Foundations of Fluids Numerical Exercises 3

Billy Hollis 201421513

2024/2025

We begin with the Poisson system

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

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

$$u(0,y) = u(1,y) = 0 \tag{3}$$

$$\partial_y(x,y)|_{y=0} = \partial_y(x,y)|_{y=1} = 0 \tag{4}$$

We first claim that the exact solution to our problem is $u_e = \sin(\pi x)\cos(\pi y)$. We observe that

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

in which case equations (1) and (2) are satisfied. Note that

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

Since $\sin(0) = \sin(\pi) = 0$, the boundary conditions (3) and (4) are satisfied and we have shown that the exact solution is indeed u_e.

We will now derive the equations (1) - (4) by considering the variation of a minimisation problem. Before we do this, we introduce some notation for the boundary of our domain Γ. Let Γ_1 denote the union of the vertical boundaries and let Γ_2 denote the union of the horizontal boundaries so that $\Gamma = \Gamma_1 \cup \Gamma_2$. Define the integral

$$I[u] = \int_\Omega \left(\frac{1}{2} |\nabla u|^2 - uf \right) d\Omega \tag{5}$$

where f is given by (2). Consider the problem in which we minimise (5) over the target space

$$\Sigma = \{u : u \ \text{ sufficiently smooth and } \ u|_{\Gamma_1} = 0\} \tag{6}$$

Let $\bar{u}$ denote the solution to equations (5) and (6) and consider the class of functions

$$u(\mathbf{x}) = \bar{u}(\mathbf{x}) + \epsilon\eta(\mathbf{x}) \tag{7}$$

where ϵ is a variable parameter and η is an arbitrary but fixed function. We must have $u|_{\Gamma_1} = 0$ and $\bar{u}|_{\Gamma_1} = 0$, which means that we must also have $\eta|_{\Gamma_1} = 0$. Of course, η must also be sufficiently smooth. We substitute (7) into (5) so

$$I[\bar{u} + \epsilon\eta] = \int_\Omega \left(\frac{1}{2} |\nabla(\bar{u} + \epsilon\eta)|^2 - \eta f \right) d\Omega \tag{8}$$

Since $\bar{u}$ is a known function and η is fixed, we conclude that I only depends on ϵ. To ensure the existence of an extreme, we must have

$$\frac{dI}{d\epsilon} = 0$$

This implies that

$$\int_\Omega (\nabla(\bar{u} + \epsilon\eta) \cdot \nabla\eta - \eta f)\, d\Omega = 0 \tag{9}$$

By considering equation (7), we conclude that I is minimised when $\epsilon = 0$ so equation (9) reduces to

$$\int_\Omega (\nabla\bar{u} \cdot \nabla\eta - \eta f)\, d\Omega = 0 \tag{10}$$

By the divergence theorem, we have

$$\int_\Omega \nabla\bar{u} \cdot \nabla\eta d\Omega = -\int_\Omega \eta\nabla \cdot (\nabla\bar{u}) d\Omega + \int_\Gamma \eta(\nabla\bar{u}) \cdot \mathbf{n} d\Gamma \tag{11}$$

where $\mathbf{n}$ is the outward unit normal vector to the surface Γ. Recall that $\eta|_{\Gamma_1} = 0$, so we have

$$\int_\Gamma \eta(\nabla\bar{u}) \cdot \mathbf{n} d\Gamma = \int_{\Gamma_2} \eta(\nabla\bar{u}) \cdot \mathbf{n} d\Gamma \tag{12}$$

Using equations (11) and (12), equation (10) becomes

$$\int_\Omega (-\nabla \cdot \nabla\bar{u} - f)\eta d\Omega + \int_{\Gamma_2} \eta(\nabla\bar{u}) \cdot \mathbf{n} d\Gamma = 0 \tag{13}$$

Let us first place an additional restriction on η. Suppose that $\eta|_{\Gamma_2} = 0$ so that $\eta = 0$ everywhere on the unit square. In this case, the second integral in equation (13) will vanish. Since $\eta = 0$ everywhere on the boundary of Ω, we can apply the Dubois-Reymond Lemma to the first integral which will give

$$-\nabla \cdot \nabla\bar{u} - f = 0 \tag{14}$$

Evidently this is identical to equation (1) and we conclude that any solution to the minimisation problem (5) subject to (6) must also solve the Poisson equation (1). Throughout the derivation we have been using the condition $\bar{u}|_{\Gamma_1} = 0$, which is exactly the boundary condition (3). This is called the essential boundary condition. Since our Poisson equation is second order in two space variables, we require two more boundary conditions. To find them, let us relax the condition that $\eta|_{\Gamma_2} = 0$. Substituting (14) into (13), we obtain

