Numerics Exercise 3

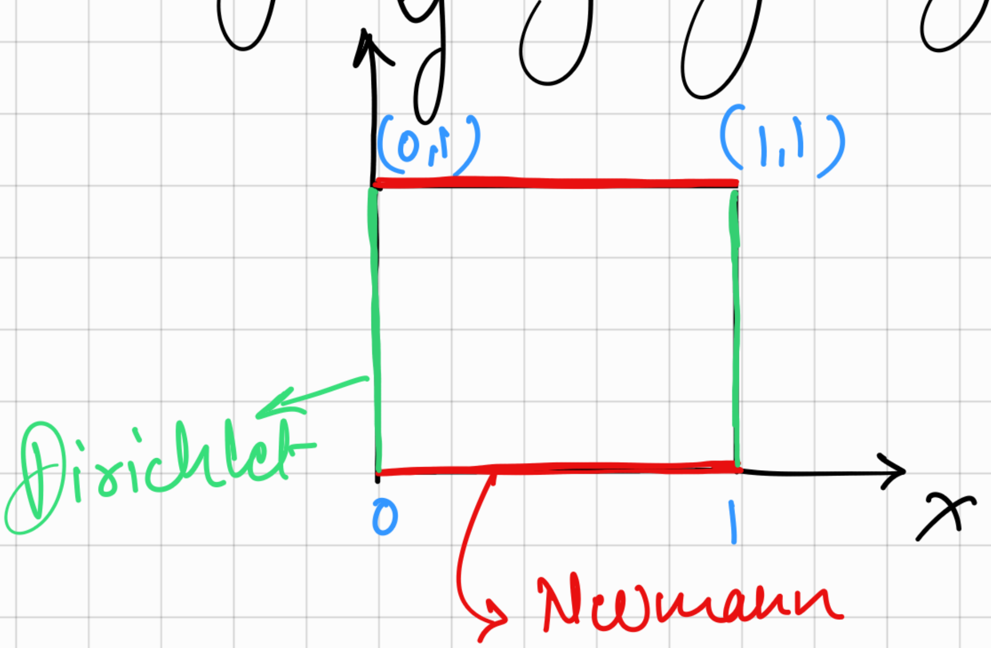
Consider Poisson System

$$-\nabla^2 u = f \quad \text{on } (x,y) \in [0,1]^2$$

$$f(x,y) = 2\pi^2 \sin(\pi x)\cos(\pi x)$$

Dirichlet BC $\rightarrow$ $u(0,y) = u(1,y) = 0$

Neumann B.C $\rightarrow$ $\left.\dfrac{\partial u(x,y)}{\partial y}\right|_{y=0} = \left.\dfrac{\partial (x,y)}{\partial y}\right|_{y=1} = 0$



The exact solution is

$$u_e(x,y) = \sin(\pi x)\cos(\pi y)$$

$\Rightarrow$ 2d - Poisson equation,

$$-\nabla^2\phi(x,y) = f(x,y)$$

$$\Rightarrow f(x,y) = -\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 2\pi^2 \sin\pi x \cos\pi y$$

$\phi(x,y) \rightarrow$ unknown function

$f(x,y) \rightarrow$ given function

$\Rightarrow$ To check the exact solution,

$$\frac{\partial u_e}{\partial x} = \pi\cos(\pi x)\cos(\pi y)$$

$$\frac{\partial^2 u_e}{\partial x^2} = -\pi^2 \sin(\pi x)\cos(\pi y)$$

$$\frac{\partial u_e}{\partial y} = -\pi \sin(\pi x)\sin(\pi y)$$

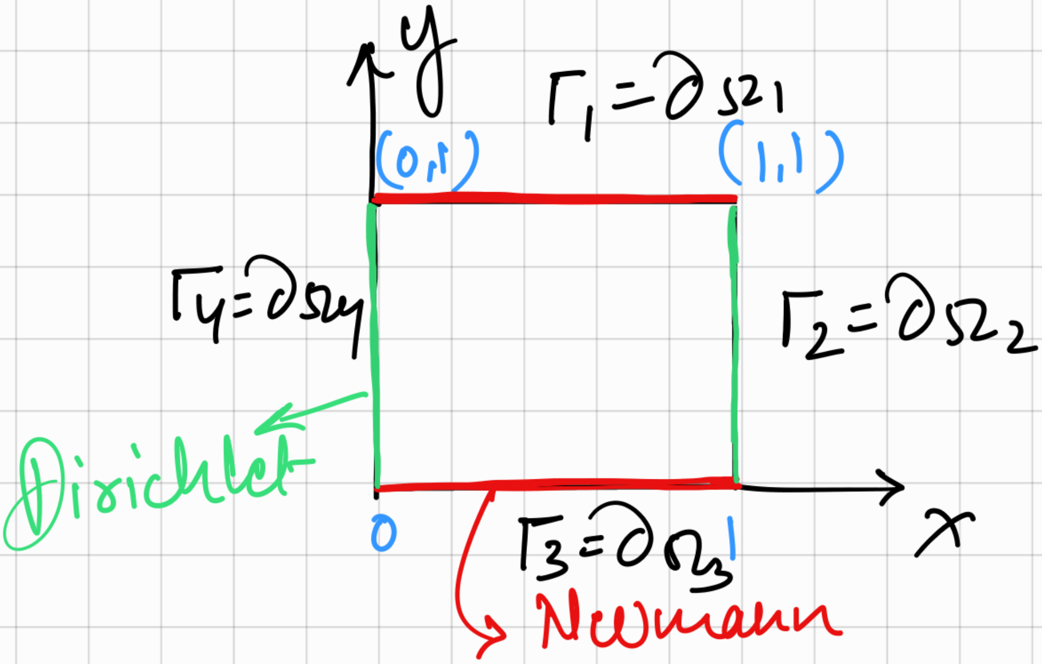
$$\frac{\partial^2 u_e}{\partial y^2} = -\pi^2 \sin(\pi x)\cos(\pi y)$$

