s - 4rs - 4s^2$$

$$S_6 = 4r - 4r^2 - 4rs$$



Alternatively the quadratic functions can be expressed in terms of the linear  $\ell_1, \ell_2, \ell_3$

$$S_1 = 2\ell_1^2 - \ell_1$$

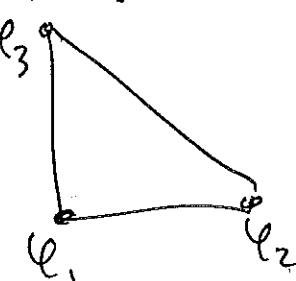
$$S_2 = 2\ell_2^2 - \ell_2$$

$$S_3 = 2\ell_3^2 - \ell_3$$

$$S_4 = 4\ell_1\ell_2$$

$$S_5 = 4\ell_1\ell_3$$

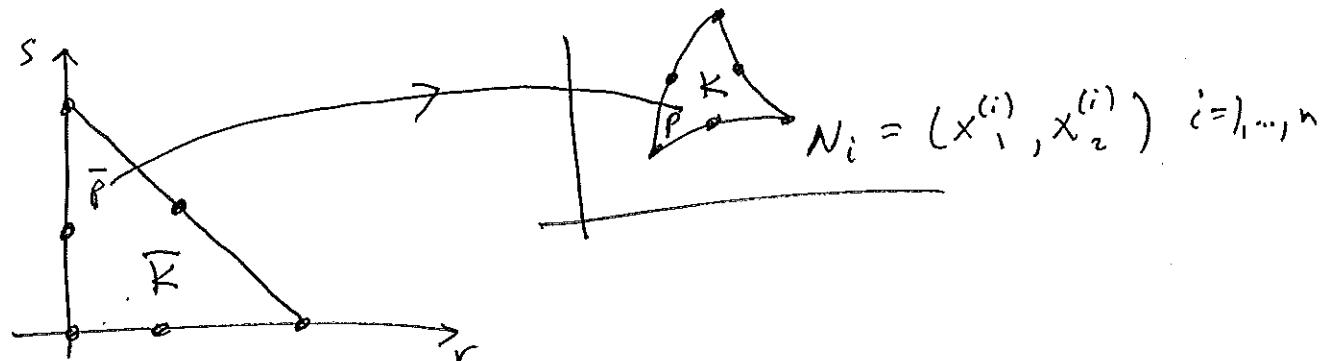
$$S_6 = 4\ell_2\ell_3$$



16

## The isoparametric map

We would like to be able to work with elements with curved boundaries. We would like to map a general element in the assembly process to a reference element.



## The isoparametric map

$$X_1(r, s) = \sum_{i=1}^n x_1^{(i)} S_i(r, s)$$

$$X_2(r, s) = \sum_{i=1}^n x_2^{(i)} S_i(r, s)$$

The map implies that  $\partial K$  is curved if the nodes  $N_i$  on an edge are not on a straight line.

The order of the curved edge is given by the order of the shape functions  $\{S_i\}_{i=1}^n$ .

Any FE function  $v$  on  $K$  is expressed as

$$v(r, s) = \sum_{i=1}^n v_i S_i(r, s), \quad v_i = v(N_i).$$

(17) In the assembly of the stiffness matrix we need to evaluate derivatives.

$$\begin{bmatrix} \frac{\partial v}{\partial x_1} \\ \frac{\partial v}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial v}{\partial r} \cdot \frac{\partial r}{\partial x_1} + \frac{\partial v}{\partial s} \cdot \frac{\partial s}{\partial x_1} \\ \frac{\partial v}{\partial r} \cdot \frac{\partial r}{\partial x_2} + \frac{\partial v}{\partial s} \cdot \frac{\partial s}{\partial x_2} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\partial r}{\partial x_1} & \frac{\partial s}{\partial x_1} \\ \frac{\partial r}{\partial x_2} & \frac{\partial s}{\partial x_2} \end{bmatrix}}_{J^{-1}} \begin{bmatrix} \frac{\partial v}{\partial r} \\ \frac{\partial v}{\partial s} \end{bmatrix}$$