$$\nabla\bar{u} \cdot \mathbf{n} = 0 \tag{15}$$

Since this equation is homogeneous, we can take $\mathbf{n} = (0, 1)^T$ for both the upper and lower part of Γ_2. This simply returns the boundary condition (4). These boundary conditions were somewhat hidden in the variational formulation and are called the natural boundary conditions. Overall, we have shown that the system (1)-(4) can be obtained by considering the variation of the minimisation problem (5)-(6). We will now proceed by working in the opposite direction and computing the weak form of the system (1)-(4). We multiply equation (1) by some test function w and then integrating over the domain Ω to obtain the expression

$$\int_\Omega (-\nabla \cdot \nabla\bar{u} - f)w d\Omega = 0 \tag{16}$$

Note that our test function w must satisfy our boundary condition on Γ_1 and must be sufficiently smooth. Taking into account our natural boundary condition, we obtain

$$\int_\Omega (-\nabla \cdot \nabla\bar{u} - f)w d\Omega + \int_{\Gamma_2} w(\nabla\bar{u}) \cdot \mathbf{n} d\Gamma = 0 \tag{17}$$

Again using the divergence theorem, we obtain the expression

$$\int_\Omega (\nabla\bar{u} \cdot \nabla w - wf)\, d\Omega = 0 \tag{18}$$

which we rearrange to give the weak form of our system

$$\int_\Omega \nabla u \cdot \nabla w d\Omega = \int_\Omega wf d\Omega \tag{19}$$

It should be fairly obvious that η and w are playing the same role although we introduced them when working in different directions. Hence we have $\eta = w$.

We will now approximate the solution in both the variational and weak formulation set-ups. We recall that

$$I[u] = \int_\Omega \left(\frac{1}{2} |\nabla u|^2 - uf \right) d\Omega \tag{20}$$

We seek a solution $u(x, y) \approx u_h(x, y)$ where

$$u_h(x, y) = \sum_{j=1}^{n} u_j \varphi_j(\mathbf{x}) \tag{21}$$

Here, each φ_j is in the same target space of the minimisation problem, namely we have that each φ_j must be sufficiently smooth and satisfy $\varphi_j(\mathbf{x})|_{\Gamma_1} = 0$ for all j. We note that n is the number of nodes minus the number of nodes lying on Γ_1. If these boundary conditions were not homogenous, we would require extra known terms to appear in the following analyses. By seeking a solution of this form and substituting into (20), our problem becomes a minimisation problem which is solved over the subset of Σ spanned by the unknowns u_j. We then require that

$$\frac{\partial I[u_h]}{\partial u_i} = 0 \text{ for } i = 1, 2, \ldots, n \tag{22}$$

This will provide a system of n equations in n unknowns for the values u_j. To this end we have

$$I[u_h] = \int_\Omega \left(\frac{1}{2} \left| \nabla \left(\sum_{j=1}^{n} u_j \varphi_j \right) \right|^2 - uf \right) d\Omega \tag{23}$$

Then by using equation (22), we obtain the system of equations

$$\sum_{j=1}^{n} u_j \int_\Omega \nabla \varphi_i \cdot \nabla \varphi_j d\Omega = \int_\Omega f \varphi_i d\Omega \tag{24}$$

for $i = 1, 2, \ldots, h$. This may be written in the more convenient matrix form $\mathbf{Su} = \mathbf{f}$ where $S_{ij} = \int_\Omega \nabla \varphi_i \cdot \nabla \varphi_j d\Omega$, $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ and $\mathbf{f} = (f\varphi_1, f\varphi_2, \ldots, f\varphi_n)^T$. Let us now perform a similar calculation using the weak form of our problem (19). Similarly to before we seek a solution of the form

$$u_h(x, y) = \sum_{j=1}^{n} a_j \varphi_j(\mathbf{x}) \tag{25}$$

where we are using the same compactly supported basis functions φ_j. Upon substitution into (20) we obtain

$$\int_\Omega \nabla \left(\sum_{j=1}^{n} a_j \varphi_j(\mathbf{x}) \right) \cdot \nabla w d\Omega = \int_\Omega w f d\Omega \tag{26}$$

In this formulation we argue slightly differently. The weak formulation of our problem must hold for all test functions w in the defined space, i.e. in the space of functions which vanish on Γ_1. In particular, we can set