$$\Rightarrow -\nabla^2 u_e = -\left(\frac{\partial^2 u_e}{\partial x^2} + \frac{\partial^2 u_e}{\partial y^2}\right)$$

$$= -\left(-2\pi^2 \sin\pi x \cos\pi y\right)$$

$$-\nabla u_e = f$$

1.



- introduce a test function

$$w(x,y) = \delta v(x,y)$$

is a variation and

Since the variable unknown is denoted by $v(x,y)$

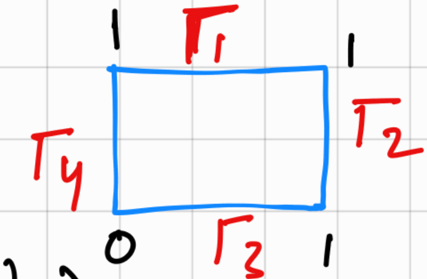
$$\Rightarrow -\int_0^1\int_0^1 w\nabla^2 v \, dx\,dy = \int_0^1\int_0^1 w f \, dx\,dy$$

$$\Rightarrow \int_0^1\int_0^1 \nabla(w)\cdot\nabla v \, d\Omega - \oint w\,\hat{n}\cdot\nabla v \, d\Gamma = \iint w f \, d\Omega$$

over $\Gamma_1, \Gamma_2, \Gamma_3$ & Γ_4

(for Both Ω_D & Ω_N)

Now:- $\oint w \, \hat{n} \cdot \nabla U \, d\Gamma$



① $\Gamma_1 = \hat{n} \cdot \nabla U = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} \partial_x U \\ \partial_y U \end{pmatrix}$

But $\partial_y U = 0$ (Neumann B.C)

$\Rightarrow \Gamma_1 = 0$

② $\Gamma_3 = \hat{n} \cdot \nabla U = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \begin{pmatrix} \partial_x U \\ \partial_y U \end{pmatrix}$

But $\partial_y U = 0$ (Neumann B.C)

$\Rightarrow \Gamma_3 = 0$

③ $\Gamma_2 = \hat{n} \cdot \nabla U = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} \partial_x U \\ \partial_y U \end{pmatrix}}_{\text{does not work}}$

and; $U(0,y) = a$, $w(0,y) = 0$

$U(1,y) = b$, $w(1,y) = 0$

Therefore $\Gamma_2, \Gamma_4 \to 0$

$\Rightarrow$ Hence,

$$\int_0^1\int_0^1 \nabla(w)\cdot\nabla u \, d\Omega - \oint_{\partial\Omega} w\,\hat{n}\cdot\nabla u \, d\Gamma = \int\int w f \,\partial\Omega$$

$\to 0$ (for Both Ω_D & Ω_N)

$$\int_0^1\int_0^1 \nabla(w)\cdot\nabla u \, dx\,dy = \int_0^1\int_0^1 w f \, dx\,dy$$

$$\int_0^1\int_0^1 \left(\nabla u\cdot\nabla(w) - w f\right) dx\,dy = 0$$

$\to$ weak form

Now :-

$\Rightarrow$ minimization of the the Ritz-Galerkin integeral:

$$F(u) = \int_0^1\int_0^1 \frac{1}{2}|\nabla u|^2 - uf\,dx\,dy = 0$$

functional

Variational ;

$$\delta F(u) = \lim_{\epsilon \to 0} \frac{F[u+\varepsilon\delta u] - F(u)}{\varepsilon}$$

