Chapter 2

Variational Formulation

2.1 Boundary Value problems

Example 2.1.1 (One dim'l problem)**.**

$$-u'' = f \text{ on } I \equiv (0,1), \text{ with B.C. } u(0) = u(1) = 0.$$

Multiply a test function $v \in H_0^1(I)$ and integrate

$$\begin{aligned}(-u'', v) &= -\int_0^1 u'' v dx \\ &= -[u'v]_0^1 + \int_0^1 u'v' dx = \int_0^1 fv dx.\end{aligned}$$

Thus we have

$$(u', v') = (f, v), \quad v \in V = H_0^1(I).$$

We will replace the space $H_0^1(I)$ by a finite dimensional space $S_h(I)$ of continuous, piecewise linear functions on I.

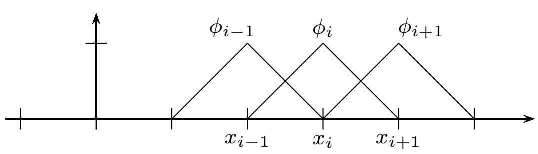


Figure 2.1: Basis in one-dimension

Let $h_i = x_i - x_{i-1},\ I_i = [x_{i-1}, x_i]$. Let $u_h = \sum_{j=1}^{n-1} c_j\phi_j$, where $\phi_i \in S_h(I)$ is the Lagrange basis function associated with the note x_i. Then substituting into (**??**), we obtain

$$\sum_{j=1}^{n-1}\int_0^1 c_j\phi_j'(x)\phi_i'(x)dx = \int_0^1 f(x)\phi_i(x)dx,\ \ i = 1, 2, \cdots, n-1.$$

Hence we obtain the matrix equation

$$A_h u_h = f_h,$$

where

$$(A_h)_{ij} = A_{ij} = \int_0^1 \phi_j'(x)\phi_i'(x)dx, \quad f_i = \frac{h_{i-1}+h_i}{2}f(x_i).$$

If we use a uniform spacing, then a typical row of A_h is $[\cdots, 0, -1, 2, -1, 0\cdots]$. This matrix is the same as the one from FDM (up to the factor of h^2).

A typical row of A_h is

$$\left(\cdots, 0, -\frac{1}{h_{i-1}}, \frac{1}{h_{i-1}} + \frac{1}{h_i}, -\frac{1}{h_i}, 0, \cdots\right)$$

Example 2.1.2. $k = 2$ quadratic basis. The basis function associate with the nodes are

$$\phi_i(x) = \begin{cases} 1 + 3\left(\dfrac{x-x_i}{h_i}\right) + 2\left(\dfrac{x-x_i}{h_i}\right)^2 \text{ if } x_{i-1} \le x < x_i \\ 1 - 3\left(\dfrac{x-x_i}{h_{i+1}}\right) + 2\left(\dfrac{x-x_i}{h_{i+1}}\right)^2 \text{ if } x_i \le x < x_{i+1}, \\ 0 \text{ otherwise} \end{cases}, i = 0, 1, \cdots, n \tag{2.1}$$

Here $h_i = x_i - x_{i-1}$. The basis function associate with the mid pts are

$$\phi_{i-1/2}(x) = \begin{cases} 1 - 4\left(\dfrac{x - x_{i-1/2}}{h_i}\right)^2 \text{ if } x_{i-1} \le x < x_i \\ 0 \text{ otherwise} \end{cases}, i = 1, \cdots, n \tag{2.2}$$

Then

$$\phi_i(x_k) = \begin{cases} 1 \text{ if } i = k \\ 0 \text{ otherwise} \end{cases}, i, k = 0, 1/2, 1, \cdots, n-1, n-1/2, n \tag{2.3}$$

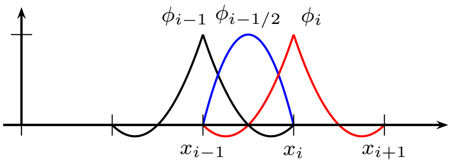


Figure 2.2: Basis in one-dimension, $k = 2$

2.1.1 The Poisson equation in $\mathbb{R}^2$

Let Ω be a bounded domain in $\mathbb{R}^2$ and $\partial\Omega$ denote its boundary.

We say a function u defined on $\overline{\Omega}$ is a classical solution of the Poisson equation with homogeneous boundary condition if $u \in C^2(\Omega), u \in C(\bar{\Omega})$ and u satisfying

$$\begin{cases} -\Delta u(x,y) &= f(x,y) \text{ for } (x,y) \in \Omega \\ u(x,y) &= 0 \text{ for } (x,y) \in \Gamma_1 \\ \frac{\partial u}{\partial \nu}(x,y) &= 0 \text{ for } (x,y) \in \Gamma_2, \end{cases} \tag{2.4}$$

where Δ is the Laplacian operator and $\Gamma_1 \cup \Gamma_2 = \partial\Omega$ and $\Gamma_1 \cap \Gamma_2$ is a set of measure zero. It is well known that for sufficiently smooth boundary, there exists a unique classical solution of (2.4) provided that Γ_1 is measurable with positive measure.

We will assume that Ω is a normal domain, i.e., it admits the application of divergence theorem:

$$\int_\Omega \frac{\partial u}{\partial x_i} dxdy = \int_{\partial\Omega} u\nu_i ds, \quad i = 1,2 \quad u \in C^1(\bar{\Omega}), \tag{2.5}$$

where ν_i are the components of unit outward normal vector to $\partial\Omega$.

Fact: Every polygonal domain or a domain with piecewise smooth boundary is a normal domain.

As a consequence of (2.5) we have the *Green's Formula.*

$$\int_\Omega v\Delta u dxdy = \int_{\partial\Omega} v\frac{\partial u}{\partial \nu} ds - \int_\Omega \nabla u \cdot \nabla v dxdy. \tag{2.6}$$

Suppose that u is a classical solution of (2.4) and that $v \in V = \{\phi \in C^\infty(\Omega) : \phi = 0 \text{ on } \Gamma_1\}$. Since $v = 0$ on Γ_1 and $\frac{\partial u}{\partial \nu} = 0$ on Γ_2, we see (2.6) yields

$$-\int_\Omega v\Delta u dxdy = \int_\Omega \nabla u \cdot \nabla v dxdy := a(u,v).$$

Hence u satisfy

$$a(u, v) = (f, v) \quad v \in V, \tag{2.7}$$

where $(f, v) = \int_\Omega fvdxdy$. $a(u, v)$ is a bilinear form defined on $H^1(\Omega)$ and is called the *Dirichlet* integral associated with the Laplace operator $-\Delta$.

The space V can be shown to be dense in $H^1_{\Gamma_1}(\Omega) = \{v \in H^1(\Omega) : v = 0$ on $\Gamma_1\}$. It is not difficult to show that (2.7) actually holds for every $v \in H^1_{\Gamma_1}(\Omega)$, i.e,

$$a(u, v) = (f, v) \quad v \in H^1_{\Gamma_1}(\Omega). \tag{2.8}$$

We now define a generalized (weak) solution $u \in H^1_{\Gamma_1}(\Omega)$ of (2.4) as a function u in $H^1_{\Gamma_1}(\Omega)$ which satisfy (2.8). We will show the existence and uniqueness of this weak solution using the Lax-Milgram theorem(Later). First, let $u, v \in H^1_{\Gamma_1}(\Omega)$, then

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v dxdy \leq \|u\|_1 \|v\|_1$$

so that condition (i) of the Lax-Milgram theorem holds. Now the second condition of the Lax-Milgram theorem holds by the Poincaré inequality. Clearly, for $f \in L^2(\Omega)$, the linear form (f, v) defines a bounded linear functional on $H^1_{\Gamma_1}(\Omega)$. Hence there exists a unique $u \in H^1_{\Gamma_1}(\Omega)$ such that (2.8) holds. Moreover $\|u\|_1 \leq c\|f\|$. More can be said about the solution if the boundary is smoother and f assures more regularity: For example, if $\partial\Omega$ is of class C^r and $f \in H^{r-2}(\Omega)$, then

$$u \in H^r(\Omega) \cap H^1_{\Gamma_1}(\Omega) \tag{2.9}$$

and

$$\|u\|_r \leq C\|f\|_{r-2}. \tag{2.10}$$

Results such as (2.9) and (2.10) are known as *elliptic regularity estimates.*

An example of FEM

Example 2.1.3 (Poisson problem-Dirichlet BC.)**.**

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega \\ u &= 0 \text{ on } \partial\Omega. \end{aligned}$$

Assume $\Omega = [0, 1]^2$ and that we divide Ω into $2n^2$ right triangles of length h. Let $S^0_h(\Omega)$ be the space of continuous, piecewise linear on each element sat-

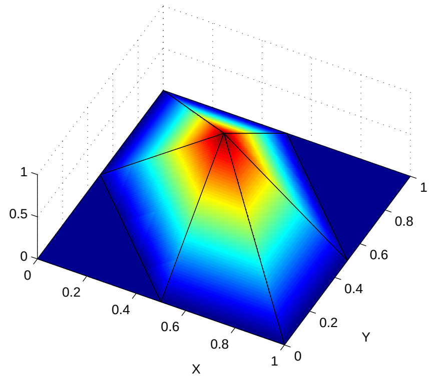


Figure 2.3: Standard nodal Lagrange local basis

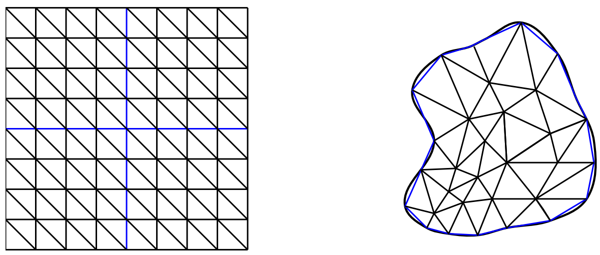


Figure 2.4: Finite element meshes

isfying zero boundary condition. Let $u_h = \sum u_j \phi_j$ where ϕ_j is the nodal(tent shape) basis function satisfying $\phi_j(x_i) = \delta_{ij}$. Then multiply ϕ_i and integrate by part to get

$$\int_\Omega \sum_j u_j \nabla\phi_j \cdot \nabla\phi_i \, dxdy = \int_\Omega f\phi_i \, dx, \text{ for each } i = 1, 2, \cdots .$$

Writing $a(\phi_i, \phi_j) = \int_\Omega \nabla\phi_j \cdot \nabla\phi_i dxdy$, we get

$$\sum_j a(\phi_i, \phi_j) u_j = (f, \phi_i).$$

In matrix form, it is

$$A\mathbf{u} = \mathbf{f}, \quad A_{ij} = a(\phi_j, \phi_i).$$

A is called the 'stiffness' matrix.

Example 2.1.4 (Neumann problem).

$$\begin{aligned} -\Delta u + u &= f \text{ in } \Omega \\ \frac{\partial u}{\partial n} &= g \text{ on } \Gamma = \partial\Omega. \end{aligned} \tag{2.11}$$

Note in this case u is unknown at the boundary. So we set $V = H^1(\Omega)$(not $H_0^1(\Omega)$!). [1]

$$(-\Delta u, v) + (u, v) = (\nabla u, \nabla v) + (u, v) - < g, v >_\Gamma = (f, v).$$

So the variational problem is: (N) Find $u \in H^1$ such that

$$a(u, v) = (f, v) + < g, v >_\Gamma, \quad v \in H^1,$$

where $a(u, v) = (\nabla u, \nabla v) + (u, v)$.

Show: If $u \in C^2$ is the solution of (N) then it is the solution of (2.11).

Proof. Let u be the solution of (N). Then for $v \in H^1$

$$\begin{aligned} a(u, v) &= (-\Delta u, v) + \int_\Gamma \frac{\partial u}{\partial n} v + (u, v) = (f, v) + < g, v >_\Gamma . \\ \int_\Omega (-\Delta u + u - f) v &= \int_\Gamma (g - \frac{\partial u}{\partial n}) v ds, \quad v \in H^1. \end{aligned}$$

Restrict to $v \in H_0^1$. Then we get

$$-\Delta u + u - f = 0 \text{ in } \Omega.$$

Hence we have

$$\int_\Gamma (g - \frac{\partial u}{\partial n}) v ds = 0, \quad v \in H^1$$

which proves $g = \frac{\partial u}{\partial n}$. The condition $\frac{\partial u}{\partial n} = g$ is called the natural boundary condition. (Look at the space V, we did not impose any condition, but we got B.C naturally from the variational formulation.) □

More general Boundary Conditions

We consider a mixed BC. i.e., on one part, the Dirichlet condition is imposed, while on the other part Neumann condition is imposed. We also consider more

[1] If $g = 0$, we have a physically insulated boundary.

general coefficients,

Example 2.1.5. Assume there exists two positive constants p_0, p_1 s.t. $0 < p_0 \leq p(x,y) \leq p_1$. Consider

$$\begin{aligned} -\nabla \cdot p\nabla u &= f \text{ in } \Omega \\ u &= 0 \text{ on } \Gamma_1 \\ \frac{\partial u}{\partial n} &= g \text{ on } \Gamma_2 := \partial\Omega \backslash \Gamma_1. \end{aligned} \tag{2.12}$$

Let $V_1 = V_{\Gamma_1} = \{v \in H^1(\Omega), v|_{\Gamma_1} = 0\}$. If $v \in V_1$,

$$\begin{aligned} (-\nabla \cdot p\nabla u, v) &= -\int_\Gamma p\frac{\partial u}{\partial n} v ds + \int_\Omega p\nabla u \cdot \nabla v dxdy \\ &= -\int_{\Gamma_2} p\frac{\partial u}{\partial n} v ds + \int_\Omega p\nabla u \cdot \nabla v \, dxdy = (f, v). \end{aligned}$$

The variational formulation is: Find u satisfying the Dirichlet condition such that

$$a(u,v) = \tilde{f}^*(v), \ \forall v \in V_1, \tag{2.13}$$

where $a(u,v) = (p\nabla u, \nabla v)$ and $\tilde{f}^*(v) = (f,v) + < pg, v >_{\Gamma_2}$.

Exercise 2.1.6. A fundamental solution of the PDE is the solution of

$$(L_y G(x,y))(x,y) = \delta_0(y-x), x, y \in \mathbb{R}^2,$$

in the distributional sense. **Greens** function of a PDE is a fundamental solution satisfying the boundary conditions. Greens functions are distributions.

(1) (10pts) In this problem we study Green's function

(a) Find a function $G \in H_0^1(I)$ satisfying

$$(G', v') = v(x_i), \qquad \text{for all } v \in H_0^1(I). \tag{2.14}$$

(b) Assuming G'' exists in some sense, interpret

$$-(G'', v) = v(x_i), \ \text{for } v \in C(I)$$

This means

$$-G''(x) = \delta(x_i)$$

where $\delta(x_i)$ the Dirac function. So we have a Green's function.

(c) Using this show we show that in one dimensional case, $u_h(x_i) = u(x_i)$ for each node.

(2) (10pts) For this and the next problem, assume $u_0 = g = 0$. Show the solution of (2.13) satisfies (2.12).

(3) (10pts) Show that we have an equivalent minimization problem : Find $u \in V_1$ such that $F(u) \le F(v)$ for all v where

$$F(v) = \frac{1}{2}a(v,v) - (f,v) - < pg, v >_{\Gamma_2} .$$

(Hint) Take derivative of $F(u+\epsilon v)$ w.r.t ϵ and set it to 0 at $\epsilon = 0$ to obtain (2.13).