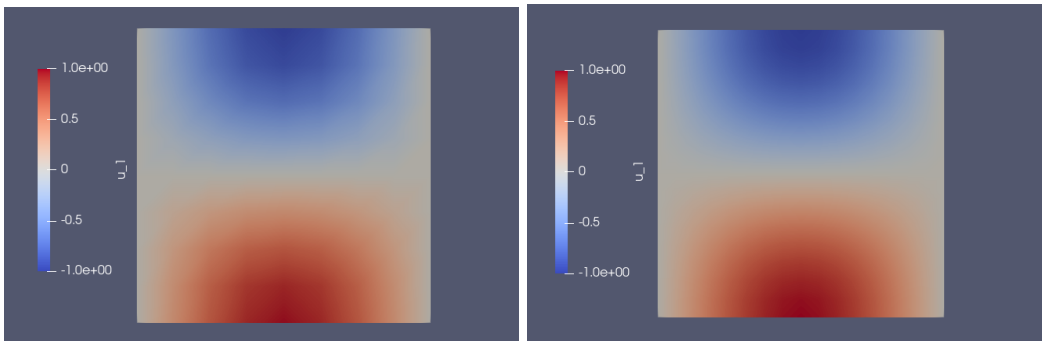
$$w = \sum_{j=1}^{n} b_j \varphi_j(\mathbf{x}) \tag{27}$$

and then choose the coefficients such that $b_j = 1$ for $j = i$ and $b_j = 0$ otherwise. Using this, we obtain

$$\sum_{j=1}^{n} a_j \int_\Omega \nabla \varphi_j \cdot \nabla \varphi_i d\Omega = \int_\Omega f \varphi_i d\Omega \tag{28}$$

for $i = 1, 2, \ldots, n$. Evidently, this is equivalent to (24), we have just used a different name for our unknowns. So far, we have shown equivalent formulations of our problem and then shown that the systems of equations to solve are indeed equivalent. We will verify that numerical implementation using each formulation yields appropriately small errors.

Let us consider the plots of the solution for various spatial resolutions.

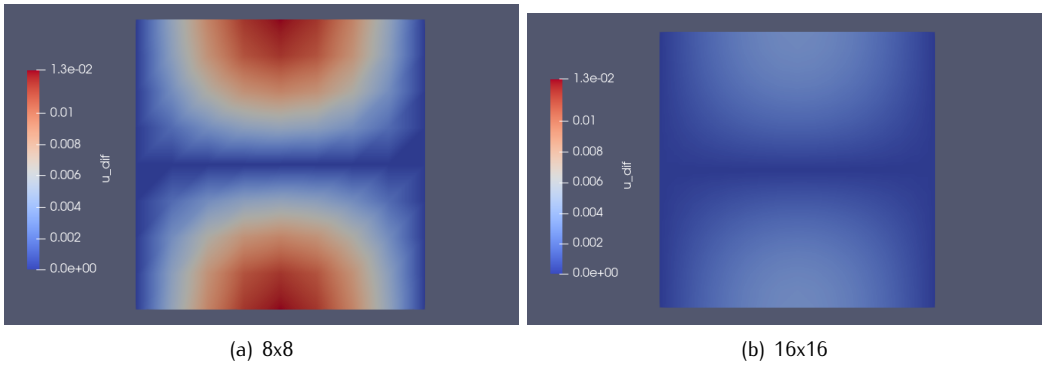


(a) 8x8 (b) 16x16

(c) 64x64 (d) 128x128

Figure 1: Approximate solutions using the weak formulation setup with varying mesh sizes.

Whilst these plots become smoother as we make the mesh more fine, this does not confirm convergence. Instead, we will consider how the value $|u_h(\mathbf{x}) - u_e(\mathbf{x})|$ changes with mesh coarseness.



(a) 8x8 (b) 16x16

(c) 64x64 (d) 128x128

Figure 2: Contour plots of $|u_h(\mathbf{x}) - u_e(\mathbf{x})|$ varying by mesh size using the weak formulation setup.

We can obtain similar plots from the implementation of the Ritz-Galerkin method but they are almost visually identical so we instead look at the errors in more detail.

From the above plots, convergence becomes a little more obvious but we will further investigating the L_2 norm error between our approximate solution plotted above and the exact solution.

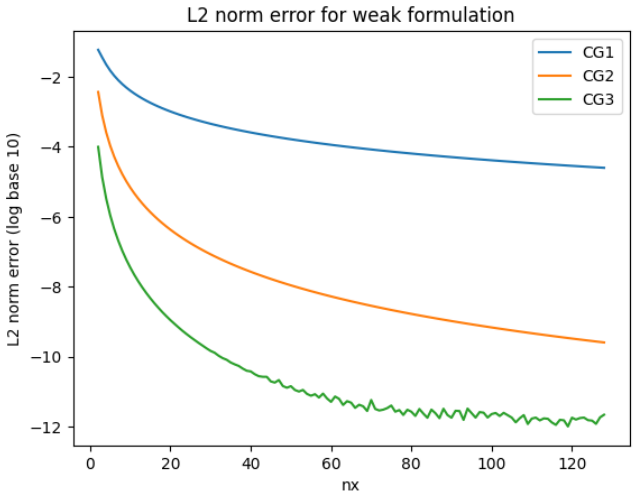


Figure 3: L_2 norm difference between approximate solution using weak formulation implementation and exact solution for various orders and mesh resolutions.