But;

$$= \lim_{\epsilon \to} \frac{d}{d\varepsilon} F[u+\varepsilon\delta u] = 0$$

$$\Rightarrow F(u+\varepsilon\delta u) = \int_0^1\int_0^1 \frac{1}{2}\left[\nabla(u+\varepsilon\delta u)\right]^2 - (u+\varepsilon\delta u) f\,dx\,dy$$

$$\Rightarrow \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} F[u+\varepsilon\delta u] = \int_0^1\int_0^1 -(\delta u) f + \underbrace{\lim_{\varepsilon \to 0} \frac{d}{d\varepsilon}\left(\frac{1}{2}\left|\nabla u + \varepsilon \delta u\right|^2\right)}$$

$$\lim_{\varepsilon \to 0} \frac{d}{d\varepsilon}\left(\varepsilon \nabla u \cdot \nabla \delta u + \underbrace{O(\varepsilon^2)}\right)$$

Ignoring square terms

$$\left(\nabla u \cdot \nabla \delta u\right) + 0$$

$$\Downarrow$$

$$\Rightarrow \lim_{\varepsilon \to 0} F[u+\varepsilon\delta u] = \int_0^1\int_0^1 -(\delta u) f + \left(\nabla u \cdot \nabla \delta u\right) dx\,dy$$

$$0 = \int_0^1\int_0^1 -(\delta u)f + (\nabla u \cdot \nabla \delta u)\,dx\,dy$$

$$0 = \int_0^1\int_0^1 (-f + \nabla^2 u)\,\delta u\,dx\,dy + \underbrace{\oint \delta u\,\hat{n}\cdot\nabla u\,d\Gamma}_{0 \text{ (Because of B.C.s)}}$$

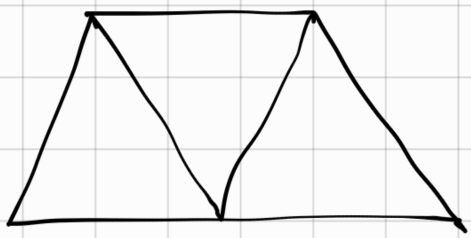
$$\Rightarrow \underbrace{\int_0^1\int_0^1 \left(\nabla u \cdot \nabla(\delta u) - f(\delta u)\right)dx\,dy = 0}_{\text{This is exactly the weak form}}$$

This shows the test function $w(x,y)$ is same as $\delta u(x,y)$

$$w(x,y) = \delta u(x,y)$$

2. Discrete Ritz-Galerkin

- $u(x,y)$ is continous but not differentiable
- $u(x,y)$ is the unknown & $u_h(x,y) \to$ FEM Approximation



$h \to$ minimum element size

$\Rightarrow u(x,y) \approx u_h(x,y)$ when $h \to 0$

$$u(x,y) = \lim_{h \to 0} u_h(x,y)$$

$$u(x,y) = \sum_{j=1}^{N_h} \hat{U}_j \phi_j(x,y)$$

$\hat{u}^n_j \rightarrow$ Coefficient

& $\phi_j(x,y) \rightarrow$ are the

basis function.

$\downarrow$

These basis function have

compact support for CG1.

$\downarrow$

Means, $\phi(x,y) = 1$ on the main node & 0 for neighbouring nodes & outside the neighbouring element

$\mapsto$ Thus compact support

$$\Rightarrow u_h(x,y) = \sum_{j=1}^{NDof} U_j \phi_j(x,y)$$

$$= \sum_{j=1}^{NDof\,(\text{Neumann})} U_j \phi_j(x,y)$$

$$\cancel{\sum_{j}^{ND\,\text{irichlet}} \hat{U}_j \phi_j(x,y)} \rightarrow 0 \;(B.C.)$$

$$u_n(x,y) = \sum_{j=1}^{N} \hat{u}_j \phi_j(x,y)$$

$\Rightarrow$ take $w(x,y) = \phi_i(x,y)$

for $i = 1, 2, \ldots N_{nodes}$

But ($N_{in} + N_{nodes}$)

from energy functional:

$$F(u) = \int_0^1\int_0^1 \frac{1}{2}|\nabla u|^2 - uf \, dx\,dy$$

Functional

Substitute $u_n(x,y)$ into $F(u)$

$$F(u_n) = \int_0^1\int_0^1 \frac{1}{2}|\nabla u_n|^2 - \int_0^1\int_0^1 f\, u_n \, dx\,dy$$

Now:

$$F(u_h) = \frac{1}{2}\int_0^1\int_0^1 \left|\nabla\left(\sum_{j=1}^{N_u} \hat{u}_j \phi_j(x,y)\right)\right|^2$$

$$- \int_0^1\int_0^1 f\left(\sum_{j=1}^{\infty} \hat{u}_j \phi_j(x,y)\right)$$

$\Rightarrow$ expanding $|\nabla u_h|^2$;

$\Rightarrow |\nabla u_h| \cdot |\nabla u_h|$,

$$\Rightarrow \sum_{j=1}^{N_u} u_j \nabla\phi_j(x,y) \cdot \sum_{i=1}^{N_u} u_i \nabla\phi_i(x,y),$$