Exercise 2.1.7. (1) (trace thm) (10pts) Let Ω be a unit square. Assuming the trace of v exists along the boundary, show that

$$\left(\int_\Gamma v^2 ds\right)^{1/2} \le C\|v\|_{H^1}, \quad \forall v \in H^1(\Omega).$$

(2) (10pts) Show that

$$\|v\|^2_{L^2(\Omega)} \le C_1|v|^2_{1,\Omega} + C_2\left|\int_\Omega v\,dx\right|^2 . \tag{2.15}$$

(Hint: first assume $v \in C^1(\Omega)$.)

Lemma 2.1.8. *Suppose that $w \in L^1(0,a)$, $w(x) \ge 0$, $a \ge b > 0$; then we have*

$$\int_0^b dr' \int_0^a dr'(r-r') \int_{r'}^r w(x)dx \le C_0 \int_0^a w(x)dx, \tag{2.16}$$

where $C_0 = (4a^2b - 4ab^2 + b^3)$ for $a \ge 2b$ and $C_0 = (a^3+b^3)/8$ for $b \le a \le 2b$.

Proof. We have, using integration by parts,

$$
\begin{aligned}
& \int_0^b dr' \int_0^a dr'(r-r') \int_{r'}^r w(x)dx \\
\leq\ & \int_0^b \left[\frac{(a-r')^2}{2} \int_{r'}^a w(x)dx + \frac{r'^2}{2} \int_0^{r'} w(x)dx \right] w(r) - \int_0^a dr \left[\int_0^b \frac{(r-r')^2}{2} dr' \right] \\
\leq\ & \left[\frac{C}{2} - \min_{0\leq r\leq a} \int_0^b \frac{(r-r')^2}{2} dr' \right] \int_0^a w(x)dx
\end{aligned}
$$

where $C = \int_0^b \max((a-r')^2, r'^2)dr' = \int_0^b (a-r')^2 dr' = (a^3-(a-b)^3)/3$, *for* $a \geq 2b$ *and* $C = \int_0^{a/2}(a-r')^2 dr' + \int_{a/2}^b r dr' = (a^3+4b^3)/3$ *for* $a \leq 2b$. $\min_{0\leq r\leq a} \int_0^b \frac{(r-r')^2}{2} dr' = \min(3br^2 - 3b^2 r + 4b^3) = -b^3/4$ *and we get (2.16).*

Lemma 2.1.9. *Let* $K = [0,1]\times[0,1]$. *We have*

$$
\|v\|_0^2 \leq \frac{1}{4}|v|_1^2 + \left| \int_K v(\xi,\eta) d\xi d\eta \right|^2, \forall v \in H^1(K) \tag{2.17}
$$

Proof. It is enough to prove (2.17) for $v \in C^2$. *For any* $(s,t),(s',t') \in K$,

$$
\begin{aligned}
& v^2(s,t) + v^2(s',t') - 2v(s,t)v(s',t') = \left(\int_{s'}^s \frac{\partial v}{\partial \xi}(\xi,t)d\xi + \int_{t'}^t \frac{\partial v}{\partial \xi}(s',\eta)d\eta \right)^2 \\
\leq\ & 2\left[(s-s') \int_{s'}^s \left(\frac{\partial v}{\partial \xi}(\xi,t) \right)^2 d\xi + (t-t') \int_{t'}^t \left(\frac{\partial v}{\partial \eta}(s',\eta) \right)^2 d\eta \right]
\end{aligned}
$$

Integrating over $(s,t),(s',t') \in K$ *on* K, *from Lemma 3.2 and now with* $a = b = 1$, *we get (2.17).*

2.2 Variational formulation of BVP

In many cases, second order BVP can be cast into a minimization problem of certain (nonlinear) functional.

Definition 2.2.1. Let V be a set in a Hilbert space. Let $B(u_0,\epsilon) = \{u \in V : \|u-u_0\| < \epsilon\}$ be a neighborhoodof u_0. Let f be a real valued function defined on V. We say $u_0 \in V$ is a *local minimizer* of f if there exists an $\epsilon > 0$ such that $f(u_0) \leq f(u), \quad \forall u \in B(u_0,\epsilon)$. If $f(u_0) < f(u), \quad \forall u \in B(u_0,\epsilon)$ we say $u_0 \in V$ is a *strong local minimizer* of f.

Definition 2.2.2. $u_0 \in V$ is called a *global minimizer* of f if $f(u_0) \leq f(u), \quad \forall u \in V$.

Definition 2.2.3. Let $u, \eta \in V$ with $\|\eta\| = 1$. Suppose there is a $t_0 > 0$ such that the function defined by $g(t) = f(u + t\eta),\ |t| < t_0$ has continuous m-th derivative, then the m-th directional derivative of f at u is

$$f^{(m)}(u; \eta) = g^{(m)}(0) = \frac{d^m f(u + t\eta)}{dt^m}|_{t=0}.$$

Definition 2.2.4. If $f^{(1)}(u_0; \eta) = 0,\ \forall \eta \in V, \|\eta\| = 1$, then f is stationary at u_0.

Theorem 2.2.5. *Suppose there exists a $u_0 \in V$ such that $f^{(1)}(u_0; \eta)$ exist for all direction η. If u_0 is a local minimizer of f, then f is stationary at u_0.*

Proof. By Taylor expansion, $f(u_0 + t\eta) = f(u_0) + tf^{(1)}(u_0; \eta) + o(t),\ \|\eta\| = 1$. Suppose $f^{(1)}(u_0; \eta)$ is nonzero, say positive for some η. then there exists a t_0 such that $tf^{(1)}(u_0; \eta) + o(t) < 0$, for $-t_0 < t < 0$. Hence every nhd of u_0 has a point $u = u_0 + t\eta$ such that $f(u) < f(u_0)$, which is a contradiction. □

Conversely we have

Theorem 2.2.6. *Suppose f is C^2 and u_0 is a stationary point of f. Suppose $f^{(2)}(u_0; \eta) \geq 0$ for all direction η. Then u_0 is a local minimizer of f.*

2.2.1 Euler- Lagrange equation

Green's identities: Let Ω be a domain in $\mathbb{R}^2$ with piecewise smooth boundary. We have for $u, v \in C'(\bar{\Omega})$,

$$\int_\Omega uv_x dxdy = \int_{\partial\Omega} uv\nu_1 ds - \int_\Omega u_x v dxdy \tag{2.18}$$

$$\int_\Omega uv_y dxdy = \int_{\partial\Omega} uv\nu_2 ds - \int_\Omega u_y v dxdy. \tag{2.19}$$

Apply this to each component of $\vec{v} = (v_1, v_2)$ and add to get

$$\int_\Omega u\nabla \cdot \vec{v} dxdy = \int_{\partial\Omega} u\vec{v} \cdot \vec{\nu} ds - \int_\Omega \nabla u \cdot \vec{v} dxdy.$$

Now if $\vec{v}$ is replaced by ∇v

$$\int_\Omega u\Delta v dxdy = \int_{\partial\Omega} u\frac{\partial v}{\partial \nu} ds - \int_\Omega \nabla u \cdot \nabla v dxdy.$$

Interchanging u and v and subtracting,

$$\int_\Omega (u\Delta v - v\Delta u)dxdy = \int_{\partial\Omega} (u\frac{\partial v}{\partial \nu} - v\frac{\partial u}{\partial \nu})ds,$$

where $\vec{\nu} = (\nu_1, \nu_2)$ is the outward unit normal vector to $\partial\Omega$.

A minimizer of a functional

Given a continuous function α defined on $\partial\Omega$, we let

$$V_\alpha = \{v \in C^2(\bar{\Omega}) : v = \alpha \text{ on } \partial\Omega\}$$

be the set of admissible functions. Then the corresponding test space is

$$V_0 = \{v \in C^2(\bar{\Omega}) : v = 0 \text{ on } \partial\Omega\}.$$

Consider a functional

$$f(u) = \int_\Omega F(x, y, u, u_x, u_y)dxdy, \ u \in V_\alpha.$$

We get a condition for stationary point for f by letting its first order directional derivative to be zero for all $\eta \in V_0$, i.e,

$$f^{(1)}(u; \eta) = \int_\Omega (\frac{\partial F}{\partial u}\eta + \frac{\partial F}{\partial u_x}\eta_x + \frac{\partial F}{\partial u_y}\eta_y)dxdy = 0. \tag{2.20}$$

Integrating by parts, we have

$$\int_\Omega \left[\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right)\right]\eta dxdy \tag{2.21}$$

$$+\int_{\partial\Omega}\left(\nu_1\frac{\partial F}{\partial u_x} + \nu_2\frac{\partial F}{\partial u_y}\right)\eta ds = 0, \quad \forall \eta \in V_0. \tag{2.22}$$

Since $\eta = 0$ on $\partial\Omega$, the line integral vanishes and hence get

$$\int_\Omega \left[\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right)\right]\eta dxdy = 0, \quad \forall \eta \in V_0 \tag{2.23}$$

which in turn implies

$$\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right) = 0 \text{ in } \Omega. \tag{2.24}$$

This is called the *Euler Lagrange equation.* The boundary condition $u = \alpha$ is called *essential boundary condition.* To find the *natural boundary condition*, consider $V = \{v \in C^2(\bar{\Omega})\}$. From (2.21) the second term is zero since $V_0 \subset V$. Thus (2.24) still holds. Now from (2.21) again, we have

$$\int_{\partial\Omega} \left(\nu_1 \frac{\partial F}{\partial u_x} + \nu_2 \frac{\partial F}{\partial u_y} \right) \eta ds = 0, \quad \forall \eta \in V.$$

Since $\eta \in V$ can have nonzero boundary conditions, we have

$$\nu_1 \frac{\partial F}{\partial u_x} + \nu_2 \frac{\partial F}{\partial u_y} = 0 \text{ on } \partial\Omega.$$

This is natural boundary condition.

Example 2.2.7. Consider a functional $f(u)$ defined on V_0 by

$$f(u) = \int \left[\frac{1}{2} p(x,y)(u_x^2 + u_y^2) + \frac{1}{2} q(x,y) u^2 - r(x,y)u \right] dxdy. \tag{2.25}$$

Here

$$F(x,y,u,u_x,u_y) = \frac{1}{2} p(x,y)(u_x^2 + u_y^2) + \frac{1}{2} q(x,y)u^2 - r(x,y)u.$$

Thus Its Euler-Lagrange equation is

$$-(pu_x)_x - (pu_y)_y + qu = r. \tag{2.26}$$

The natural boundary condition is

$$p(\nu_1 u_x + \nu_2 u_y) = p\frac{\partial u}{\partial \nu} = 0 \text{ on } \partial\Omega. \tag{2.27}$$

We have shown the solution of the pde (2.26) is a minimizer (in the classical sense) of the functional (2.25) on V_0. It turns out that the minimizer u satisfies natural boundary condition. The existence, uniqueness of the solution will be shown in the next section in Sobolev setting.

In some problems the boundary condition may be imposed on some part of the boundary, say, on $\Gamma_1 \subset \partial\Omega$. Let

$$V_g = \{v \in C^2(\bar{\Omega}) : v = g \text{ on } \Gamma_1\}.$$

In this case the space of test function is

$$V_{\Gamma_1} = \{v \in C^2(\bar{\Omega}) : v = 0 \text{ on } \Gamma_1\}.$$

Then from (2.21), we have

$$\nu_1 \frac{\partial F}{\partial u_x} + \nu_2 \frac{\partial F}{\partial u_y} = 0 \text{ on } \Gamma_2 = \Gamma - \Gamma_1.$$

Boundary conditions of different type on different portion of boundary are called *mixed boundary conditions.*

More general boundary condition

Suppose u satisfies $p\frac{\partial u}{\partial \nu} + \sigma u = 0$ on $\partial\Omega$. Now we have to modify $f(u)$ so that its stationary point satisfies the given boundary condition. Let

$$f(u) = \int_\Omega F(x, y, u, u_x, u_y)dxdy + \int_{\partial\Omega} G(x, y, u)ds, \;\; u \in C^2(\bar{\Omega}).$$

The condition for stationary point is

$$\int_\Omega \left[\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right)\right] \eta dxdy \tag{2.28}$$

$$+ \int_{\partial\Omega} \left(\nu_1 \frac{\partial F}{\partial u_x} + \nu_2 \frac{\partial F}{\partial u_y} + \frac{\partial G}{\partial u}\right) \eta ds = 0, \quad \forall \eta \in V_0. \tag{2.29}$$

Now the natural boundary condition is

$$\nu_1 \frac{\partial F}{\partial u_x} + \nu_2 \frac{\partial F}{\partial u_y} + \frac{\partial G}{\partial u} = 0 \text{ on } \partial\Omega.$$

Example 2.2.8. Let $G = \frac{1}{2}\sigma(x, y)u^2$. Then the natural boundary condition becomes

$$\frac{\partial u}{\partial \nu} + \frac{\sigma}{p} u = 0.$$

This kind of boundary condition is called *boundary condition of third type* or *Robin condition.*

Exercise 2.2.9. What change have to be made if we want the B.C. holds only on Γ_1(a portion of the boundary) ?

2.2.2 Existence, Uniqueness of a Weak solution

In this section we rephrase previous discussions in a weaker sense. We deal weak solutions in Sobolev spaces.

Example 2.2.10. Consider

$$\begin{aligned} -\nabla \cdot p\nabla u + qu &= f \text{ in } \Omega \\ u &= 0 \text{ on } \Gamma_1 \\ \frac{\partial u}{\partial n} &= 0 \text{ on } \Gamma_2 := \partial\Omega \backslash \Gamma_1. \end{aligned} \tag{2.30}$$

Let

$$V = \{v \in H^1(\Omega); v = 0 \text{ on } \Gamma_1\},$$

where $\Gamma_1 \subset \partial\Omega$ is nontrivial and $f \in L^2(\Omega)$. The associated functional is

$$F(u) = \frac{1}{2}a(u,u) - \tilde{f}(u),$$

where

$$a(u,v) = \int_\Omega p\nabla u \cdot \nabla v + quv\, dxdy, \quad v \in V \tag{2.31}$$

$$\tilde{f}(v) = \int_\Omega fv\, dxdy, \quad v \in V. \tag{2.32}$$

Its minimizer should satisfy (2.20). Thus the variational form for this problem is

$$a(u,v) = \tilde{f}(v), \quad v \in V. \tag{2.33}$$

For the existence and the uniqueness, we need a theorem. Let H be a Hilbert space equipped with a norm $\|\cdot\|$.

Theorem 2.2.11. *[Lax-Milgram] Let $a(\cdot,\cdot) : H \times H \to \mathbb{R}$ be a bilinear form satisfying*

(1) $|a(u,v)| \le C\|u\|_H\|v\|_H$ for all $u, v \in H$. (Bounded)

(2) $\rho\|u\|_H^2 \le a(u,u)$ for some constant $\rho > 0$. (Coercive)

(3) $\tilde{f}$ is a bounded linear functional.

Then there exists a unique solution $u \in H$ satisfying

$$a(u,v) = \tilde{f}(v).$$

Furthermore, there is a positive constant C s.t.

$$\|u\|_H \le C\|\tilde{f}\|. \tag{2.34}$$

Also, if a is symmetric, then it is a unique minimizer of the functional

$$F(u) = \frac{1}{2}a(u,u) - \tilde{f}(u).$$

From this result we can prove the existence, uniqueness of the solution of the pde.

Now we can verify the a form in the above examples satisfy the conditions of Lax-Milgram.

Lemma 2.2.12. *(Friedrich's first inequality) If Γ_1 is nontrivial, there is a constant $\alpha > 0$ such that*

$$\int_\Omega |\nabla u|^2 dxdy \ge \alpha \int_\Omega u^2 dxdy, \quad v \in V.$$

If $\Gamma_1 = \partial\Omega$, we have Poincaré inequality.