We have  $J = \begin{bmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_2}{\partial r} \\ \frac{\partial x_1}{\partial s} & \frac{\partial x_2}{\partial s} \end{bmatrix}$  is the Jacobian matrix

$$J_{11} = \frac{\partial x_1}{\partial r} = \sum_{i=1}^n \frac{\partial s_i}{\partial r} x_1^{(i)} \quad \text{and so on}$$

Partial derivatives of  $v$  can therefore be computed by inverting the Jacobian  $J$ .

### Numerical quadrature

Given the map  $(x_1, x_2) \rightarrow (r, s)$  we have

$$\int_K f(x_1, x_2) dx = \int_E f(r, s) \det(J(r, s)) dr ds ,$$

which is the classical change of variable formula.

The RHS will be approximated using numerical quadrature

$$\int_E f(r, s) \det(J(r, s)) dr ds \approx \sum_{q=1}^n w_q f(r_q, s_q) \det(J(r_q, s_q))$$

Gauss points are typically used, optimal for polynomials.

(18)

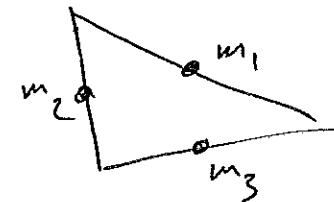
## Exotic elements

### The Crouzeix-Raviart element

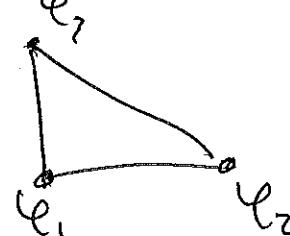
It's a linear element, only continuous at mid-points.

$$P = P_1(\mathbf{k})$$

$$\mathbf{L}_i(\mathbf{r}) = \mathbf{v}(m_i), \quad i=1,2,3$$



CR elements can be expressed in terms of linear Lagrangean elements



$$S_1^{CR} = -\mathbf{l}_1 + \mathbf{l}_2 + \mathbf{l}_3$$

$$S_2^{CR} = \mathbf{l}_1 - \mathbf{l}_2 + \mathbf{l}_3$$

$$S_3^{CR} = \mathbf{l}_1 + \mathbf{l}_2 - \mathbf{l}_3$$

CR elements are used in fluid mechanics.

Note that  $S_i^{CR} \notin H^1$

### The Raviart-Thomas element

The RT element is a vector valued element approximating  $H(\text{div}; \Omega) = \{\mathbf{v} \in (\mathbf{L}^2(\Omega))^2 : \nabla \cdot \mathbf{v} \in L^2(\Omega)\}$  rather than e.g.  $H^1(\Omega)$ .

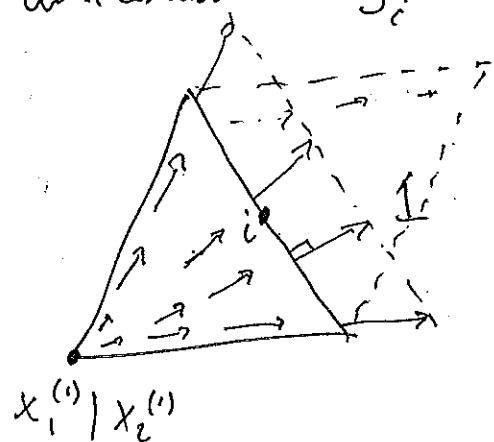
Green's formula can be used to show that functions in  $H(\text{div}; \Omega)$  are normal continuous.

$$\text{Let } P = [P_0(k)]^2 + [x_1, x_2]^T P_0(k) \Rightarrow v = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} + b \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\text{Furthermore } L_i(v) = \frac{(n^{E_i}, v)_{L^2(E_i)}}{|E_i|}, i=1,2,3$$

where  $n^{E_i}$  is unit normal on edge  $E_i$  on  $K$ .