$\Rightarrow$ Substitute it in functional $F(u_h)$

$$F(u_h) = \frac{1}{2}\int_0^1\int_0^1 \sum_{j=1}^{N_u}\sum_{i=1}^{N_u} u_j u_i \nabla\phi_j \cdot \nabla\phi_i \, dx\,dy$$

$$- \int_0^1\int_0^1 f \sum_{j=1}^{N_u} u_j \phi_j(x,y)\, dx\, dy ,$$

$\Rightarrow$ To minimize it, $\frac{\partial F(u_n)}{\partial u_j} = 0$

$$\Rightarrow \sum_{j=1}^{N_u} u_i \int_0^1\int_0^1 \nabla\phi_j \cdot \nabla\phi_i \, dx\, dy$$

$$- \sum_{j=1}^{N_u} \int_0^1\int_0^1 f \phi_j \, dx\, dy = 0 ,$$

In matrix form,

$$\boxed{A_{ij} u_i = b_i}$$

where $A_{ij} = \sum_{j=1}^{N_u} \int_0^1\int_0^1 \nabla u_j \nabla \phi_j$

$\Rightarrow$

$$b_i = \sum_{i=1}^{N_{nodes}} \int_0^1\int_0^1 f \phi_i \, dx\, dy$$

for $i = 1, 2, \ldots N_{nodes}$
& $j = 1, 2, \ldots N_n$

Now for discrete weak formulation,

$$\int_0^1\int_0^1 \left(\nabla U \cdot \nabla(w) - wf\right)dx\,dy = 0$$

$\hookrightarrow$ weak form

$$\Rightarrow \int_0^1\int_0^1 \nabla\left(\sum_{j=1}^{N_{nn}} u_j \phi_j\right) \cdot \nabla\phi_i \, dx\,dy$$

$$= \iint f \phi_i \, dx\,dy$$

$$\Rightarrow \sum_{j=1}^{N_n} \int_0\int_0 \nabla u_j \nabla \phi_j \cdot \nabla \phi_i \, dx\,dy$$

$$= \sum_{i=1}^{N_{node}} \iint f \phi_i \, dx\,dy$$

In matrix form,

$$A_{ij} u_j = b_i$$

where $A_{ij} = \sum_{j=1}^{N_u} \int_0^1\int_0^1 u_j \nabla \phi_j$

$\Rightarrow$

$$b_i = \sum_{i=1}^{N_{nodes}} \int_0^1\int_0^1 f \phi_i \, dx\, dy$$

for $i = 1, 2, \ldots N_{nodes}$

& $j = 1, 2, \ldots N_u$

$\hookrightarrow$ for both processes

the discrete forms are same

$\Rightarrow$ Therefore final discrete form

$$\sum_{j=1}^{N_{nodes}} A_{ij} u_j = b_i \quad \text{for } i = 1, 2, \dots N_{nodes}$$

$\Rightarrow$ unknown coefficient vector

u_i,

$$u = [u_1, u_2 \cdots u_N]^T$$

$\Rightarrow$ Now variation in discrete energy functional.

$$F(u) = \frac{1}{2} \int_0^1\!\!\int_0^1 |\nabla u_h|^2 \, dx\,dy - \int_0^1\!\!\int_0^1 f \, u_h \, dx\,dy$$

Substituting

$$\nabla u_h = \nabla \left(\sum_{j=1}^{N_n} u_j \phi_j(x, y) \right)$$

taking variation of δu_n

$$\therefore \nabla u_n = \nabla u_n + \varepsilon \nabla \delta u_n$$

$$= \sum_{j=1}^{N_u} u_j \nabla \phi_j + \varepsilon \sum_{j=1}^{N_u} \delta u_j \nabla \phi_j$$

Now

$$|\nabla u_n|^2 = |\nabla u_n|^2 + 2\varepsilon \nabla u_n \cdot \nabla \delta u_n + \varepsilon^2 |\nabla \delta u_n|^2 .$$

Substitute in $F(u_n)$

$$F(u_n + \varepsilon \delta u_n) = \int_0^1 \int_0^1 \frac{1}{2} |\nabla (u_n + \varepsilon \delta u_n)|^2 - (u_n + \varepsilon \delta u_n) f \, dx dy$$

$$\lim_{\varepsilon \to 0} \frac{dF}{d\varepsilon}(u_n + \varepsilon \delta u_n) = \int_0^1 \int_0^1 -(\delta u_n) f$$

$$+ \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon}\left(\frac{1}{2}\left|\nabla u_n + \varepsilon \delta u_n\right|^2\right)$$