By the Lemma, $a(u,u)$ is coercive. Boundedness is easy to show. Now by Lax-Milgram lemma, there exists a unique solution $u \in V$ such that

$$a(u,v) = \tilde{f}(v), \quad v \in V.$$

One can easily see that u satisfies (2.30). Also,

$$\|\tilde{f}\| \le \|f\|_0 \text{ (the rhs of (2.30))}$$

Example 2.2.13. If we choose $H = H^1_{\Gamma_1} = \{v \in H^1, v|_{\Gamma_1} = 0\}$ and consider

$$-\Delta u = f \text{ in } \Omega \tag{2.35}$$
$$u = 0 \text{ on } \Gamma_1 \tag{2.36}$$
$$\frac{\partial u}{\partial \nu} = 0 \text{ on } \partial\Omega - \Gamma_1. \tag{2.37}$$

Example 2.2.14. (Robin condition) Next we consider Robin problem.

$$
\begin{aligned}
-\Delta u &= f \text{ in } \Omega && (2.38)\\
\frac{\partial u}{\partial \nu} + \sigma u &= 0 \text{ on } \Gamma = \partial\Omega, && (2.39)
\end{aligned}
$$

where $\sigma \in C(\Gamma), f \in L^2(\Omega),\ 0 < \sigma_0 \le \sigma(x,y) \le \sigma_1,\ (x,y) \in \Gamma$. Then its corresponding variational problem is

$$
\begin{aligned}
a(u,v) &= \int_\Omega \nabla u \cdot \nabla v\, dxdy + \int_{\partial\Omega} \sigma uvds, \quad v \in V = H^1(\Omega) && (2.40)\\
\tilde{f}(v) &= \int_\Omega fv\, dxdy, \quad v \in V. && (2.41)
\end{aligned}
$$

To show that this is equivalent to the p.d.e. above, we need some preliminary concepts as below.

Lemma 2.2.15. *(Trace theorem) For any $u \in H^1(\Omega)$, the restriction of u to $\partial\Omega$ exists and belongs to $L^2(\partial\Omega)$ and satisfies*

$$\int_{\partial\Omega} u^2 ds \le \beta \|u\|_1^2, \quad \forall u \in H^1(\Omega).$$

Lemma 2.2.16. *(Friedrich's second inequality) If $\Gamma_1 \subset \partial\Omega$ is nontrivial, there is a constant $\alpha > 0$ such that*

$$\int_\Omega |\nabla u|^2 dxdy + \int_{\Gamma_1} u^2 ds \ge \alpha \|u\|_1^2,\ \forall u \in H^1(\Omega).$$

By Friedrich's 2nd inequality, one can easily show the coerciveness:

$$a(u,u) \ge \alpha \min(\sigma_0, 1) \|u\|_1^2,\ \forall u \in H^1(\Omega).$$

(Note the difference between two versions of Friedrich's inequality.) For boundedness, we note that

$$a(u,u) \le \int_\Omega |\nabla u|^2 dxdy + \sigma_1 \int_\Gamma u^2 ds \le \|u\|_1^2.$$

by the trace theorem.

Remark 2.2.17. Note that if $\sigma(x,y) \equiv 0$ on Γ then (2.38) becomes pure Neumann problem, and a is not coercive. To see this, we note Green's second identity:

$$\int_\Omega (u\Delta v - v\Delta u)\, dxdy = \int_{\partial\Omega} (u\frac{\partial v}{\partial \nu} - v\frac{\partial u}{\partial \nu})ds.$$

Set $v = 1$. Then

$$-\int_\Omega \Delta u\, dxdy = \int_\Omega g dxdy = -\int_{\partial\Omega} \frac{\partial u}{\partial \nu} ds = 0.$$

Hence the partial differential equation has a solution only if g satisfies the *consistency condition*: $\int_\Omega f\, dxdy = 0$.

2.2.3 More general coefficient

Example 2.2.18. Let $V = H_{\Gamma_1} = \{v \in H^1(\Omega), v = 0 \text{ on } \Gamma_1 \subset \partial\Omega\}$. Consider : find $u \in V$ satisfying

$$a(u, v) = \tilde{f}(v), \; \forall v \in V, \tag{2.42}$$

where

$$a(u, v) := \int_\Omega \sum_{i,j=1}^{2} a_{ij} \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} dxdy + \int_\Omega cuv\, dxdy, \quad v \in V \tag{2.43}$$

$$\tilde{f}(v) = \int_\Omega fv\, dxdy, \quad v \in V, \tag{2.44}$$

and $a_{ij} = a_{ji} \in C(\bar{\Omega}),\; c \in C(\bar{\Omega}), c \geq 0,\; f \in L^2(\Omega)$. Assume there exists a constant $\lambda > 0$ such that

$$\sum_{i,j} a_{ij}\xi_i\xi_j \geq \lambda \sum_i \xi_i^2, \quad \forall (x, y) \in \Omega, \xi_i \in \mathbb{R}.$$

(This is equivalent to: eigenvalues of $\{a_{ij}\}$ are positive.) Hence we have $a(u, u) \geq \rho\|u\|_1^2, u \in V$. The corresponding boundary value problem is

$$\mathcal{L}u = f \text{ in } \Omega \tag{2.45}$$

$$u = 0 \text{ on } \Gamma_1 \tag{2.46}$$

$$\frac{\partial u}{\partial \nu_{\mathcal{L}}} = 0 \text{ on } \partial\Omega - \Gamma_1, \tag{2.47}$$

where $\mathcal{L}u = -\sum_{i,j} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + cu$ and $\frac{\partial u}{\partial \nu_{\mathcal{L}}}$ is the conormal derivative defined by

$$\frac{\partial u}{\partial \nu_{\mathcal{L}}} = \sum_{i,j} \nu_i a_{ij} \frac{\partial u}{\partial x_j}. \tag{2.48}$$

Exercise 2.2.19. (1) Derive (2.48) from (2.42).

2.2.4 Inhomogeneous boundary condition

Consider

$$\begin{aligned} \mathcal{L}u &= f \text{ in } \Omega & (2.49)\\ u &= g \text{ on } \partial\Omega, & (2.50) \end{aligned}$$

where $\mathcal{L}u = -(pu_x)_x - (pu_y)_y + qu$ on Ω. Take $V = H_0^1(\Omega)$ and let

$$\begin{aligned} a(u, v) &= \int_\Omega (pu_x v_x + pu_y v_y + quv)dxdy, \ v \in V & (2.51)\\ \tilde{f}(v) &= \int_\Omega fv\, dxdy. & (2.52) \end{aligned}$$

Let $H_g^1(\Omega) = \{u \in H^1(\Omega); u = g \text{ on } \partial\Omega\}$.

2.2.5 Question

(1) Given a function $u \in H^1(\Omega)$ how do we define its restriction to $\partial\Omega$?

(2) Given a function $g \in H^1(\Omega)$ does there exist $u \in H^1(\Omega)$ such that the restriction of u to $\partial\Omega$ is g?

The first question is answered by the *trace* and *trace inequality.* The answer to the next question is true if and only if $g \in H^{1/2}(\partial\Omega)$.

Now the functional to be minimized is

$$F(u) = \frac{1}{2}a(u, u) - \tilde{f}(u), \quad u \in H_g^1(\Omega).$$

Since $H_g^1(\Omega)$ is not a linear space, Lax-Milgram lemma does not apply. Instead, for any fixed $u_g \in H_g^1(\Omega)$, we can write

$$H_g^1(\Omega) = \{u \in H^1(\Omega) : u = w + u_g, w \in H_0^1(\Omega)\}$$

and define

$$\begin{aligned} F_g(w) &= \tilde{f}(w+u_g)-\tilde{f}(u_g) && (2.53)\\ &= \frac{1}{2}a(w+u_g,w+u_g)-\tilde{f}(w+u_g)-\frac{1}{2}a(u_g,u_g)-\tilde{f}(u_g) && (2.54)\\ &= \frac{1}{2}a(w,w)+a(u_g,w)-\tilde{f}(w) && (2.55)\\ &= \frac{1}{2}a(w,w)-\tilde{f}_g(w) && (2.56)\end{aligned}$$

where $\tilde{f}_g(w) := \tilde{f}(w)-a(u_g,w)$. Now Lax-Milgram lemma asserts that there exists a unique element $u_0 \in H_0^1(\Omega)$ (minimizer of $\tilde{f}_g(w)$) such that

$$a(u_0,v)=\tilde{f}_g(v), \quad v \in H_0^1(\Omega)$$

and hence

$$a(u_0+u_g,v)=\tilde{f}(v), \quad v \in H_0^1(\Omega).$$

Let $u=u_0+u_g$, then $u \in H_g^1(\Omega)$ and satisfies

$$a(u,v)=\tilde{f}(v), \quad v \in H_0^1(\Omega).$$

It is easy to verify that u_0 minimizes F_g over $H_0^1(\Omega)$. Since

$$a(u,v)-\tilde{f}(v)=(\mathcal{L}u-f,v), \quad v \in H_0^1(\Omega),$$

u is the generalized solution of (2.49), (2.50).

Nonhomogenous Robin Condition

In this case, we have instead of (2.50),

$$u_\nu+\sigma(x,y)u=\xi(x,y) \text{ on } \partial\Omega, \qquad (2.57)$$

where $\sigma(x,y)\in C(\partial\Omega), \xi(x,y)\in L^2(\partial\Omega)$ and $0<\sigma_0 \le \sigma(x,y) \le \sigma_1$.

If $u \in H^1(\Omega)$ is a weak solution of (2.49),(2.57), then integration by parts,

$$\begin{aligned} (\mathcal{L}u-f,v) &= a(u,v)-\tilde{f}(v)-\int_{\partial\Omega} pu_\nu v ds && (2.58)\\ &= a(u,v)-\tilde{f}(v)-\int_{\partial\Omega} p(\xi-\sigma u)v ds,\ v\in H^1(\Omega). && (2.59)\end{aligned}$$

Let

$$\begin{aligned} \hat{a}(u,v) &= a(u,v) + \int_{\partial\Omega} p\sigma uv\, ds, \quad u,v \in H^1(\Omega) && (2.60)\\ \hat{f}(v) &= \tilde{f}(v) + \int_{\partial\Omega} p\xi v\, ds,\ v \in H^1(\Omega). && (2.61) \end{aligned}$$

Then we have

$$\hat{a}(u,v) = \hat{f}(v),\ \forall v \in H^1(\Omega). \tag{2.62}$$

Apply Lax-Milgram lemma (Check conditions) to (2.60) to get a unique $u \in H^1(\Omega)$ satisfying (2.62).

One can show the solution of this problem satisfies (2.49) and (2.57). Indeed, using integration by parts and setting $v \in H_0^1(\Omega)$, we can see that u satisfies (2.49). Hence we have

$$\int_{\partial\Omega} p(u_\nu + \sigma u - \xi)v ds = 0, \quad v \in H^1(\Omega).$$

This shows that u satisfies the nonhomogenous B.C. (2.57).

Exercise 2.2.20. (1) (10pts) Apply Lax-Milgram to show the existence of solution for (2.35).

(2) (10pts) Show that the solution of (2.2.18) satisfies (2.43)-(2.44).

2.3 Ritz-Galerkn Method

For the simplicity of presentation, we assume Ω is a polygonal domain. Consider

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega && (2.63)\\ u &= 0 \text{ on } \partial\Omega. && (2.64) \end{aligned}$$

Assume Ω is partitioned by triangles. Let S_h be a finite dimensional subspaces of $H_0^1(\Omega)$. The finite element problem is to find a $u_h \in S_h$ satisfying

$$a(u_h, v_h) = \tilde{f}(v_h), \quad v_h \in S_h. \tag{2.65}$$

W. Ritz(1908) used polynomials of higher degree to solve the variational problem in a finite dimensional subspace.

(1) Ritz, W. Neue Methode zur Losung gewisser Randwertaufgaben. Gesellschaft der Wissenschaften zu Gottingen: Math.-physik. Klasse: Nachrichten. Gottingen, 1908..

(2) Galerkin, B. G. Sterzhni i plastinki: Riady v nekotorykh voprosakh uprugogo ravnovesiia sterzhnei i plastinok. Vestnik inzhenerov, 1915, vol. 1, no. 19, pp. 897-908

Let $u_h = \sum_{j=1}^N u_j\phi_j$. Then (2.65) becomes

$$a(\sum_{j=1}^N u_j\phi_j, \phi_i) = f(\phi_i),\ i = 1, ..., N.$$

In matrix form we have

$$A \cdot \vec{u} = \vec{f}, \tag{2.66}$$

where $A_{ij} = a(\phi_j, \phi_i)$, $\vec{u} = (u_1, \cdots, u_N)$ and $\vec{f_i} = \int f\phi_i dxdy$. Such a matrix A is called the *stiffness matrix* and $\vec{f}$ is called a *load vector*. A is SPD and hence the system has a unique solution.

2.3.1 Choice of S_h and its basis

Thus far we have no assumption on the shape (support, degrees, etc.) of basis functions. Basic idea is to choose a nice basis $\{\phi_i\}$ for S_h so that

(1) Easy to construct A

(2) A is sparse (To save storage and computations)

(3) The condition number of the matrix A is not too large.

Often, we use continuous functions which are piecewise linear on triangular elements. Note that in this case, $A_{ij} \neq 0$ only if the node i, j are adjacent. Thus the matrix is sparse.

2.3.2 Inhomogeneous boundary conditions

Consider solving

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega \\ u &= g \text{ on } \partial\Omega. \end{aligned}$$

From earlier discussions, the variational formulation is to find a $u = u^0 + u_g$, where $u^0 \in H_0^1$ satisfies

$$a(u^0, v) = \tilde{f}(v) - a(u_g, v), \quad v \in H_0^1,$$

where u_g is any function in $H_g^1(\Omega) = \{u \in H^1(\Omega), u = g \text{ on } \partial\Omega\}$. Then the FEM amounts to finding $u_h^0 \in S_h$ such that

$$a(u_h^0, v_h) = \tilde{f}(v_h) - a(u_g^h, v_h), \quad v_h \in S_h. \tag{2.67}$$

Here u_g^h is an finite element approximation to u_g. In matrix form,

$$A \cdot \vec{u} = \vec{f^*}, \tag{2.68}$$

where $f_i^* = \tilde{f}(\phi_i) - a(u_g^h, \phi_i)$. Assume the number of unknowns on the boundary is p and let $\{\phi_j\}_{j=N+1}^{N+p}$ are the piecewise linear basis associated with the boundary. One often try to approximate u_g in the form $u_g^h = \sum_{j=N+1}^{N+p} c_j \phi_j$ so that

$$u_g^h(x_j, y_j) = g(x_j, y_j), \ (x_j, y_j) \in \partial\Omega_h(\text{boundary nodes}). \tag{2.69}$$

In particular, if ϕ_j are Lagrange basis so that $\phi_j(x_i, y_i) = \delta_{ij}$, then u_g ca be replaced by $\sum_{j=N+1}^{N+p} u_j \phi_j$.