$$\text{Furthermore } S_i^{RT_0} = \frac{|E_i|}{2|K|} \begin{bmatrix} x_1 - x_1^{(i)} \\ x_2 - x_2^{(i)} \end{bmatrix} \quad i=1,2,3$$



### The Nedelec element

A vector valued element approximating

$$H(\text{curl}; \Omega) = \{ v \in [L^2(\Omega)]^2 : \nabla \times v \in L^2(\Omega) \}$$

$$\text{Let } P = [P_0(k)]^2 + \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix} P_0(k)$$

$$\text{Let } L_i(v) = \frac{(t^{E_i}, v)_{L^2(E_i)}}{|E_i|}, i=1,2,3$$

where  $t^{E_i}$  is a unit tangent vector on edge  $E_i$ .  
(counter clockwise)

$S_i^{NO} = |E_i| (\psi_i \nabla \psi_k - \psi_k \nabla \psi_i)$ , where  $\psi_i$  are the usual hat functions with cyclic permutation of the indices  $\{i, j, k\}$  over  $\{1, 2, 3\}$ .

Nedelec elements are tangential continuous.



## ⑩ Chapter 9 Non-linear problems

- \* Repetition of iterative methods
- \* Non-linear Poisson equation
- \* The bistable equation
- \* Numerical approximation of the Jacobian

### Iterative methods

We first consider Picard or fixed point iteration

$$x = g(x)$$

Given  $x^{(0)}$ , a first guess, we let  $x^{(k+1)} = g(x^{(k)})$

(if  $\|g(x) - g(y)\| \leq L \|x - y\|$  with  $L < 1$

$g$  is a contraction map. In that case we have

$$\|x^{(k+1)} - x\| = \|g(x^{(k)}) - g(x)\| \leq L \|x^{(k)} - x\| \leq L^{k+1} \|x^{(0)} - x\|$$

$\rightarrow 0$  as  $k \rightarrow \infty$ .

### Newton iteration

Now consider  $g(x) = 0$ . Let  $x = x^0 + \delta x$ , where  $x^0$  is initial guess. Using Taylor expansion we get

$$0 = g(x) = g(x^0 + \delta x) = g(x^0) + g'(x^0) \delta x + O(\delta x^2)$$

$$\Rightarrow \delta x = -g'(x^0)^{-1} g(x^0), \quad x^{(k+1)} = x^k - g'(x^k)^{-1} g(x^k)$$

(21) It can be shown that

$$\|x^{(k+1)} - x\| \leq C \|x^{(k)} - x\|^2 \text{ for } x^{(k)} \text{ sufficiently close to } x.$$

### The non-linear Poisson equation

$$\text{Consider } \begin{cases} -\nabla \cdot (a(u)\nabla u) = f \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

We multiply the equation by a test function  $v \in V = H_0^1(\Omega)$  and use Green's formula: find  $u \in V$

$$: (a(u)\nabla u, \nabla v) = (f, v) \quad \forall v \in V.$$

We now use Newton's method and let  $u = u^0 + \delta u$

$$\Rightarrow (a(u^0 + \delta u)\nabla(u^0 + \delta u), \nabla v) = (f, v), \quad \forall v \in V$$

$$a(u^0 + \delta u) = a(u^0) + a'_u(u^0) \frac{\delta u}{\delta u} + O((\delta u)^2)$$

We drop quadratic terms to get: find  $\delta u \in V$

$$: (a(u^0)\nabla \delta u + a'_u(u^0)\nabla u^0 \cdot \delta u, \nabla v) = (f, v) - (a(u^0)\nabla u^0, \nabla v) \quad \forall v \in V$$

Given  $\delta u$  we let  $u' = u^0 + \delta u$ .

### Finite element approximation

Let  $\mathcal{M} = \{K\}$  be a mesh of  $\Omega$  and  $V_h \subset V$  the space of cont. piecewise linear basis functions.