$\downarrow$

making linear scale

$$(\nabla u \cdot \nabla \delta v) + 0$$

(Ignoring square term)

$$\Rightarrow \lim_{\varepsilon \to 0} F(u_n + \varepsilon \delta u_n)$$

$$= \int_0^1 \int_0^1 \left(-(\delta v_n) f + (\nabla v_n \cdot \nabla \delta u_n)\right) dx\,dy$$

$$\Rightarrow \int_0^1 \int_0^1 (\nabla u_n \cdot \nabla \delta v_n)\, dx\,dy$$

$$- \int_0^1 \int_0^1 (\delta v_n) f \, dx\,dy$$

$$\Rightarrow \int_0^1\int_0^1 \sum_{j=1}^{Nu} u_j \nabla\phi_j \cdot \sum_{i=1}^{Nu} \delta u_i \nabla\phi_i \, dx\,dy$$

$$- \int_0^1\int_0^1 \sum_{i=1}^{Nu} \delta u_i \nabla\phi_i f \, dx\,dy$$

$$\Rightarrow \sum_{j=1}^{Nu}\sum_{i=1}^{Nu} u_j \delta u_i \int_0^1\int_0^1 \nabla\phi_j \bar{\phi}_i \, dx\,dy$$

$$= \sum_{i=1}^{Nu} \delta u_i \int_0^1\int_0^1 f \nabla\phi_i \, dx\,dy$$

Since δu_n is arbitrary

$$\Rightarrow \sum_{j=1}^{Nu}\sum_{i=1}^{Nu} u_j u_i \int_0^1\int_0^1 \nabla\phi_j \cdot \nabla\phi_i \, dx\,dy$$

$$= \sum_{i=1}^{Nu} u_i \int_0^1\int_0^1 f \nabla\phi_i \, dx\,dy$$

$$\Rightarrow \sum_{j=1}^{N_n} \sum_{i=1}^{N_n} u_j \int_0^1 \int_0^1 \nabla \phi_j \cdot \nabla \phi_i \, dx\, dy - \sum_{i=1}^{N_n} \int_0^1 \int_0^1 f \nabla \phi_i \, dx\, dy$$

$\Rightarrow$ In matrix form.