2.4 Finite Element Method -A Concrete Ritz-Galerkin method

2.4.1 Finite element basis functions

We assume $\bar{\Omega}$ is subdivided by a non-overlapping elements, say triangles or rectangles of certain regular shape. For a mesh generator, see

http://www.cs.cmu.edu/q̃uake/triangle.html

Notations

- L: Total number of elements
- M: Total number of nodes
- T: Number of nodes in a single element
- K_ℓ: $\ell = 1, 2, \cdots, L$ the elements
- $N_i = (x_i, y_i)$: the nodes
- $\hat{K}$: the standard(reference) element

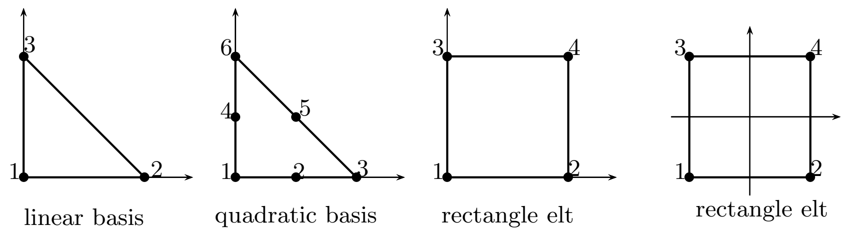


Figure 2.5: Reference element and nodes

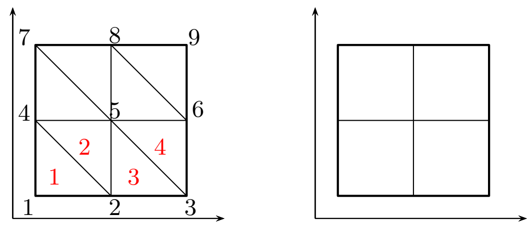


Figure 2.6: Global numbering of Elements(in red) and nodes

Since the support of ϕ_i is usually a small subset of $\bar{\Omega}$, we say that ϕ_i has *local support*. A rough geometrical description of ϕ_i is that of a "tent" centered about N_i. The floor of the tent is the support of ϕ_i. If $N_i \notin \partial\Omega = \Gamma$, ϕ_i vanishes on the boundary of its support.

By definition, a function u belongs to $S_M = S_h$ if and only if it can be expressed as

$$u(x, y) = \sum_{i=1}^{M} u_i \phi_i(x, y), \quad (x, y) \in \bar{\Omega}.$$

Each u is a continuous, piecewise polynomial over $\bar{\Omega}$.

Interpolation

Define the interpolation $I_h : C(K) \to S^h(K)$ by the conditions $(I_h u)(a_i) = u(a_i)$ for all $i = 1, \cdots, T$. Piecewise linear element was introduced by Courant(1943), Prager and Synge(1947). It is often better than higher order functions when the solution has less regularity. For any $u \in C(\bar{\Omega})$ let $u_I \in S_M$ be defined by

$$u_I(x,y) = \sum_{i=1}^{M} u(N_i)\phi_i(x,y), \quad (x,y) \in \bar{\Omega}.$$

The function u_I is the interpolant of u in S_M.

For error analysis we need to view u_I as interpolant of $u \in H^p(\Omega)$. But a function in $H^p(\Omega)$ is actually an equivalence class of functions defined a.e, $u(N_i)$ is not well-defined.

An equivalent class can contain at most one continuous function. If it does contain such function, we shall define $u(N_i)$ to be the value of that function at N_i. According to the Sobolev embedding theorem,

$$H^p(\Omega) \subset C^0(\Omega), \quad p \geq 1 \text{ if one space dimension}$$

$$H^p(\Omega) \subset C^0(\Omega), \quad p \geq 2 \text{ if two or three space dimension .}$$

Define k and m to be the greatest integers such that

$$P_k(\bar{\Omega}) \subset V_M \subset C^{m-1}(\bar{\Omega}) \tag{2.70}$$

is satisfied, where $P_k(\bar{\Omega})$ denote the space of all polynomials of degree k on $\bar{\Omega}$)(global, not piecewise). In this case, if $u \in P_k(\bar{\Omega})$ $u_I(x,y) = u(x,y)$, $\forall (x,y) \in \Omega$, i.e, all polynomial of degree $\leq k$ are interpolated exactly in V_M. "This property makes k the fundamental parameter in the error analysis".

Regarding m, m must be ≥ 1(continuous). If $m = 1$, $V_M \not\subset C^1(\Omega)$. To achieve C^1, one has to enlarge basis function(Assign more than one basis function at some of the nodes) so that not only u_I interpolates the values of u but certain derivatives of u_I interpolates the corresponding derivatives of u. This is known as "Hermite interpolation" in contrast to Lagrange interpolation which interpolates values of u only.

In general, a function in $C^{m-1}(\bar{\Omega})$ does not does not belong to $H^m(\Omega)$ since there exists nowhere differentiable continuous functions. If u is piecewise

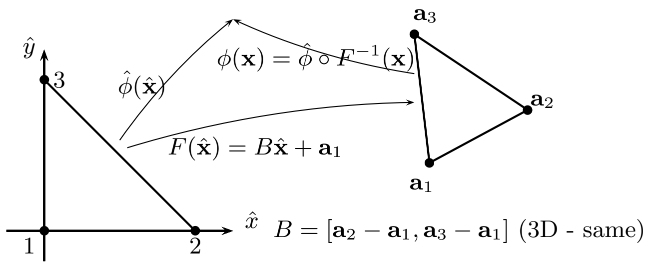


Figure 2.7: Reference triangle and the mapping

polynomial in $C^{m-1}(\bar{\Omega})$ then it belongs to

$$H^m(\Omega)$$

Also,

$$H^{m+1}(\Omega) \subset C^{m-1}(\Omega)$$

In general, this does not hold since there exists nowhere differentiable continuous functions. If $m = 1$, then V_M is appropriate for 2nd order elliptic problem, called "conforming". Referring to the figure (2.7), the nodal basis functions are

$$\hat{\phi}_1 = 1 - \hat{x} - \hat{y}, \ \hat{\phi}_2 = \hat{x}, \ \hat{\phi}_3 = \hat{y}. \tag{2.71}$$

Example 2.4.1. Piecewise linear basis on triangular element. First, on the standard refernce basis element $\hat{K}$,

$$\hat{\phi}_r(\hat{x}, \hat{y}) = \hat{c}_r^1 + \hat{c}_r^2\hat{x} + \hat{c}_r^3\hat{y}, r = 1, 2, 3.$$

Let the local basis function on a general element K_ℓ be give by

$$\phi_r^\ell(x, y) = c_{r,\ell}^1 + c_{r,\ell}^2 x + c_{r,\ell}^3 y, (x, y) \in K_\ell, \ \ell = 1 \cdots, L, \ r = 1, 2, 3.$$

It is nothing but the restriction of global basis function $\phi_{i_r^\ell}(x, y), i = 1, \cdots, M$. Since $\mathbf{x}$ is related to $\hat{\mathbf{x}}$ by the affine map (See figure 2.6)

$$\begin{bmatrix} x \\ y \end{bmatrix} = F_\ell \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} b_{11}\hat{x} + b_{12}\hat{y} + a_1 \\ b_{21}\hat{x} + b_{22}\hat{y} + a_2 \end{bmatrix} \equiv B_\ell \hat{\mathbf{x}} + \hat{d}_\ell, \tag{2.72}$$

we have

$$\phi_r^\ell(x, y) = \hat{\phi}_r(\hat{x}, \hat{y}) = \hat{\phi}_r \circ F_\ell^{-1}(x, y).$$

We use conventional counterclockwise ordering as local ordering. Is is easy to see that

$$B = [\mathbf{a}_2 - \mathbf{a}_1, \mathbf{a}_3 - \mathbf{a}_1].$$

Example 2.4.2. Piecewise quadratic basis on triangular elements(fig. 2.5).

$$\phi_r(x,y) = c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2.$$

The corresponding standard basis functions on half of $\hat{K} = [-1,1]^2$ are:

$$\left\{ \begin{array}{llll} \hat{\phi}_1(\hat{x},\hat{y}) & = (1-\hat{x}-\hat{y})(1-2\hat{x}-2\hat{y}), & \hat{\phi}_4(\hat{x},\hat{y}) = 4\hat{x}\hat{y} \\ \hat{\phi}_2(\hat{x},\hat{y}) & = \hat{x}(2\hat{x}-1), & \hat{\phi}_5(\hat{x},\hat{y}) = 4(1-\hat{x}-\hat{y})\hat{y} \\ \hat{\phi}_3(\hat{x},\hat{y}) & = \hat{y}(2\hat{y}-1), & \hat{\phi}_6(\hat{x},\hat{y}) = 4(1-\hat{x}-\hat{y})\hat{x} \end{array} \right\}$$

The same transformation (2.72) maps $\hat{K}$ onto K_ℓ with mid-edge of $\hat{K}$ to the mid-edge of K_ℓ. We used the ordering of vertex nodes first and then the mid-edge opposite to the nodes 1,2 and 3 points as nodes 4,5 and 6.

Example 2.4.3. (Piecewise bilinear basis on rectangular elements) Let

$$\phi_r(x,y) = c_1 + c_2 x + c_3 y + c_4 xy.$$

Then the standard basis functions on $\hat{K}$ are

$$\hat{\phi}_1(\hat{x},\hat{y}) = \frac{1}{4}(1-\hat{x})(1-\hat{y}), \qquad \hat{\phi}_2(\hat{x},\hat{y}) = \tfrac{1}{4}(1+\hat{x})(1-\hat{y}) \tag{2.73}$$

$$\hat{\phi}_3(\hat{x},\hat{y}) = \frac{1}{4}(1+\hat{x})(1+\hat{y}), \qquad \hat{\phi}_4(\hat{x},\hat{y}) = \tfrac{1}{4}(1-\hat{x})(1+\hat{y}). \tag{2.74}$$

2.4.2 Assembly of stiffness matrix

Consider

$$-\nabla \cdot p\nabla u + qu = f \text{ in } \Omega \tag{2.75}$$

$$u = g \text{ on } \Gamma_1 \tag{2.76}$$

$$u_\nu + \sigma u = \xi \text{ on } \Gamma_2, \quad \Gamma = \Gamma_1 \cup \Gamma_2. \tag{2.77}$$

Let

$$H_g^1(\Omega) = \{v \in H^1(\Omega) : v = g \text{ on } \Gamma_1\} \text{ (affine space)}$$

$$H_{\Gamma_1}^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_1\} \text{ (linear space)}$$

Find $u \in H_g^1(\Omega)$ such that

$$a(u,v) = \tilde{f}(v), \quad \forall v \in H^1_{\Gamma_1}(\Omega), \tag{2.78}$$

where

$$\begin{aligned} a(u,v) &= \iint (pu_x v_x + pu_y v_y + quv)\, dx + \int_{\Gamma_2} p\sigma uv\, ds && (2.79)\\ &= (p\nabla u, \nabla v) + (qu, v) + < p\sigma u, v >_{\Gamma_2} && (2.80)\end{aligned}$$

and

$$\begin{aligned} \tilde{f}(v) &= (f, v) + < p\xi, v >_{\Gamma_2} && (2.81)\\ &= \tilde{f}(v) + < p\xi, v >_{\Gamma_2} && (2.82)\end{aligned}$$

Then as shown before, the solution u minimizes the functional

$$f(u) = \frac{1}{2} a(u,u) - \tilde{f}(u) \quad \forall u \in H_g^1(\Omega). \tag{2.83}$$

If u_g is any function in $H_g^1(\Omega)$ one have

$$H_g^1(\Omega) = H_0^1(\Omega) + u_g = \{u \in H^1(\Omega) : u = u_0 + u_g, u_0 \in H^1_{\Gamma_1}(\Omega)\}$$

and (2.78) is equivalent to finding $u_0 \in H^1_{\Gamma_1}(\Omega)$ such that

$$a(u_0, v) = \tilde{f}(v) - a(u_g, v), \ \forall v \in H^1_{\Gamma_1}(\Omega) \tag{2.84}$$

Let $\mathcal{T}_h = \{K_\ell\}$ be a triangulation of the domain Ω and let

2.5 Outline of Programming

$$V_N = span\{\phi_i \text{ linear on each element },\phi_i \in H^1_{\Gamma_1}(\Omega)\}.$$

Let us list some notations:

- L: Total number of elements
- M: Total number of nodes
- Γ_1: the part of boundary where Dirichlet condition is imposed

- M_1: the number of nodes on Γ_1
- $\mathcal{J}_1$: the index set of nodes on Γ_1 (Dirichlet condition)
- Γ_2: the part of boundary where Neumann condition is imposed
- L_2: the number of element edges on Γ_2
- N: number of nodes in $\Omega \cup \Gamma_2$ (Total unknowns)
- $\mathcal{J}$: index set of nodes in $\Omega \cup \Gamma_2$
- Q: number of integration points in each element

To form a finite element matrix, we need to replace (2.84) by a finite dimensional analog: Find $u_h := u_{h,0} + u_g^h$ s.t.

$$a(u_{h,0}, v) = \tilde{f}(v) - a(u_g^h, v), \ \forall v \in V_N, \tag{2.85}$$

where

$$u_{h,0} = \sum_{j \in \mathcal{J}} u_j \phi_j$$

and u_g is replaced by a P. L. function satisfying the BC(at least weakly.) We usually choose

$$u_g^h = \sum_{i \in \mathcal{J}_1} g_i \phi_i \text{ so that } u_g^h(N_i) = g(N_i).$$

Even though $u_g^h \notin H_g^1(\Omega)$, it would suffice our purpose.

A tip for the boundary nodes

When $\bar{\Gamma}_1 \cap \bar{\Gamma}_2 \neq \emptyset$, it is important for application to

(1) place nodes at the end points of $\bar{\Gamma}_1$.

(2) assign nodes to $\bar{\Gamma}_1$ rather than to Γ_2.

Assembly of stiffness matrix

We usually compute so called *element stiffness matrix* and sum them over all elements to assemble the global stiffness matrix, denoted by A here. From (2.85) we have

$$a(u_{h,0}, \phi_i) = \sum_{j \in \mathcal{J}} u_j \int_\Omega (p \nabla \phi_j \cdot \nabla \phi_i + q \phi_j \phi_i) + \sum_{j \in \mathcal{J}} u_j \int_{\Gamma_2} p \sigma \phi_j \phi_i ds, \ i \in \mathcal{J},$$

and

$$F_i^* := \tilde{f}(\phi_i) - a(u_g^h, \phi_i) = \int_\Omega f\phi_i + \int_{\Gamma_2} p\xi\phi_i \, ds - \sum_{j\in\mathcal{J}_1} g(N_j)a(\phi_j, \phi_i), \ i \in \mathcal{J}.$$

Hence we get

$$A\mathbf{u} = F^*, \ \mathbf{u} = (u_j)_{j\in\mathcal{J}}$$

where

$$A_{ij} = \int_\Omega (p\nabla\phi_j \cdot \nabla\phi_i + q\phi_j\phi_i) + \int_{\Gamma_2} p\sigma\phi_j\phi_i ds.$$

Some general issues:

(1) Input data : $\Omega, \Gamma_1, \Gamma_2, p, g_1, \xi$, coefficients, etc.

(2) Construction and representation of $\mathcal{T}_h$

(3) Computation of element stiffness matrix a^K and f^K

(4) Assembly of global stiffness matrix A, F^*

(5) Linear solver for the system $A\mathbf{u} = F^*$

(6) Presentation of result. Discrete L^2, H^1-error. Numerical Table, order of convergence, graphics.

Remark on (2): quasi uniform—essentially the same size, but it is desirable to vary the size of triangle—adaptive or successive refinement. Conforming: vertex should not lie in the interior of an edge.

We need to compute A_{ij}.