(2)

Find  $\delta u_n \in V_h$  s.t.

$$(a(u^0) \nabla \delta u_n + a'_h(u_h^0) \nabla u_h^0 \cdot \delta u_n, \nabla v) = (f, v) - (a(u^0) \nabla u_h^0, \nabla v)$$

$v \in V_h$ .

We assume  $u^0 \in V_h$  which is a sensible assumption.

$$\text{We now let } \delta u_n = \sum_{j=1}^n d_j \varphi_j, \quad \text{span}(\{\varphi_j\}_{j=1}^n) = V_h$$

$$\sum_{j=1}^n d_j (a(u_h^0) \nabla \varphi_j + a'_h(u_h^0) \varphi_j \nabla u_h^0, \nabla \varphi_i) = (f, \varphi_i) - (a(u_h^0) \nabla u_h^0, \nabla \varphi_i) \quad i=1, \dots, n$$

$$\text{We let } J_{ij} = (a(u_h^0) \nabla \varphi_j, \nabla \varphi_i) + (a'_h(u_h^0) \nabla u_h^0 \cdot \varphi_j, \nabla \varphi_i)$$

$$\text{and } r_i = (f, \varphi_i) - (a(u_h^0) \nabla u_h^0, \nabla \varphi_i), \quad i=1, \dots, n$$

$$\underbrace{\mathbf{J} \mathbf{d}}_{} = \mathbf{r}$$

An algorithm can now be derived with  
stopping criteria  $\|\mathbf{J} \mathbf{d}_k\| < \varepsilon$ .

$$u_h^{(k+1)} = u_h^{(k)} + \delta u_h^{(k)}$$

If we replace  $\mathbf{J}$  by  $A_{ij} = (a(u_h^0) \nabla \varphi_j, \nabla \varphi_i)$   
in Newton we get Picard iteration, see p. 233.

(23) The bistable equation

$$\left\{ \begin{array}{ll} \text{Consider } \begin{cases} u_t - \varepsilon \Delta u = u - u^3 & \text{in } \Omega \times (0, T] \\ n \cdot \nabla u = 0 & \text{in } \partial \Omega \times (0, T] \\ u = u_0 & \text{in } \Omega, t=0 \end{cases} \end{array} \right.$$

Weak form find  $u \in H^1(\Omega)$  s.t

$$(u_t, v) + \varepsilon (\nabla u, \nabla v) = (f(u), v) \quad \forall v \in H^1(\Omega)$$

with  $f(u) = u - u^3$ .

FEM let  $u_n = \sum_{j=1}^n \varepsilon_j \psi_j$ ,  $j=1, \dots, n$ ,  $v = \psi_i$ :

$$M \dot{z} + A z = b(z), \quad M_{ij} = (\psi_j, \psi_i), \quad A_{ij} = \varepsilon (\nabla \psi_j, \nabla \psi_i)$$

$$b_i(z) = (f(u_n), \psi_i)$$

Time discretization

$$\text{BE: } M \frac{z_l - z_{l-1}}{k_l} + A z_l = b(z_l) \quad \text{or}$$

$$(M + k_l A) z_l = M z_{l-1} + k_l b(z_l),$$

Picard Iteration

$$z_l^{(t)} = (M + k_l A)^{-1} (M z_{l-1} + k_l b(z_l^{(t-1)}))$$

converged sol from previous step

$z_l^{(0)}$  is naturally chosen to be  $z_{l-1}$ .

④ Numerical approximation of the Jacobian

Let  $J_{ij} = \frac{\partial r_i}{\partial z_j}$   $i, j = 1, \dots, n$

let  $J_{:,i}$  be column  $i$  of  $J$ . Then

$$J_{:,i} = \frac{r(\bar{z} + \varepsilon e_i) - r(\bar{z})}{\varepsilon} \quad \text{for some } \varepsilon > 0$$

This is a simple but expensive approx. of  $J$ .