$$\boxed{A_{ij} u_j = b_i}$$

Hence, variation in first yields the later,

Question 4: Contour plots

1. Figures 1,2,3,4 and 5 show the solution contour plots of $u_h(x, y)$ under different space resolution at $p = 1$.

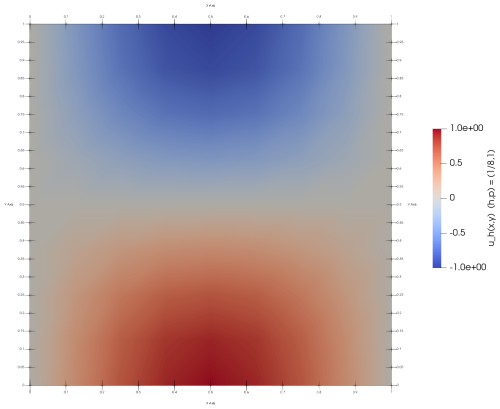


Figure 1: $u_h(x, y)$ $(h, p) = (1/8, 1)$

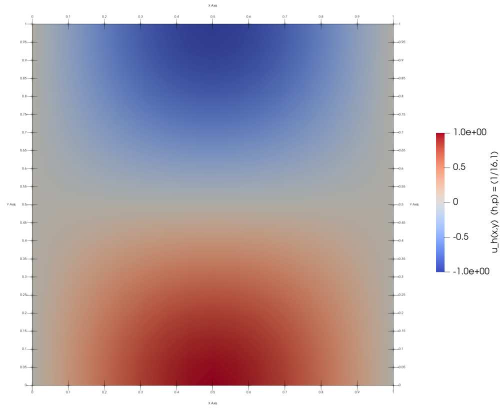


Figure 2: $u_h(x, y)$ $(h, p) = (1/16, 1)$

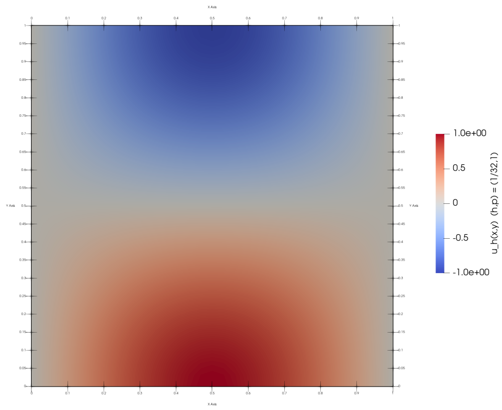


Figure 3: $u_h(x, y)$ $(h, p) = (1/32, 1)$

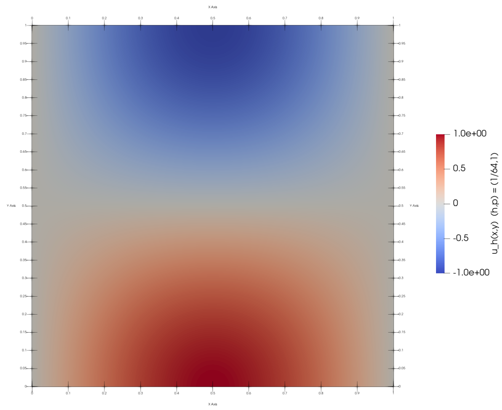


Figure 4: $u_h(x, y)$ $(h, p) = (1/64, 1)$

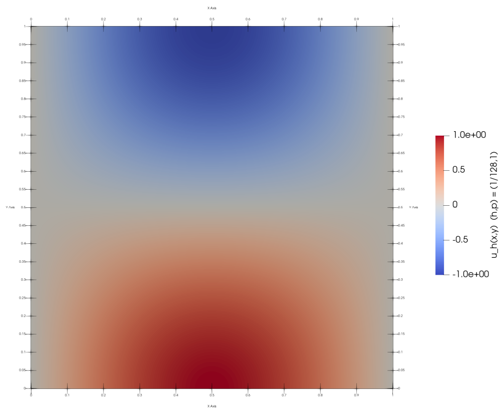


Figure 5: $u_h(x, y)$ $(h, p) = (1/128, 1)$

2. Figure 6 to 20 shows the Error contour of error $|u_h - u_e|$ for spatial resolution of 8×8, 16×16, 32×32, 64×64, 128×128 at $p = 1, 2, 3$ in each case.