(1) $A_{ij} := a(\phi_j, \phi_i)_\Omega, \ i, j = 1, \cdots, M$

(2) $F_i := F(\phi_i)_\Omega, \ i = 1, \cdots, M$

(3) $A_{ij} := A_{ij} + a(\phi_j, \phi_i)_{\Gamma_2}, \ i, j = 1, \cdots, M$

(4) $F_i := F_i + F(\phi_i)_{\Gamma_2}, \ i = 1, \cdots, M$

(5) For $j \in \mathcal{J}_1$:

  (a) $F_i := F_i - g(N_j)A_{ij}, \ i \in \mathcal{J}$
  (b) $A_{ij} := 0 = A_{ji}, \ i \in \mathcal{J}$
  (c) $A_{ji} := 0, \ i \in \mathcal{J}_1, \ i \neq j$

(d) $F_j := g(N_j); A_{jj} = 1$

$$
A = \begin{array}{cc}
\begin{array}{cc} \qquad N \qquad\qquad & \qquad N_1 \qquad \end{array} & \\
\left[\begin{array}{ccc|ccc}
A_{11} & * & * & & & \\
* & \ddots & * & & O & \\
* & * & A_{NN} & & & \\
\hline
 & & & 1 & 0 & 0 \\
 & O & & 0 & \ddots & 0 \\
 & & & 0 & 0 & 1
\end{array}\right] & \begin{array}{c} N \\ \\ N_1 \end{array}
\end{array}
\tag{2.86}
$$

Step (1), (2) are related to the interior nodes, Step (3), (4) are related to the natural BC. while step (5) is related to the essential BC. Since A_{ij} is symmetric, the computation in (1), (3) and (5) are done for $j = 1, \cdots, i$ only to save time and memory. (For details see Axellson p. 185)

Remark 2.5.1. (1) We used full matrix notation A_{ij} for simplicity of presentation; However, one need to exploit the sparseness of the matrix to save memory. So it may be nice to provide a $M \times 5$(for 5 point stencil) matrix and store nonzero $A_{i,j}$. One can write a class file to define matrix-function that looks like $A(i, j)$ but has single array structure of length $5M$.

(2) The true number of unknowns are N, not M. The (d) of step (5) means we append the following trivial equations next to the $N \times N$ equations;

$$u(N_i) = g(N_i), \quad i \in \mathcal{J}_1$$

2.5.1 Computation of $a(\phi_i, \phi_j)$ elementwise.

Note that

$$a(\phi_i, \phi_j) = \sum_K \int_K (p\nabla\phi_i \cdot \nabla\phi_j + q\phi_i\phi_j) + \sum_K \int_{\bar{K}\cap\Gamma_2} p\sigma\phi_i\phi_j ds := \sum_K a_{ij}^K,$$

where the summation runs through the common support of ϕ_i and ϕ_j. We compute this entry by computing the contribution of $a^K(\phi_i, \phi_j)$, called the *element stiffness matrix* for each element K. However, note the index given here is global index. In the code we use local index corresponding to the global index.

Notations:

Z: $M \times 2$ matrix, $Z(i,j), j = 1,2$, are the (x,y) coordinates of the node i - vertex coordinates table.

T: $3 \times L$ matrix, $T(\alpha, \ell), \alpha = 1,2,3$, denotes the global node number of local α-th node of ℓ-th triangle - element node table.

A triangulation $\mathcal{T}_h$ may be represented by two matrices $Z : 2 \times M$ matrix and $T : 3 \times L$ matrix.

Definition 2.5.2 (lexicographic order)**.** If $A = \{1,2\}$ and $B = \{a,b,c\}$. Then the lexicographic order of the set $A \times B$ is

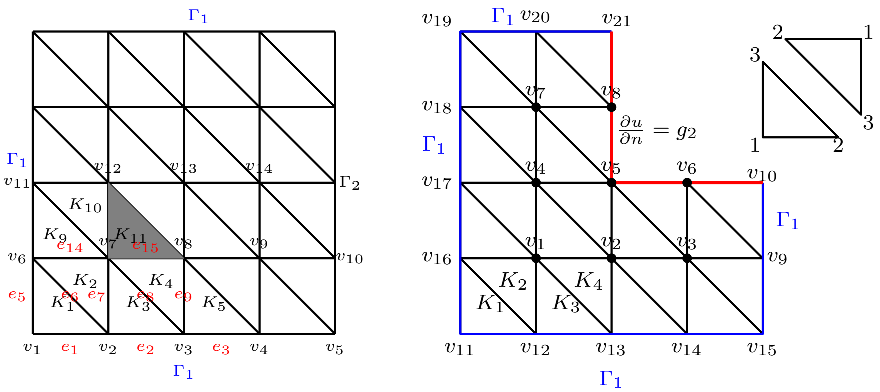
$$(1,a),(1,b),(1,c),(2,a),(2,b),(2,c).$$

Example 2.5.3. Let us divide a unit square by 4×4 uniform meshes where each sub-rectangle is subdivided by the diagonal of slope -1. See figure 2.5.3. Label all the vertex nodes linearly starting for the bottom row as $1,2,3,\cdots,25$, and lexicographically as $(1,1),(1,2),\cdots,(2,1),(2,2),\cdots$ also. Label the elements from lower left corner as $K_1\backslash K_2$, $K_3\backslash K_4$, $K_5\backslash K_6,\cdots$.

Compute the element stiffness matrix for each triangle, and add all the contribution to three vertices as K runs through all element. If $i = j$, K runs through all element having the node i as a vertex. If $i \neq j$, K runs through all element having the line segment $\bar{ij}$ as an edge. In this way, we assemble the global matrix A collecting the contribution form each element.

For example, let $K = K_{11}$. The global indices for its vertices are $7, 8$ and 12. For the element matrix we need to compute $a^K(\phi_i, \phi_j)$ for $i,j = 7,8,12$.

$$\begin{aligned}
a^K_{11} = a_K(\phi_7,\phi_7) &= \int_K (-\frac{1}{h}, -\frac{1}{h}) \cdot (-\frac{1}{h}, -\frac{1}{h}) = 1 \\
a^K_{12} = a_K(\phi_7,\phi_8) &= \int_K (-\frac{1}{h}, -\frac{1}{h}) \cdot (\frac{1}{h}, 0) = -\frac{1}{2} \\
a^K_{13} = a_K(\phi_7,\phi_{12}) &= \int_K (-\frac{1}{h}, -\frac{1}{h}) \cdot (0, \frac{1}{h}) = -\frac{1}{2} \\
a^K_{22} = a_K(\phi_8,\phi_8) &= \int_K (\frac{1}{h}, 0) \cdot (\frac{1}{h}, 0) = \frac{1}{2} \\
a^K_{23} = a_K(\phi_8,\phi_{12}) &= \int_K (\frac{1}{h}, 0) \cdot (0, \frac{1}{h}) = 0 \\
a^K_{33} = a_K(\phi_{12},\phi_{12}) &= \int_K (0, \frac{1}{h}) \cdot (0, \frac{1}{h}) = \frac{1}{2}
\end{aligned}$$



(a) Rectangular domain (b) L-shape, mixed BC, unknowns at •

Figure 2.8: Label of elements and vertices

Here we used the notation :

$$a_K(\phi_i, \phi_j) = \int_K p \nabla \phi_i \cdot \nabla \phi_j \, dx. \tag{2.87}$$

The element stiffness matrix $a^{K_{11}}$ (corresponding to the vertices 7,8 and 12) is

$$\begin{bmatrix} 1, & -\frac{1}{2}, & -\frac{1}{2} \\ -\frac{1}{2}, & \frac{1}{2}, & 0 \\ -\frac{1}{2}, & 0, & \frac{1}{2} \end{bmatrix} = a^{K_{11}}_{\alpha,\beta}$$

Generate element matrices for all element $K_\ell, \ell = 1, 2 \cdots, L$ add its contribution to all pair of vertices (i, j). Note that T is the $3 \times L$ matrix whose ℓ-th column denotes the three *global* indices of **vertices** of ℓ-th element. For example, $T(\cdot, 11) = [7, 8, 12]^t$.

$$Z^T = \begin{bmatrix} 0 & 0.25 & 0.5 & 0.75 & 1.0 & 0 & 0.25 & \cdots \\ 0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.25 & 0.25 & \cdots \end{bmatrix} \tag{2.88}$$

$$T = \begin{bmatrix} 1 & 7 & 2 & 8 & \cdots & 12 & 7 & \cdots \\ 2 & 6 & 3 & 7 & \cdots & 11 & 8 & \cdots \\ 6 & 2 & 7 & 3 & \cdots & 7 & 12 & \cdots \end{bmatrix} \tag{2.89}$$

If we use CR element, we will need(later) a $3 \times L$ matrix E whose ℓ-th column denotes the three *global* indices of **edges** of ℓ-th element.

$$E = \begin{bmatrix} 1 & 7 & 2 & 9 & \cdots & 7 & \cdots \\ 6 & 14 & 8 & 15 & \cdots & 8 & \cdots \\ 2 & 6 & 7 & 8 & \cdots & 12 & \cdots \end{bmatrix} \tag{2.90}$$

Computation of element stiffness matrix

Let $K_\ell \in \mathcal{T}_h$ be a fixed element. Then $T(\alpha, \ell)$, $\alpha = 1, 2, 3$, are the global numbering of vertices of K_ℓ. The x_i-coordinates of vertices are $Z(T(\alpha, \ell), i), i = 1, 2$. We note that the local element matrix is given by

$$a^\ell_{\alpha,\beta} := a^{K_\ell}_{\alpha,\beta} = \int_{K_\ell} (p\nabla\phi_\alpha \cdot \nabla\phi_\beta + q\phi_\alpha\phi_\beta)dx + \int_{\bar{K}_\ell \cap \Gamma_2} p\sigma\phi_\alpha\phi_\beta \, ds,$$

$$b^\ell_\alpha = \int_{K_\ell} f\phi_\alpha \, dx - \sum_{j \in \mathcal{J}_1} g(N_j) a_{K_\ell}(\phi_j, \phi_\alpha) + \int_{\bar{K}_\ell \cap \Gamma_2} p\xi\phi_\alpha \, ds, \ \alpha = 1, 2, 3.$$

Then the assembly of global stiffness matrix is as follows:

Assembly of global stiffness matrix

Assume $\Gamma_2 = \emptyset$ for simplicity. Initially, set $A(i, j) = 0, \mathbf{F}(i) = 0, i, j = 1, \cdots, M$. The for $\ell = 1, \cdots, L$ do the following:

$$A_{(T(\alpha,\ell),T(\beta,\ell))} = A_{i,j} = \int_\Omega (p\nabla\phi_j \cdot \nabla\phi_i + q\phi_j\phi_i)dx = \sum_\ell a^\ell_{\beta\alpha}. \tag{2.91}$$

Since $(i, j) = (T(\alpha, \ell), T(\beta, \ell))$, the global index 7 corresponds to the local index of K_2, K_3, K_4, K_{11}, f K_{10}, K_9 are given in the following table.

The global index 8 corresponds to the local index K_4, etc. are also given. With the notation in (2.87) we see

Table 2.1: correspdence

global index	element					
7	K_2	K_3	K_4	K_{11}	K_{10}	K_9
7	1	3	2	1	3	2
8	K_4	K_5	K_6	K_{11}	K_{12}	K_{13}
8	1	3	2	2	3	1

$$\begin{aligned}A_{7,7} &= a^{K_2}(\phi_7,\phi_7)+a^{K_3}(\phi_7,\phi_7)+a^{K_4}(\cdot,\cdot)+a^{K_9}(\cdot,\cdot)+a^{K_{10}}(\cdot,\cdot)+a^{K^{11}}(\cdot,\cdot)\\ &= a_{11}^{K_2}+a_{33}^{K_3}+a_{22}^{K_4}+a_{22}^{K_9}+a_{33}^{K_{10}}+a_{11}^{K_{11}}=\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+1+1+\frac{1}{2}=4\\ A_{7,8} &= a_{12}^{K_4}+a_{21}^{K_{11}}=-\frac{1}{2}-\frac{1}{2}=-1.\end{aligned}$$

In other words, add contributions from L-th element stiffness matrix to its corresponding basis. For example, on the K_{11}, the entry $a_{\alpha\beta}^{K_{11}}$ has contribution to $a_{7,7}, a_{7,8}, a_{7,12}, a_{8,7}, a_{8,8}, a_{8,12}, a_{12,7}, a_{12,8}, a_{12,12}$. Hence the code looks like this:

for $\ell = 1, 2, \cdots, L$ (outer loop)
  for $\alpha = 1, 2, 3$ do
    for $\beta = 1, 2, 3$ do
      $A(T(\beta,\ell), T(\alpha,\ell)) = A(T(\beta,\ell), T(\alpha,\ell)) + a_{\alpha\beta}^{\ell}$
    end
    $\mathbf{F}(T(\alpha,\ell)) = \mathbf{F}(T(\alpha,\ell)) + b_{\alpha}^{\ell}$,
  end
end (L loop)

For FEM software see the National Institute of Standards and Technology (USA); http://www.netlib.org, http://gams.nist.gov
Finite element method (FEM) is a powerful and popular numerical method on solving partial differential equations (PDEs), with flexibility in dealing with complex geometric domains and various boundary conditions. MATLAB (Matrix Laboratory) is a powerful and popular software platform using matrix-based script language for scientific and engineering calculations. This project is on the development of an finite element method package in MATLAB based on an innovative programming style: sparse matrixlization. That is to reformulate algorithms in terms of sparse matrix operations to make use of the unique strength of MATLAB on fast matrix operations. iFEM, the resulting package, is a good balance between simplicity, readability, and efficiency. It will benefit not only education but also future research and algorithm development on finite element method.

This package can be downloaded from http://ifem.wordpress.com/

2.6 Mid term Take Home Exam

Mid term take home exam Due Oct 28

Exercise 2.6.1.

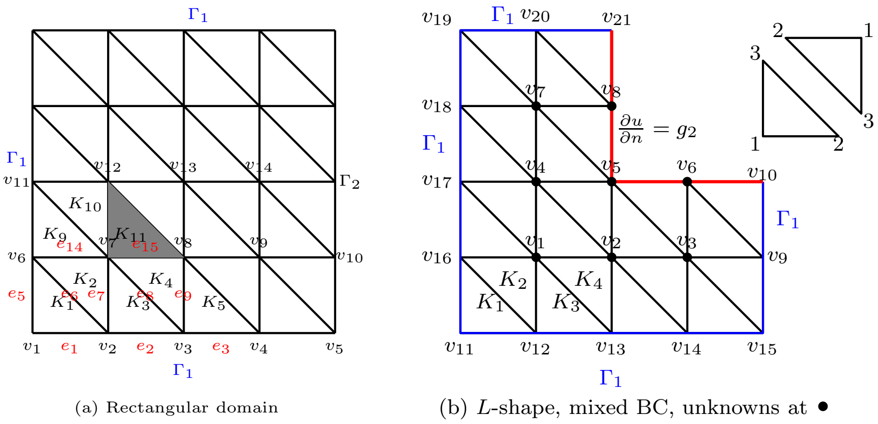


Figure 2.9: Label of elements and vertices

Refer to figure 2.8.1 (a). Write a FEM code for the problem (2.75 -2.77) when $\Omega = [-1, 1]^2$, with the following data. Use the uniform grids of $h = 2^{-k}, k = 2, 3, \cdots, 6$. For all problems, $q = \sigma = \xi = 0$.

(1) Compute the 3×3 element stiffness matrix $A_{K_{11}}$ for the case $p = 1 + x + 2y^2$ when $k = 2$(4×4 grid shown in the note above.)

(2) Print the entries of $F(i), i = 4, 5, 6, 7$ when $k = 2$ for $p = 1 + x + 2y^2$, $f = -(1 + 4y)$ and $g = 1 + x + y$ on the boundary.

(3) Solve the problem for the case $p = 1+x+2y^2$, $g = 0$ (Dirichlet condition on $\partial\Omega$) and $u = x(1 - x)y(1 - y)$.