Broyden's method

In order to evaluate  $r$ ,  $n$  times the following formula can be applied (approximation)

$$J^{(k)} = J^{(k-1)} + \frac{r^{(k)} - r^{(k-1)} - J^{(k-1)}(\bar{z}^{(k)} - \bar{z}^{(k-1)})}{\|\bar{z}^{(k)} - \bar{z}^{(k-1)}\|^2} (\bar{z}^{(k)} - \bar{z}^{(k-1)})^T$$

## (25) Chapter 10 Transport problems

- \* The transport equation
- \* Stabilization
- \* Galerkin Least Squares
- \* Error analysis

### Transport problem

Consider  $\begin{cases} -\varepsilon \Delta u + b \cdot \nabla u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$ .

To get a well posed problem we assume  $\nabla \cdot b = 0$ .

This is often referred to as the convection-diffusion problem, where  $\varepsilon$  is diffusion coefficient and  $b$  is convection coefficient.

### Weak form:

Find  $u \in V = H_0^1(\Omega)$  s.t

$$a(u, v) = (\varepsilon \nabla u, \nabla v) + (b \cdot \nabla u, v) = (f, v) \quad \forall v \in V$$

### FEM:

Let  $V_h \subset V$ : Find  $u_h \in V_h$  s.t  $a(u_h, v) = (f, v) \quad \forall v \in V_h$

with  $u_h = \sum_{j=1}^n z_j \varphi_j$  we get  $(A + C)z = b$  with

$$A_{ij} = c(\nabla \varphi_j, \nabla \varphi_i) \quad C_{ij} = (b \cdot \nabla \varphi_j, \varphi_i) \quad b_i = (f, \varphi_i)$$

i.e. non-symmetric matrix  $(A + C)$ .

⑯

## Stabilization

$$\begin{aligned}
 \text{Note that } (\mathbf{b} \cdot \nabla u, u) &= (\nabla u, \mathbf{b} u) = - (u, \nabla \cdot (\mathbf{b} u)) + (u, \mathbf{n} \cdot \mathbf{b} u) \\
 &= - (u, \mathbf{b} \cdot \nabla u) - (u, (\nabla \cdot \mathbf{b}) u) + (u, \mathbf{n} \cdot \mathbf{b} u)_{\partial \Omega} = \\
 &= - (u, \mathbf{b} \cdot \nabla u) \Rightarrow (u, \mathbf{b} \cdot \nabla u) = 0.
 \end{aligned}$$

$$\begin{aligned}
 \varepsilon \|\nabla u\|_{L^2(\Omega)}^2 &= \varepsilon (\nabla u, \nabla u) + (\mathbf{b} \cdot \nabla u, u) = (f, u) \leq C \|f\|_{L^2(\Omega)} \|\nabla u\|_{L^2(\Omega)} \\
 \Rightarrow \|\nabla u\|_{L^2(\Omega)} &\leq C \varepsilon^{-1} \|f\|_{L^2(\Omega)}
 \end{aligned}$$

As  $\varepsilon \rightarrow 0$  we loose control over gradients.

This leads to boundary layers.

FEM have big problems with layers

## Least Squares stabilization

Consider  $Lu = f$

A least squares approximation would be

$$L^* L u = L^* f \text{ or on weak form find } u \in V:$$

$$(*) (Lu, Lv) = (f, Lv) \quad \forall v \in V$$

The Galerkin LS method is a combination of standard FEM and (\*) namely, find  $u \in V$  s.t

$$(Lu, v + \delta Lv) = (f, v + \delta Lv) \quad \forall v \in V$$

where  $\delta$  is a given parameter.

## ② GLS (Galerkin Leastsquares)

Recall  $L = -\varepsilon \Delta + b \cdot \nabla$

Find  $u \in V_h$  s.t.

$$a_h(u_h, v) = l_h(v) \quad \forall v \in V_h \text{ where}$$

$$a_h(u, v) = a(u, v) + \delta \sum_{k \in \Omega} (-\varepsilon \Delta u + b \cdot \nabla u, -\varepsilon \Delta v + b \cdot \nabla v)_{L^2(k)}$$

$$l_h(v) = ((v) + \delta \sum_{k \in \Omega} (f, -\varepsilon \Delta v + b \cdot \nabla v)_{L^2(k)})$$

We integrate elementwise since  $\Delta v$  is not globally defined for  $v \in V_h$ .