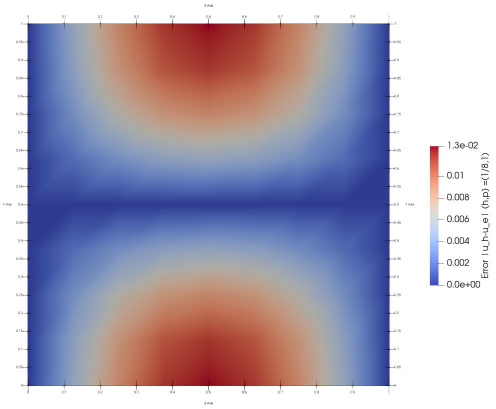


Figure 6: $Error|u_h - u_e|$ $(h,p) = (1/8, 1)$

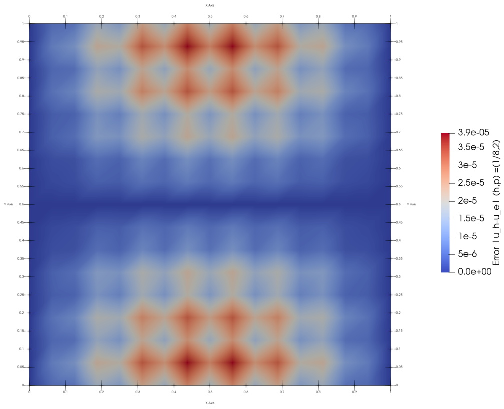


Figure 7: $Error|u_h - u_e|$ $(h,p) = (1/8, 2)$

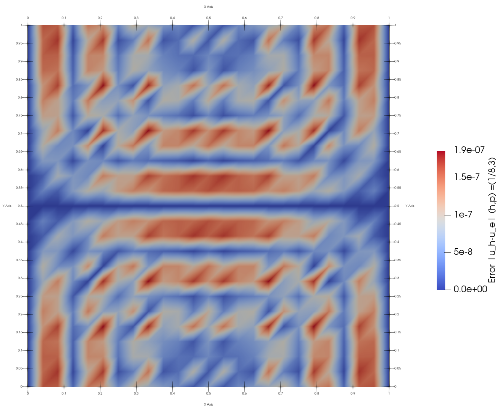


Figure 8: $Error|u_h - u_e|$ $(h,p) = (1/8, 3)$

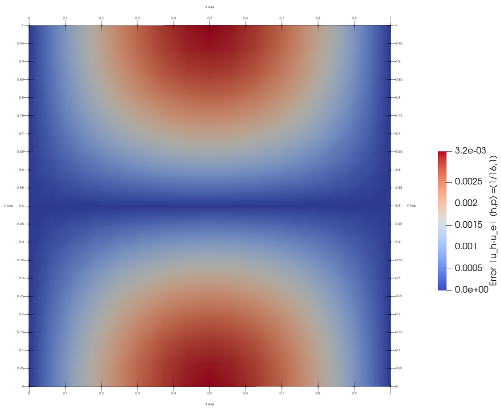


Figure 9: $Error|u_h - u_e|$ $(h,p) = (1/16, 1)$

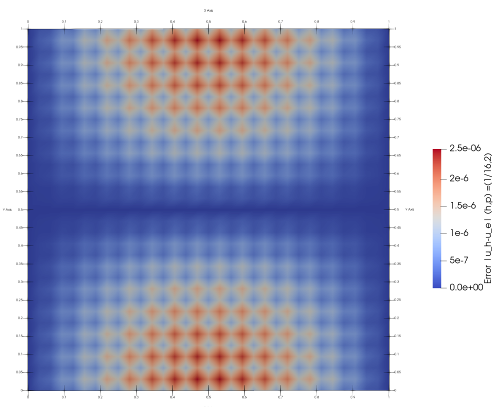


Figure 10: $Error|u_h - u_e|$ $(h,p) = (1/16, 2)$

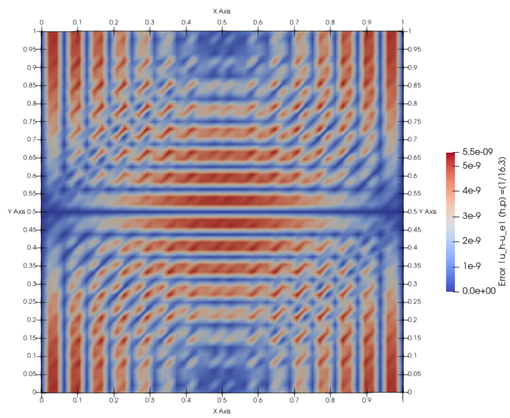


Figure 11: $Error|u_h - u_e|$ $(h,p) = (1/16, 3)$

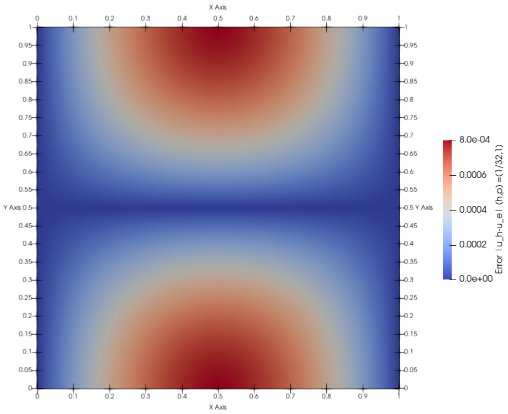


Figure 12: $Error|u_h - u_e|$ $(h,p) = (1/32, 1)$

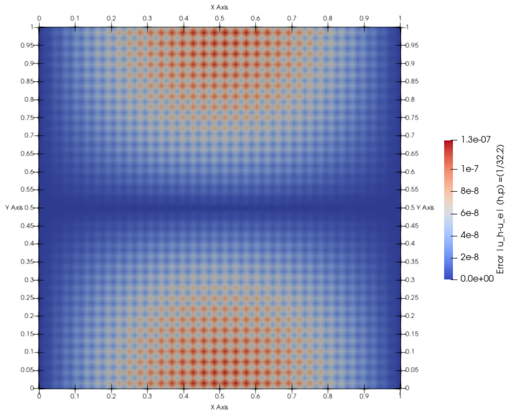


Figure 13: $Error|u_h - u_e|$ $(h,p) = (1/32, 2)$

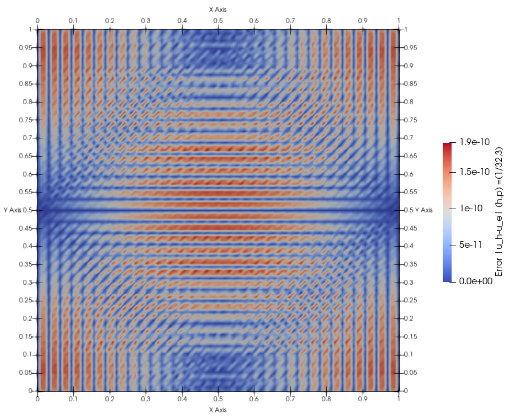


Figure 14: $Error|u_h - u_e|$ $(h,p) = (1/32, 3)$

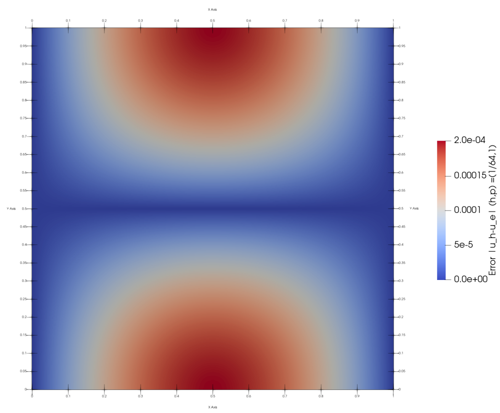


Figure 15: $Error|u_h - u_e|$ $(h,p) = (1/64, 1)$

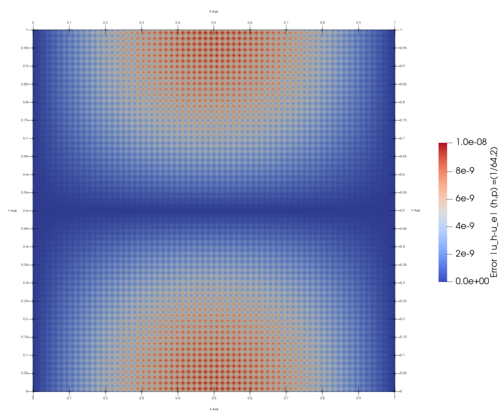


Figure 16: $Error|u_h - u_e|$ $(h,p) = (1/64, 2)$

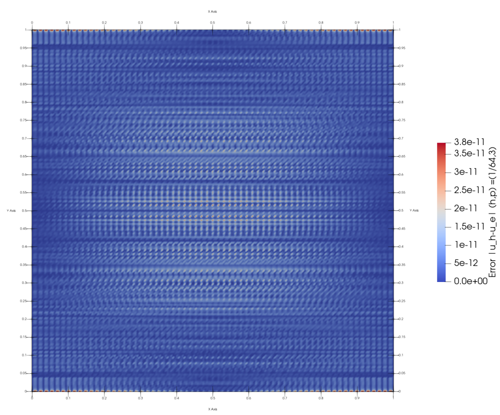


Figure 17: $Error|u_h - u_e|$ $(h,p) = (1/64, 3)$

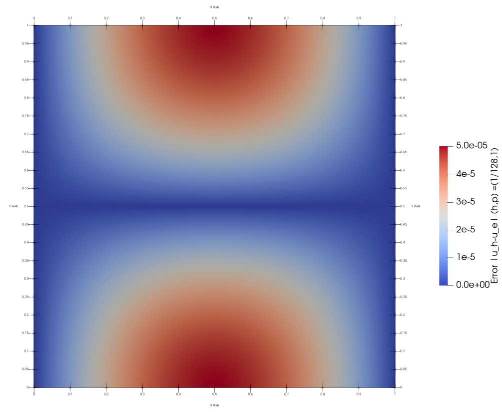


Figure 18: $Error|u_h - u_e|$ $(h,p) = (1/128,1)$

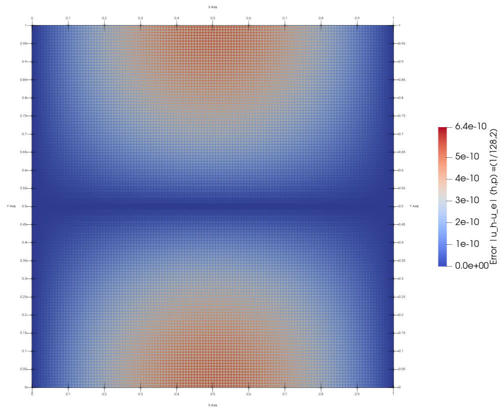