(4) Solve the problem for the case $p = 1$, $u = 1 + x + 9y$, $f = 0$. $g = u|_{\partial\Omega}$ (Dirichlet condition on $\partial\Omega$). Do you find any special phenomena?

(5) Draw the graph of solution u_h of (3).

(6) Choose an exact solution of the form $u(x, y) = r^\gamma = (x^2 + y^2)^{\gamma/2}$ for (b) with appropriately chosen $\gamma > 0$.

To solve the linear system use either Gauss-Seidel method or conjugate gradient method. Report discrete L^2-norm defined by $\|u-u_h\|_h := \sqrt{h^2 \sum_i (u - u_h)^2(c_i)}$. Here c_i is the centroid of each element (triangle). Write the Table in a easily verifiable manner (systematically) for $h = 1/2^k$. Submit the paper report and the coding(by email).

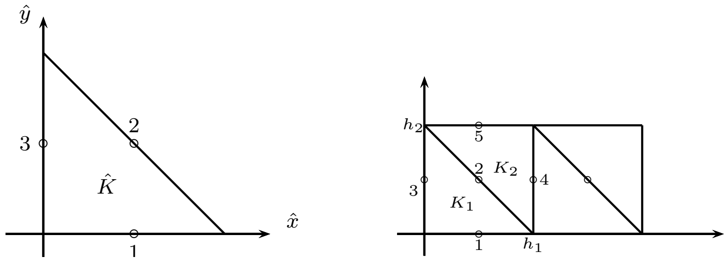


Figure 2.10: reference element and real domain

2.6.1 P_1 Nonconforming space of Crouzeix and Raviart

We introduce a P_1 nonconforming finite element method for $-\Delta u = f$. Assume a quasi uniform triangulation of the domain by triangles is given. Consider the space of all piecewise linear functions which is continuous only at mid point of edges. Here the degree of freedom is located at mid point of edges.

Let N_h be the space of all functions which is linear on each triangle and whose degrees of freedoms are determined

$$\begin{cases} u_h(m)|_L = u_h(m)|_R & \text{when } m \text{ is a mid point of interior edges} \\ u_h(m) = 0 & \text{when } m \text{ is a mid point of boundary edges} \end{cases}$$

Since u_h is discontinuous, the $a(u_h, v_h) = \int_\Omega \nabla u_h \cdot \nabla v_h \, dx$- is not well defined. So we define a discrete form a_h as follows:

$$a_h(u_h, v_h) = \sum_{T \in \mathcal{T}_h} \int_T \nabla u_h \cdot \nabla v_h \, dx \tag{2.92}$$

The P_1-nonconforming fem is: Find $u_h \in N_h$ such that

$$a_h(u_h, v_h) = f(v_h), \quad \forall v_h \in N_h.$$

Note that in general

$$a_h(u, v_h) \neq f(v_h).$$

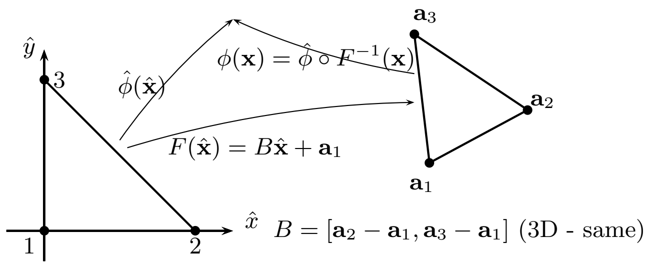


Figure 2.11: Reference triangle and the mapping

We define a discrete norm on N_h by

$$\|u_h\|_{1,h} = a_h(u_h, u_h)^{1/2}.$$

Assume the reference element $\hat{K} = K_1$ with $h_1 = h_2$. Then the basis functions are

$$\hat{\phi}_1 = 1 - 2y, \quad \hat{\phi}_2 = 2x + 2y - 1, \quad \hat{\phi}_3 = 1 - 2x. \tag{2.93}$$

Consider K_2. By mapping $F_{K_2}(\hat{\mathbf{x}}) = B_K\hat{\mathbf{x}} + \mathbf{b}$ we have

$$\begin{aligned} F_{K_2}\begin{pmatrix}\hat{x}\\ \hat{y}\end{pmatrix} &= \begin{pmatrix}0 & -h_1\\ h_2 & h_2\end{pmatrix}\begin{pmatrix}\hat{x}\\ \hat{y}\end{pmatrix} + \begin{pmatrix}h_1\\ 0\end{pmatrix} \\ \phi(\mathbf{x}) &= \hat{\phi} \circ (F_K^{-1}(\mathbf{x})) = \hat{\phi}\left[\frac{1}{h_1h_2}\begin{pmatrix}h_2 & h_1\\ -h_2 & 0\end{pmatrix}\begin{pmatrix}x - h_1\\ y\end{pmatrix}\right] \\ &= \hat{\phi}\begin{pmatrix}\frac{x-h_1}{h_1} + \frac{y}{h_2}\\ -\frac{x-h_1}{h_1}\end{pmatrix} \end{aligned}$$

According to mapping $\hat{\phi}_1$ corresp to ϕ_4, $\hat{\phi}_2$ corresp to ϕ_5 and $\hat{\phi}_3$ corresp to ϕ_2. Since

$$\hat{\phi}_1 = 1 - 2\hat{y}, \quad \hat{\phi}_2 = 2\hat{x} + 2\hat{y} - 1, \quad \hat{\phi}_3 = 1 - 2\hat{x} \text{ on } \hat{k} = \hat{K}_1$$

we have

$$\phi_4 = 1 + 2\frac{x - h_1}{h_1}, \quad \phi_5 = \frac{2x}{h_2} - 1, \quad \phi_2 = 1 - 2\frac{x + y - h_1}{h_1} \text{ on } K_2$$

Theorem 2.6.2 (Second Strang lemma)**.** *Under conditions given above, there* *exists a constant C independent of v_h such that*

$$\|u-u_h\|_h \leq C\left(\inf_{v_h\in V_h}\|u-v_h\|_h+\sup_{w_h\in V_h}\frac{|a_h(u,w_h)-f(w_h)|}{\|w_h\|_h}\right). \tag{2.94}$$

Proof. Let v_h be an arbitrary element in V_h. Then

$$\begin{aligned}\alpha\|u_h-v_h\|_h^2 &\leq a_h(u_h-v_h,u_h-v_h)\\ &= a(u-v_h,u_h-v_h)+f(u_h-v_h)-a_h(u,u_h-v_h)\\ &\leq M\|u-v_h\|_h\|u_h-v_h\|_h+|f(u_h-v_h)-a_h(u,u_h-v_h)|.\end{aligned}$$

So

$$\begin{aligned}\alpha\|u_h-v_h\|_h &\leq CM\|u-v_h\|_h+\frac{|f(u_h-v_h)-a_h(u,u_h-v_h)|}{\|w_h\|_h}\\ &\leq CM\|u-v_h\|_h+\sup_{w_h\in V_h}\frac{|f(w_h)-a_h(u,w_h)|}{\|w_h\|_h}.\end{aligned}$$

Now result follows from this and the triangle inequality

$$\|u-v_h\|_h\leq\|u-v_h\|_h+\|u_h-v_h\|_h.$$

□

2.6.2 Integration using reference element

In practice K is in a general position. Hence we show how to compute the integral

$$a_{ji}^K=\int_K(p\nabla\phi_i\cdot\nabla\phi_j+q\phi_i\phi_j)dxdy$$

through a mapping to a fixed "nice" reference element $\hat{K}$. Let

$$\mathbf{x}=\begin{pmatrix}x\\ y\end{pmatrix}=F(\hat{\mathbf{x}})=\begin{pmatrix}f_1(\hat{x},\hat{y})\\ f_2(\hat{x},\hat{y})\end{pmatrix} \tag{2.95}$$

be a one-to-one invertible map $\hat{K}\to K$, Then any function and $g(\mathbf{x})$ is related to a function $\hat{g}(\hat{\mathbf{x}})$ defined on the reference element $\hat{K}$ by

$$g(\mathbf{x})=g(F(\hat{\mathbf{x}}))=\hat{g}(\hat{\mathbf{x}}).$$

In particular, if it is affine then $F(\hat{\mathbf{x}}) = B\hat{\mathbf{x}} + \mathbf{b}$. For a scalar function g, we see

$$\int_K g(\mathbf{x})dxdy = \int_{\hat{K}} g(F(\hat{\mathbf{x}}))|J_K|d\hat{x}d\hat{y}, \tag{2.96}$$

where $J_K = det(DF_K) = det(B)$. But for a gradient of a function, it is more complicated: Noting that

$$\hat{\nabla}\hat{g} = B^T\nabla g, \tag{2.97}$$

we see

$$\int_K (p\nabla\phi_i \cdot \nabla\phi_j + q\phi_i\phi_j)dxdy = \int_{\hat{K}} (\hat{p}(B^{-T}\hat{\nabla}\hat{\phi}_i) \cdot (B^{-T}\hat{\nabla}\hat{\phi}_j) + \hat{q}\hat{\phi}_i\hat{\phi}_j)|J|d\hat{x}d\hat{y}, \tag{2.98}$$

where J is the determinant of B. To save computational cost in computing (2.98) we do as follows: A little of algebra shows that (2.98) is

$$a_{ji}^K = \int_{\hat{K}} \left\{ \frac{\hat{p}}{|J|} \left[E_1\hat{\phi}^i_{\hat{x}}\hat{\phi}^j_{\hat{x}} - E_2(\hat{\phi}^i_{\hat{y}}\hat{\phi}^j_{\hat{x}} + \hat{\phi}^i_{\hat{x}}\hat{\phi}^j_{\hat{y}}) + E_3\hat{\phi}^i_{\hat{y}}\hat{\phi}^j_{\hat{y}} \right] + |J|\hat{q}\hat{\phi}^i\hat{\phi}^j \right\} d\hat{x}d\hat{y},$$

where

$$J = x_{\hat{x}}y_{\hat{y}} - x_{\hat{y}}y_{\hat{x}}, \quad E_1 = x_{\hat{y}}^2 + y_{\hat{y}}^2 \tag{2.99}$$

$$E_2 = x_{\hat{x}}x_{\hat{y}} + y_{\hat{x}}y_{\hat{y}}, \quad E_3 = x_{\hat{x}}^2 + y_{\hat{x}}^2. \tag{2.100}$$

Things to consider

(1) Use banded storage

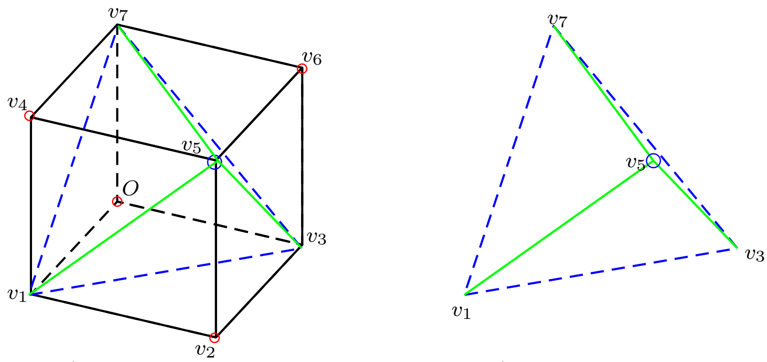
(2) Use as many modules as possible.

(3) Iterative method or direct method ?

(4) Output. Check the error by discrete L^2, H^1-inner product. Graphics.

2.7 Three dim cube

We provide a method to subdivide a 3D cube into tetrahedra to triangularize a box.



5 tetra w/ vertex O, v_4, v_2, v_6, v_5. Inner tetra w/vertices v_1, v_3, v_7, v_5

Figure 2.12: Subdivision of Cube into 5 tetrahedra

2.7.1 Subdivision into 5 tetrahedra

Here 4 sub tetra with red vertex (90 deg corner) plus the one interior (Green, blue).

From fig. 2.9, we see 5 Sub-tetra hedra with vertices at

$$(O, v_1, v_3, v_7),\ (v_4, v_1, v_5, v_7),\ (v_2, v_1, v_3, v_5),\ (v_6, v_3, v_7, v_5),\ (v_5, v_1, v_3, v_7)$$

2.7.2 Subdivision into 6 tetrahedra

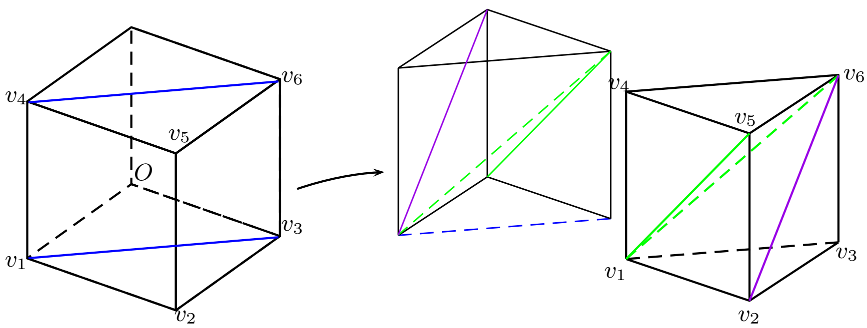


Figure 2.13: Subdivision of Cube into 6 tetrahedra

From fig. 2.10, we see 3 sub-tetra hedra with vertices at

$$(v_1, v_5, v_4, v_6),\ (v_1, v_2, v_5, v_6),\ (v_1, v_2, v_3, v_6)$$

2.8 Three dim cube -MG

Above subdivision of 3D cube into tetrahedra cannot(? at least difficult) be used to apply MG since the k-level of tetrahedra are obtained from k-level of box. Not easy to see nested ness of tetrahedra. So using the boxes are much easier to apply MG. Then the problem with IFEM. How to describe cutting the box with a surface ?

2.8.1 Numerical Integration

Abramowitz, Stegun. A software package in the public domain by Gautschi. We replace integral by certain weighted sum of function values:

$$I = \int_K g(x,y)dxdy \approx \sum_{i=1}^{Q} w_i g(x^i, y^i)$$

where w_i and (x^i, x^i) are independent of θ.

2.8.2 Quadrature for the interval $I = [-1, 1]$

$$\int_I g(x,y)dxdy \approx \sum_{i=1}^{Q} w_i g(x^i)$$

(1) (Gauss 2 pts). $x^i = -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$, weight $w_i = 1, 1$.

(2) (Gauss 3 pts). $x^i = -\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}$, weight $w_i = \frac{5}{9}, \frac{8}{9}, \frac{5}{9}$.

2.8.3 Quadrature for a triangle

We assume the reference triangle $\hat{K}$ is the right triangle with vertices at $(0,0)$, $(1,0)$ and $(0,1)$.