For piecewise linear basis functions we have  $\Delta v|_K = 0$

$$\text{we get } a_h(u, v) = \varepsilon(\nabla u, \nabla v) + (b \cdot \nabla u, v) + \delta(b \cdot \nabla u, b \cdot \nabla v)$$

$$l_h(v) = (f, v) + \delta(f, b \cdot \nabla v)$$

$a_h(v, v) = \varepsilon \|\nabla v\|^2 + \delta \|b \cdot \nabla v\|^2$  i.e. we directly have coercivity in the  $\|v\| = a_h(v, v)^{1/2}$  norm.

However we need to keep  $\delta$  big enough for small  $\varepsilon$ . A good choice is

$$\delta = \begin{cases} ch^2, & \varepsilon > h \\ ch/\|b\|_{L^2(h)}, & \varepsilon < h \end{cases}$$

For Continuity of  $a_h$  and  $l_h$  see Exe 10.7.

## ⑦8 Apriori error estimate

We note that  $a_h(u, v) = (u, v) \quad \forall v \in V_h$

using Galerkin orthogonality we get

$$a_h(u - u_h, v) = 0 \quad \forall v \in V_h.$$

$$\text{Now let } e = u - u_h = u - \pi u + (\pi u - u_h)$$

$$\text{We have } \|u - \pi u\|^2 \leq \varepsilon h^2 \|u\|_{H^2}^2 + \delta \|b\|_{L^\infty}^2 \cdot h^2 \|u\|_{H^2}^2$$

$$\text{and for } \varepsilon < h, \delta = C_h / \|b\|_{L^\infty} \text{ we get } \|u - \pi u\| \leq C_b h^{3/2} \|u\|_{H^2}$$

It remains to study  $\|\pi u - u_h\|$ .

$$\begin{aligned} \|\pi u - u_h\| &= a_h(u - \pi u, u_h - \pi u) = (\varepsilon \nabla u - \pi u, \nabla u_h - \pi u) + (b \cdot \nabla u - \pi u, u_h - \pi u) \\ &\quad + \delta \sum_{e \in E} (-\varepsilon \Delta u + b \cdot \nabla u - \pi u, b \cdot \nabla u_h - \pi u)_{L^2(e)} = \textcircled{I} + \textcircled{II} + \textcircled{III} \end{aligned}$$

$$\textcircled{I} \leq \|u - \pi u\| \cdot \|u_h - \pi u\| \leq C h^{3/2} \|u\|_{H^2} \cdot \|u_h - \pi u\|$$

$$\begin{aligned} \textcircled{II} &\leq (b \cdot \nabla u - \pi u, u_h - \pi u) = -(u - \pi u, b \cdot \nabla(u_h - \pi u)) \leq \|u - \pi u\| \cdot \|b \cdot \nabla u - \pi u\| \\ &\leq h^2 \|u\|_{H^2} \cdot \|b\|_{L^\infty} \cdot \delta^{1/2} \cdot \|u_h - \pi u\| \leq C_b h^{3/2} \|u\|_{H^2} \cdot \|u_h - \pi u\| \end{aligned}$$

$$\textcircled{III} \leq \delta^{1/2} \varepsilon \|u\|_{H^2} \cdot \|u_h - \pi u\| + \delta^{1/2} \|b\|_{L^\infty} \cdot h \|u\|_{H^2} \cdot \|u_h - \pi u\|$$

$$\leq C_b h^{3/2} \|u\|_{H^2} \cdot \|u_h - \pi u\| + C_b h^{3/2} \|b\|_{L^\infty} \cdot \|u_h - \pi u\|$$

$$\therefore \|\pi u - u_h\| \leq C_b h^{3/2} \|u\|_{H^2} \quad \text{i.e. } \|u - u_h\| \leq C_b h^{3/2} \|u\|_{H^2}$$