Figure 19: $Error|u_h - u_e|$ $(h,p) = (1/128,2)$

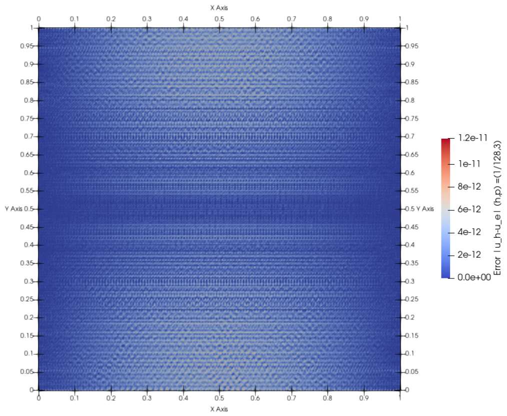


Figure 20: $Error|u_h - u_e|$ $(h,p) = (1/128,3)$

- From Figure 6 to 20, as the n or $1/h$ (Spatial resolution) increases, The error $|u_h - u_e|$ increase too in order of $O(h^2)$ when $p = 1$.
- When refining p, it reduces the error more effectively than refining n. Therefore, it is advisable to prioritize the refinement of p to minimize error. However, it is essential to find the optimal combination of both (h,p) to achieve the lowest error. Increasing p while keeping h at a smaller value can result in missing important features due to insufficient spatial resolution.
- From Figures 17 and 20, both exhibiting an error on the order of $O(10^{-11})$, we observe that for the combinations of $(h,p) = (1/128,3)$ and $(h,p) = (1/64,3)$, higher values of p result in minimized error. Additionally, the effect of increasing the value of n becomes negligible.

Question 5

I have commented in the $Poissons_eq_v2(Commented).py$ file, which is uploaded to Github.

Question 6

As per your feedback on numerics exercises 1 and 2 I need to redo a few FVM problems again, I will redo it by next week and submit it all together.