Example 2.8.1. $Q = 1$. (exact for P_1). Quadrature point for $\hat{K}$ is $(\frac{1}{3}, \frac{1}{3})$, $w = \frac{1}{2}$.

$$\int_K gdxdy \approx |K| g(\frac{1}{3}, \frac{1}{3}).$$

Example 2.8.2. $Q = 3$. (exact for P_2). Quadrature points are $(\frac{1}{2}, 0)$, $(\frac{1}{2}, \frac{1}{2})$ and $(0, \frac{1}{2})$ $w = \frac{1}{6}$. Thus

$$\int_K gdxdy \approx \frac{|K|}{3} \left[g(\frac{1}{2}, 0) + g(\frac{1}{2}, \frac{1}{2}) + g(0, \frac{1}{2}) \right].$$

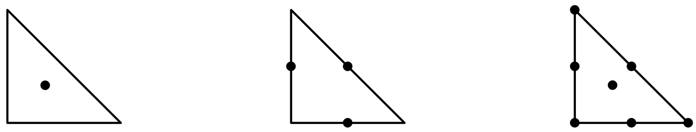


Figure 2.14: Quadrature points for triangle

Another one: Gaussian quadrature of degree 2 for the standard triangle

$$\int_K g dx dy \approx \frac{|K|}{3}\left[g(\frac{1}{6},\frac{1}{6}) + g(\frac{2}{3},\frac{1}{6}) + g(\frac{1}{6},\frac{2}{3})\right].$$

Example 2.8.3 (Quadrature for triangle)**.** $Q = 7$ (exact for P_3).

$$\frac{|K|}{60}\left(3\sum_{i=1}^{3} g(v_i) + 8\sum_{i<j}^{3} g(v_{ij}) + 27g(v_{123})\right).$$

Another $Q = 4$

$$-\frac{27}{96}g(\frac{1}{3},\frac{1}{3}) + \frac{25}{96}\left[g(\frac{2}{15},\frac{11}{15}) + g(\frac{2}{15},\frac{2}{15}) + g(\frac{11}{15},\frac{2}{15})\right].$$

Here $\frac{2}{15}, \frac{11}{15}$ can be replaced by $\frac{1}{5}, \frac{3}{5}$ resp.

2.8.4 Quadrature for a Rectangle

Example 2.8.4. $\hat{Q} = [-1,1] \times [-1,1]$.

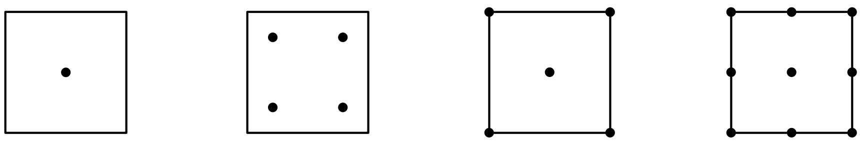
(1) $Q = 1$. (Gaussian quadrature) The point is $(0,0)$ and $w = 4$. It is exact for Q_1.

(2) $Q = 4$. (Product of quadrature) The points are $(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$ and $w = 1$. Exact for Q_3.

(3) $Q = 5$.

$$\frac{1}{3}\left[g(-1,-1) + g(-1,1) + g(1,-1) + g(1,1) + 8g(0,0)\right]$$

is exact for P_3.



$Q = 1(Q_1)$ $Q = 4(Q_3)$ Gaussian $Q = 5(P_3)$ $Q = 9(Q_3)$ Simpson

Figure 2.15: Quadrature points for the rectangle

(4) $Q = 9$. (Product of Simpson's rule).

$$\frac{1}{9}\left[\sum g(\pm 1, \pm 1) + 4\sum (g(\pm\frac{1}{2}, 0) + g(0, \pm\frac{1}{2})) + 16g(0,0)\right]$$

is exact for Q_3.

(5) $Q = 9$. (Product of Gauss 3 pts). $x_{ij} = -\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}$, weight $w_{ij} = \frac{5}{9}, \frac{8}{9}, \frac{5}{9}$.

$$\left[\sum_{i,j=1}^{3} w_{ij} g(x_{ij})\right]$$

is exact for Q_5.

2.8.5 Quadrature for a circle

Example 2.8.5 (Quadrature for the circle of radius h)**.** $Q = 4$(exact up to Q_3).

$$\iint_C g(x,y)dxdy = \pi h^2 \sum_{i=1}^{4} w_i g(x_i, y_i) + O(h^4),$$

where $(x_i, y_i) = (\pm h/2, \pm h/2)$ and $w_i = 1/4$.

Example 2.8.6 (Quadrature for the circle of radius h)**.** $Q = 7$(exact to P_5)

$$(x_1, y_1) = (0,0) \qquad (x_i, y_i) = (\pm\sqrt{2/3}, 0), \; i = 2, 3 \tag{2.101}$$

$$(x_i, y_i) = (\pm h/\sqrt{6}, \pm h/\sqrt{2}), \; i = 4, \cdots, 7 \tag{2.102}$$

$w_1 = 1/4, w_i = 1/8, i = 2:7.$

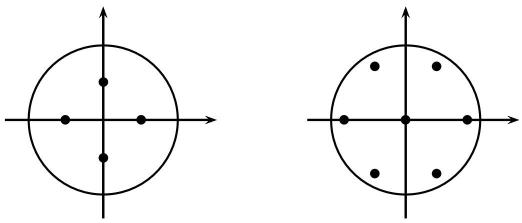


Figure 2.16: Quadrature points for the circle

2.9 Stokes Equation

Notations.

$$\begin{array}{rcl} \mathrm{curl}(\mathbf{curl}\,\phi) & = & -\Delta\phi, \quad n=2 \\ \left.\begin{array}{r} \mathbf{curl}\,(\mathbf{curl}\,\mathbf{v}) \\ \mathbf{curl}\,(\mathrm{curl}\mathbf{v}) \end{array}\right\} & = & -\Delta\mathbf{v} + \mathrm{grad}\,(\nabla\mathbf{v}) \quad \begin{array}{r} n=3 \\ n=2. \end{array} \end{array}$$

We define

$$\mathbf{grad}\,p = \begin{pmatrix} \partial p/\partial x_1 \\ \partial p/\partial x_2 \end{pmatrix}, \quad \mathbf{div}\,\boldsymbol{\tau} = \begin{pmatrix} \partial\tau_{11}/\partial x_1 + \partial\tau_{12}/\partial x_2 \\ \partial\tau_{21}/\partial x_1 + \partial\tau_{22}/\partial x_2 \end{pmatrix},$$

$$\mathrm{div}\,\mathbf{v} = \partial v_1/\partial x_1 + \partial v_2/\partial x_2, \quad \mathbf{Grad}\,\mathbf{v} = \begin{pmatrix} \partial v_1/\partial x_1 & \partial v_1/\partial x_2 \\ \partial v_2/\partial x_1 & \partial v_2/\partial x_2 \end{pmatrix}.$$

$$\mathrm{curl}\mathbf{v} = \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}, \qquad \mathbf{curl}\,\eta = \left(\tfrac{\partial\eta}{\partial y}, -\tfrac{\partial\eta}{\partial x}\right)$$

$$\mathbf{curl}\,\mathbf{v} = \begin{pmatrix} \mathbf{curl}\,v_1 \\ \mathbf{curl}\,v_2 \end{pmatrix}.$$

We also define

$$\boldsymbol{\delta} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\chi} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad \mathrm{tr}(\boldsymbol{\tau}) = \boldsymbol{\tau} : \boldsymbol{\delta}.$$

For two matrices $\boldsymbol{\tau}, \boldsymbol{\delta}$ we write

$$\boldsymbol{\tau} : \boldsymbol{\delta} = \sum_{i=1}^{2}\sum_{j=1}^{2} \delta_{ij}\tau_{ij}.$$

Also define *deviatoric or deformation* tensor

$$\boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T), \text{ and } \boldsymbol{\epsilon}_{ij}(\mathbf{u}) = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}).$$

Let Ω be a domain in $\mathbb{R}^n (n = 2, 3)$ with its boundary $\Gamma := \partial\Omega$. The Navier-Stokes equations for a viscous fluid is are as follows:

$$\begin{aligned} \frac{\partial u_i}{\partial t} + \sum_{j=1}^{n} u_j \frac{\partial u_i}{\partial x_j} - 2\nu \sum_j \frac{\partial \boldsymbol{\epsilon}_{ij}(\mathbf{u})}{\partial x_j} + \frac{\partial p}{\partial x_i} &= f_i (1 \le i \le n) \text{ in } \Omega && (2.103)\\ \operatorname{div}\mathbf{u} &= 0 \text{ (incompressible)} && (2.104)\\ \mathbf{u} &= \mathbf{g} \text{ on } \Gamma. && (2.105) \end{aligned}$$

Here $\mathbf{u}$ is the velocity of the fluid, $\nu > 0$ is the viscosity and p is the pressure; (Here we assume p and ν are normalized so we may assume $\rho = 1$) and the vector $\mathbf{f}$ represents body forces per unit mass. If we introduce the stress tensor $\sigma_{ij} := -p\delta_{ij} + 2\nu\boldsymbol{\epsilon}_{ij}(\mathbf{u})$ we have a simpler form :

$$\begin{aligned} \tfrac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \operatorname{div}\boldsymbol{\sigma} &= \mathbf{f} \text{ in } \Omega,\\ \operatorname{div}\mathbf{u} &= 0 \text{ in } \Omega,\\ \mathbf{u} &= \mathbf{g} \text{ on } \Gamma. \end{aligned} \qquad (2.106)$$

Here the first term is interpreted as

$$(\mathbf{u}\cdot\nabla)\mathbf{v} = \mathbf{e}_i \sum_j u_j \frac{\partial v_i}{\partial x_j} = \sum_j u_j \frac{\partial \mathbf{v}}{\partial x_j}.$$

Note that if $\operatorname{div}\mathbf{u} = 0$, the following identity holds

$$\sum_j \frac{\partial \boldsymbol{\epsilon}_{ij}(\mathbf{u})}{\partial x_j} = \frac{1}{2}\sum_j \left(\frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial^2 u_j}{\partial x_i \partial x_j}\right) = \frac{1}{2}\Delta u_i, \quad \text{for each } i \qquad (2.107)$$

so that the equation can be written as

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \mathbf{grad}\, p = \mathbf{f} \text{ in } \Omega,\\ \operatorname{div}\mathbf{u} = 0 \text{ in } \Omega,\\ \mathbf{u} = \mathbf{g} \text{ on } \Gamma. \end{cases} \qquad (2.108)$$

We only consider the steady-state case and assume that $\mathbf{u}$ is so small that we can ignore the non-linear convection term $u_j \frac{\partial u_i}{\partial x_j}$. Thus, we have the Stokes equation:

$$\begin{cases} -\nu\Delta\mathbf{u} + \mathbf{grad}\, p &= \mathbf{f} \text{ in } \Omega, \\ \operatorname{div}\mathbf{u} &= 0 \text{ in } \Omega, \\ \mathbf{u} &= \mathbf{g} \text{ on } \Gamma. \end{cases} \tag{2.109}$$

2.10 Newton for NS System

Steady state Navier Stokes equation is following

$$\begin{cases} -\mu\Delta\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p & = \mathbf{f} \quad \text{on} \quad \Omega \\ \nabla\cdot\mathbf{u} & = g \quad \text{on} \quad \Omega \\ \mathbf{u} & = \mathbf{q} \quad \text{on} \quad \partial\Omega. \end{cases}$$

Newton iteration is following. Initial guess $\mathbf{u}^0 = 0, p^0 = 0$ are given. Suppose $\mathbf{u}^n$ and p^n is given. We do the Newton iteration until the residual is smaller than the tolerance.

Suppose $\mathbf{u}^{n+1} = \mathbf{u}^n + \tilde{\mathbf{u}}$ and $p^{n+1} = p^n + \tilde{p}$.

$$\begin{cases} -\mu\Delta(\mathbf{u}^n + \tilde{\mathbf{u}}) + (\mathbf{u}^n + \tilde{\mathbf{u}})\cdot\nabla(\mathbf{u}^n + \tilde{\mathbf{u}}) + \nabla(p^n + \tilde{p}) & = \mathbf{f} \\ \nabla\cdot(\mathbf{u}^n + \tilde{\mathbf{u}}) & = g \end{cases}$$

By omitting high order term respect to $\tilde{\mathbf{u}}$ we obtain linear system for $\tilde{\mathbf{u}}$

$$\begin{cases} -\mu\Delta\tilde{\mathbf{u}} + \mathbf{u}^n\cdot\nabla\tilde{\mathbf{u}} + \tilde{\mathbf{u}}\cdot\nabla\mathbf{u}^n + \nabla\tilde{p} & = r_1^n \\ \nabla\cdot\tilde{\mathbf{u}} & = r_2^n, \end{cases}$$

where we define residual $\mathbf{r}^n$ of the system of the residual at $n-th$ iteration by $\mathbf{r}^n = (r_1^n, r_2^n) = (f + \mu\Delta\mathbf{u}^n - \mathbf{u}^n\cdot\nabla\mathbf{u}^n - \nabla p^n, g - \nabla\cdot\mathbf{u}^n)$. Or write it as

$$\begin{aligned} -\mu\Delta\mathbf{u}^{n+1} + \mathbf{u}^n\cdot\nabla\mathbf{u}^{n+1} + \mathbf{u}^{n+1}\cdot\nabla\mathbf{u}^n + \nabla p^{n+1} &= \mathbf{f} + \mathbf{u}^n\cdot\nabla\mathbf{u}^n \\ \nabla\cdot\mathbf{u}^{n+1} &= 0 \\ \mathbf{u}^{n+1} &= 0 \text{ on } \Gamma. \end{aligned}$$

We discretize the problem on finite element space $M_h = V_h \times Q_h$. V_h may or may not be subspace of $H_0^1(\Omega)$, and Q_h is subspace of $L^2(\Omega)\backslash\{0\}$. For velocity space V_h we use Q1-nonconforming element and pressure space Q_h we use P0 conforming element. After fully discretizing the system we obtain linear system

$$\begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_h \\ \tilde{p}_h \end{pmatrix} = \begin{pmatrix} r_{1,h}^n \\ r_{2,h}^n . \end{pmatrix} \tag{2.110}$$

with nonsymmetric matrix A_h. When μ gets smaller it means nonlinear term of the Navier-Stokes equation dominates. At the discretized level, the A_h gets highly nonsymmetric and system gets difficult to solve. The system (2.111) can be solved by Uzawa method.

$[\mathbf{u}^{n+1}, p^{n+1}] = \mathbf{Newton}[\mathbf{u}^n, p^n]$

(1) Solve for $\tilde{\mathbf{u}}$ and $\mathbf{p}$ with given $(\mathbf{u}^n, p^n)$

$$\begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_h \\ \tilde{p}_h \end{pmatrix}. = \begin{pmatrix} r_{1,h}^n \\ r_{2,h}^n \end{pmatrix}. \tag{2.111}$$

(2) Compute $(\mathbf{u}^{n+1}, p^{n+1}) = (\mathbf{u}^n + \tilde{\mathbf{u}}, p^n + \tilde{p})$

Process Newton iteration until $\mathbf{r}^n$ is smaller than tolerance.

2.11 The Equations of Elasticity

Notations: Let $\mathbf{H} = (H^1(\Omega))^3$ and $\mathbf{H}_0 = (H_0^1(\Omega))^3$. For $\mathbf{u} = (u_1, u_2, u_3) \in \mathbf{H}$, we let

$$\operatorname{div} \mathbf{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}, \ \|(u_1, u_2, u_3)\|_{\mathbf{H}}^2 = \sum_i \|u_i\|_{H_0^1}^2.$$

Let

$$\boldsymbol{\epsilon}_{ij}(\mathbf{v}) = \frac{1}{2}(\partial_j v_i + \partial_i v_j) = \frac{1}{2}(\nabla \mathbf{v} + \nabla \mathbf{v}^T) \tag{2.112}$$

be the linearized strain tensor and the stress tensor be

$$\sigma_{ij}(\mathbf{v}) = \lambda \left(\sum_{k=1}^{3} \boldsymbol{\epsilon}_{kk}(\mathbf{v}) \right) \delta_{ij} + 2\mu\epsilon_{ij}(\mathbf{v}), (1 \le i, j \le 3) \tag{2.113}$$

or simply

$$\boldsymbol{\sigma}(\mathbf{v}) = 2\mu\boldsymbol{\epsilon}(\mathbf{v}) + \lambda tr(\boldsymbol{\epsilon}(\mathbf{v}))\boldsymbol{\delta}. \tag{2.114}$$

Stress is defined as force per unit area.