As we can see in Figure 3, increasing the mesh resolution decreases the difference between the approximate solution and the exact solution for first and second order continuous Galerkin implementations of the weak formulation. The same is true for third order although the behaviour becomes a little more unpredictable for finer meshes. Suppose we wish to obtain a solution which is accurate to order 10^{-8}. Then we could use CG2 with $nx \approx 15$ or CG3 with $nx \approx 43$. Combined with Table 1, we see that a significantly greater value of nx is needed to obtain comparatively accurate results when using CG1. Hence, increasing the order is more useful than in increasing resolution in many cases. Whilst this gives an indication of the convergence, let us investigate in more detail.

Resolution	L_2 norm error
8x8	0.006213900940246132
16x16	0.0015930107731038869
32x32	0.00040075733647192444
64x64	0.0001003464040232397
128x128	2.5096426249750656e-05
256x256	6.274721113367353e-06
512x512	1.5687200485843606e-06
1024x1024	3.8914130879272917e-07

Table 1: Weak formulation L_2 error for various spatial resolutions using CG1.

With these more specific values, it is more clear that the convergence is approximately quadratic for CG1, where doubling the resolution decreases the error by a factor of ≈ 4. We can perform a similar investigation into CG2.

Resolution	L_2 norm error
8x8	1.6573873958586682e-05
16x16	1.0427252041804732e-06
32x32	6.527731850134832e-08
64x64	4.081585069002153e-09
128x128	2.5548043762165057e-10
256x256	1.7317616892713087e-11
512x512	6.4779756469743625e-12
1024x1024	2.2004481876126444e-11

Table 2: Weak formulation L_2 error for various spatial resolutions using CG2.

Performing similar calculations, one finds that the convergence is approximately quartic. In a similar fashion to what we observe in Figure 3, the error is not strictly decreasing when we increase the resolution for CG2. One may plot figures which are almost identical to those shown in Figure 2 but for the results we obtain from using the Ritz-Galerkin implementation. Instead, we consider the following plot which is analogous to Figure 3.

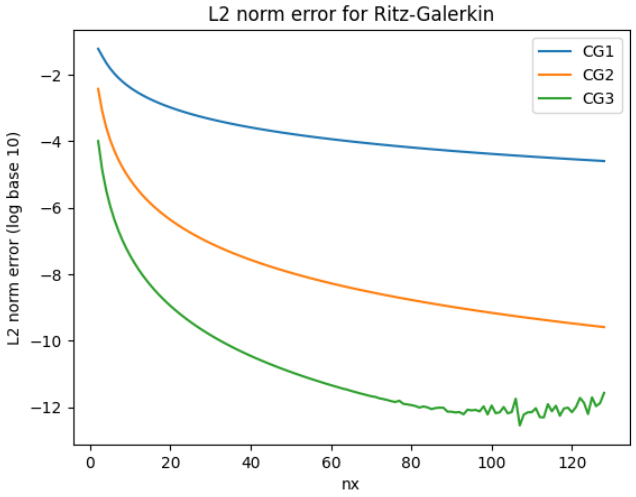


Figure 4: L_2 norm difference between approximate solution using Ritz-Galerkin implementation and exact solution for various orders and mesh resolutions.

We see a similar trend, whereby increasing the order improves the accuracy of the solution but once again seeing some instability for third order. As we explained when developing the theory of both methods, we speculated that both methods should yield appropriately similar results.

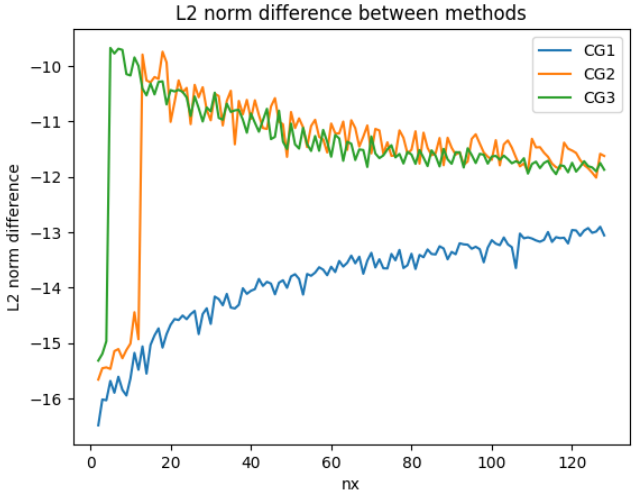


Figure 5: L_2 norm difference between approximate solution using weak formulation implementation and exact solution for various orders and mesh resolutions.

Here we observe that when using CG1, the two methods yield closer results than when using higher orders. Of course, this figure is purely to show the claimed equivalence of the formulations. In practice, when we solve our system, we would choose one formulation or the other and then choose the order of the CG and spatial resolution based on our tolerance for error as described previously.