We use the following notation (matrix dot product)

$$\boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}) = \sum_{i,j=1}^{3} \epsilon_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{v}).$$

Pure displacement problem

The equation of elasticity in pure displacement problem is

$$\begin{aligned} -\mathbf{div}\,\{2\mu\boldsymbol{\epsilon}(\mathbf{u}) + \lambda\, tr(\boldsymbol{\epsilon}(\mathbf{u}))\boldsymbol{\delta}\} = \mathbf{f}, &\quad \text{in } \Omega, \\ \mathbf{u} = \mathbf{g}, &\quad \text{on } \partial\Omega, \end{aligned} \tag{2.115}$$

where λ and μ are Lamé constants. Since $\mathbf{div}\,(\nabla\mathbf{u})^T = \mathbf{div}\,((\mathrm{div}\,\mathbf{u})\boldsymbol{\delta})$, the first equation of (2.115) becomes

$$-\mathbf{div}\,\{\mu\nabla\mathbf{u} + (\mu+\lambda)(\mathrm{div}\,\mathbf{u})\boldsymbol{\delta}\} = \mathbf{f}. \tag{2.116}$$

Green's formula

For any tensor $\boldsymbol{\sigma} = (\sigma_{ij})$ we have

$$-\int_\Omega (\partial_j\sigma_{ij})v_i\,d\mathbf{x} \;=\; \int_\Omega \sigma_{ij}\partial_j v_i\,d\mathbf{x} - \int_\Gamma \sigma_{ij}v_i\nu_j\,ds. \tag{2.117}$$

If $\sigma_{ij} = \partial_j u_i$ then we have

$$-\int_\Omega (\partial_j^2 u_i)v_i\,d\mathbf{x} \;=\; \int_\Omega \partial_j u_i\partial_j v_i\,d\mathbf{x} - \int_\Gamma \partial_j u_i v_i\nu_j\,ds. \tag{2.118}$$

Summation give

$$-\int_\Omega \mathrm{div}\,(\nabla\mathbf{u})\cdot\mathbf{v}\,d\mathbf{x} \;=\; \int_\Omega \nabla\mathbf{u} : \nabla\mathbf{v}\,d\mathbf{x} - \int_\Gamma \mathbf{v}^T\nabla\mathbf{u}\cdot\mathbf{n}\,ds. \tag{2.119}$$

If, on the other hand, $\boldsymbol{\sigma} = (\mathrm{div}\,\mathbf{u})\,\boldsymbol{\delta}$, then we have (after Einstein summation notation is used)

$$\begin{aligned} -\int_\Omega (\partial_j(\mathrm{div}\,\mathbf{u})\delta_{ij})v_i\,d\mathbf{x} &= \int_\Omega \mathrm{div}\,\mathbf{u}\,\delta_{ij}\partial_j v_i\,d\mathbf{x} - \int_\Gamma \mathrm{div}\,\mathbf{u}\,\delta_{ij}v_i\nu_j\,ds \\ &= \int_\Omega \mathrm{div}\,\mathbf{u}\,\mathrm{div}\,\mathbf{v}\,d\mathbf{x} - \int_\Gamma \mathrm{div}\,\mathbf{u}(\mathbf{v}\cdot\mathbf{n})\,ds \\ &= \int_\Omega \mathrm{div}\,\mathbf{u}\,\mathrm{div}\,\mathbf{v}\,d\mathbf{x} - \int_\Gamma tr(\boldsymbol{\epsilon}(\mathbf{u}))\mathbf{v}\cdot\mathbf{n}\,ds \end{aligned} \tag{2.120}$$

Hence by (2.119) and (2.120), the weak form for the pure displacement problem is: find $\mathbf{u} \in \mathbf{H}$ satisfying the BC and for all $\mathbf{v} \in \mathbf{H}_0$

$$a(\mathbf{u}, \mathbf{v}) := \mu \int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v}\, dx + (\mu + \lambda) \int_\Omega \operatorname{div} \mathbf{u} \operatorname{div} \mathbf{v}\, dx = f(\mathbf{v}). \qquad (2.121)$$

Pure traction problem

The pure traction problem is

$$\begin{aligned} -\mathbf{div}\, \{2\mu \boldsymbol{\epsilon}(\mathbf{u}) + \lambda\, tr(\boldsymbol{\epsilon}(\mathbf{u}))\boldsymbol{\delta}\} = \mathbf{f}, &\quad \text{in } \Omega, \\ \boldsymbol{\sigma}(\mathbf{u}) \cdot \mathbf{n} = \mathbf{g}, &\quad \text{on } \partial\Omega = \Gamma, \end{aligned} \qquad (2.122)$$

with compatibility condition:

$$\int_\Omega \mathbf{f} \cdot \mathbf{v} dx + \int_{\partial\Omega} \mathbf{g} \cdot \mathbf{v} ds = 0, \text{ for } \mathbf{v} \in RM := \{(a + by, c - bx)\}.$$

Here

$$\boldsymbol{\sigma}(\mathbf{u}) = 2\mu \boldsymbol{\epsilon}(\mathbf{u}) + \lambda\, tr(\boldsymbol{\epsilon}(\mathbf{u}))\boldsymbol{\delta}.$$

Multiply $\mathbf{v} \in (H^1(\Omega))^n$ and integrate by part (use (2.117) with $\boldsymbol{\epsilon}$ in place of $\boldsymbol{\sigma}$) to the first term of (2.122), we see

$$\begin{aligned} & -2\mu \int_\Omega \sum_{i,j} \frac{\partial \epsilon_{ij}(\mathbf{u})}{\partial x_j} v_i \\ = \; & -2\mu \int_{\partial\Omega} \sum_{i,j} \epsilon_{ij}(\mathbf{u}) n_j v_i + 2\mu \int_\Omega \sum_{i,j} \epsilon_{ij}(\mathbf{u}) \frac{\partial v_i}{\partial x_j}. \end{aligned}$$

Using symmetry of $\boldsymbol{\epsilon}_{ij}(\mathbf{u})$, we get

$$\sum_{i,j} \epsilon_{ij}(\mathbf{u}) \frac{\partial v_j}{\partial x_i} = \sum_{i,j} \epsilon_{ji}(\mathbf{u}) \frac{\partial v_j}{\partial x_i} = \sum_{i,j} \epsilon_{ij}(\mathbf{u}) \frac{\partial v_i}{\partial x_j}. \qquad (2.123)$$

Hence

$$\begin{aligned}
& -2\mu\int_\Omega \sum_{i,j} \frac{\partial \epsilon_{ij}(\mathbf{u})}{\partial x_j} v_i \\
= & -2\mu\int_{\partial\Omega} \sum_{i,j} \epsilon_{ij}(\mathbf{u}) n_j v_i + \mu\int_\Omega \sum_{i,j} \epsilon_{ij}(\mathbf{u})(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}) \\
= & -2\mu\int_{\partial\Omega} \sum_{i,j} \epsilon_{ij}(\mathbf{u}) n_j v_i + 2\mu\int_\Omega \sum_{i,j} \epsilon_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{v}) \\
= & -2\mu\int_{\partial\Omega} \boldsymbol{\epsilon}(\mathbf{u})\mathbf{n}\cdot\mathbf{v} + 2\mu\int_\Omega \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}). \qquad (2.124)
\end{aligned}$$

Meanwhile the second term of (2.122) gives

$$\begin{aligned}
& \lambda\int_\Omega \operatorname{div}\mathbf{u}\operatorname{div}\mathbf{v}\,dx - \lambda\int_{\partial\Omega} tr(\boldsymbol{\epsilon}(\mathbf{u}))\boldsymbol{\delta}\mathbf{n}\cdot\mathbf{v}\,ds \\
& = \lambda\int_\Omega \operatorname{div}\mathbf{u}\operatorname{div}\mathbf{v}\,dx - \lambda\int_{\partial\Omega} tr(\boldsymbol{\epsilon}(\mathbf{u}))\mathbf{v}\cdot\mathbf{n}\,ds. \qquad (2.125)
\end{aligned}$$

Hence we get

$$a(\mathbf{u},\mathbf{v}) := 2\mu\int_\Omega \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v})\,dx + \lambda\int_\Omega \operatorname{div}\mathbf{u}\operatorname{div}\mathbf{v}\,dx = \bar{\mathbf{f}}(\mathbf{v}), \qquad (2.126)$$

where

$$\bar{\mathbf{f}}(\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v}\,dx + \int_\Gamma \mathbf{g}\cdot\mathbf{v}\,ds. \qquad (2.127)$$

Notice that the boundary terms of (2.124) and (2.125) are included in rhs as $\mathbf{g}$. Thus formally we are solving

$$-\sum_{j=1}^{3}\int_\Omega (\partial_j\sigma_{ij}(\mathbf{u}))v_i = \int_\Omega \bar{f}_i v_i\,d\mathbf{x},\ \ i = 1,2,3. \qquad (2.128)$$

This together with the BC, we can check the weak form is equivalent to the pure traction case.

Compatibility condition: use Fredholm alternative (or duality in Banach operator):

$$R(A)^\perp = N(A^*).$$

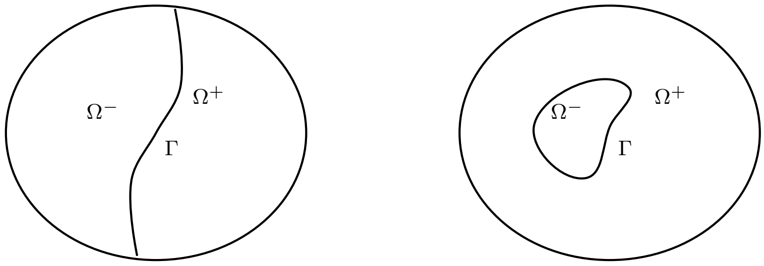


Figure 2.17: A domain Ω for the interface problem

$$\begin{aligned}\bar{\mathbf{f}}(\mathbf{v}) = 0 \Leftrightarrow A^*\mathbf{v} = 0 \Leftrightarrow \\ \Leftrightarrow \boldsymbol{\epsilon}(\mathbf{v}) = 0 \text{ and } \operatorname{div}\mathbf{v} = 0 \Leftrightarrow \mathbf{v} \in RM.\end{aligned}$$

If $\mathbf{v} \in RM$ then $\boldsymbol{\epsilon}(\mathbf{v}) = \operatorname{div}\mathbf{v} = 0$. Hence $\int_\Omega \mathbf{f}\cdot\mathbf{v}dx + \int_{\partial\Omega}\mathbf{g}\cdot\mathbf{v}ds = 0$. Conversely, Let $\boldsymbol{\epsilon}(\mathbf{v}) = 0$ and $\operatorname{div}\mathbf{v} = 0$. Then

$$\begin{pmatrix} 2v_{11} & v_{12}+v_{21} \\ v_{12}+v_{21} & 2v_{22}\end{pmatrix} = 0 \text{ and } v_{11}+v_{22} = 0.$$

Hence

$$v_1 = g_1(y) + a,\ \ v_2 = g_2(x) + c$$

for some constant a, c and ftns g_1, g_2 (not $\mathbf{g}$ above.) Further, $v_{12} + v_{21} = 0$ implies

$$g_1'(y) + g_2'(x) = 0$$

and hence

$$g_1'(y) = b = -g_2'(x).$$

This shows the compatibility condition holds precisely when $\mathbf{v} \in RM := \{(a + by, c - bx)\}$.

Exercise 2.11.1. What happens to the bilinear form if the domain consists of two connected parts $\Omega = \Omega^+ \cup \Omega^-$ where the Lamé constants are different on each of the domain, i.e., $\mu = \mu\pm,\ \lambda = \lambda\pm$ on $\Omega^\pm$? Also state an appropriate conditions.

Relation to Stokes equation

Meanwhile if we introduce $p = -\lambda tr(\boldsymbol{\epsilon}(\mathbf{u}))$ then

$$\begin{aligned} -\mathbf{div}\,\{2\mu\boldsymbol{\epsilon}(\mathbf{u})\} + \nabla p = \mathbf{f}, &\quad \text{in } \Omega,\\ \mathbf{div}\,\mathbf{u} = -\frac{p}{\lambda}, &\quad \text{in } \Omega, \\ \mathbf{u} = \mathbf{g}, &\quad \text{on } \partial\Omega. \end{aligned} \tag{2.129}$$

This is a mixed form. If $\lambda \to \infty$ then $\mathbf{div}\,\mathbf{u} = 0$ hence we get Stoke problem.

2.12 Final take home exam - Due Dec 24

Consider the Stokes equation

$$\begin{cases} -\nu\Delta\mathbf{u} + \mathbf{grad}\,p &= \mathbf{f} \text{ in } \Omega = [0,1]^2, \\ \operatorname{div}\mathbf{u} &= 0 \text{ in } \Omega, \\ \mathbf{u} &= \mathbf{g} \text{ on } \Gamma, \end{cases} \tag{2.130}$$

The FEM form is (Q_h): find a pair (u_h, p_h) in $X_h \times M_h$ such that

$$\begin{aligned} a(\mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h) = (\mathbf{f}, \mathbf{v}_h) - a(\mathbf{u}_g, \mathbf{v}), &\quad \mathbf{v} \in X_h \\ b(\mathbf{u}_h, q) = 0, &\quad q \in M_h \end{aligned}$$

Normalize p_h so that it belongs to $M_h = L^2_0(\Omega)$.

Here spaces are P_1 nonconforming FEM of Crouzeix-Raviart for velocity and piecewise constant for pressure.

(1) Choose your own exact solution pair $(\mathbf{u}, p)$ which satisfies $\operatorname{div}\mathbf{u} = 0$ and obtain $\mathbf{f}$ and $\mathbf{g}$ by plugging them into the equation. For example, you can choose any smooth scalar function $\phi(x, y)$ and take $\operatorname{curl}\phi$. Solve it by FEM.

  (a) Report $\|u - u_h\|_0$, $\|u - u_h\|_1$ and $\|p - p_h\|_0$ for $h = 1/4, 1/8, \cdots, 1/2^6$. Write down the table. Make sure you are getting correct answer.

  (b) Draw the graph of one component, say u_1 of the velocity field $\mathbf{u}_h = (u_1, u_2)$.

(2) Solve the Stokes equation with $\mathbf{f} = 0$ and the following BC and various

$\nu = 1.0, 0.1, 0.01, 0.005.$

$$\mathbf{g} = (g_1, g_2) = (const, 0) \text{ on } y = 1, \mathbf{g} = 0 \text{ otherwise}$$

(a) Draw the stream line (Use matlab or any graphic package) for each case $\nu = 1.0, 0.1, 0.01, 0.005$ or smaller.

Solver - Uzawa method

Solve it by Standard Uzawa method. Try various $\epsilon > 0$ see which one is good. Let p_h^0 given. Solve for $m = 0, 1, \cdots$, until $\|p_h^{m+1} - p_h^m\|$ is sufficiently small.

$$\begin{aligned} a(\mathbf{u}_h^{m+1}, \mathbf{v}) + b(\mathbf{v}, p_h^m) &= (\mathbf{f}, \mathbf{v}_h) - a(\mathbf{u}_g, \mathbf{v}), \quad \mathbf{v} \in X_h \\ b(\mathbf{u}_h^{m+1}, q) &= \frac{1}{\epsilon}(p_h^{m+1} - p_h^m, q), \quad q \in M_h \end{aligned}$$

Normalize p_h^{m+1} each step so that it belongs to $M = L_0^2(\Omega)$.