TMA4212 - Exploring the powers of FEMs

Halvard Emil Sand-Larsen, Henrik Grenersen and Eirik Fagerbakke

March 2023

Contents

0 Introduction

In this problem, we will consider the following PDE, which models the concentration $u(x)$ in a 1D-tube

$$-\partial_x(\alpha(x)\partial_x u) + \partial_x(b(x)u) + c(x)u = f$$

Here, the terms α, b and c are first allowed to vary as functions of x, but will later be considered constant in order to simplify some calculations. Respectively, these terms represent the diffusion coefficient (which is positive), the fluid velocity and the decay rate, whilst f is a source term.

Throughout this project we will first focus on deriving analytical results concerning the existence and uniqueness of a solution to the above presented problem. Later, we will use a numerical approach to determine solutions of some test problems using FEMs.

1 Existence of a unique solution

1.1 Definitions and preliminary results

We start by providing some preliminary results, which will be used in the proofs. By Hölder's inequality, we have

$$\|uv\|_{L^1} \leq \|u\|_{L^2}\|v\|_{L_2} \tag{1}$$

The norm of $u(x)$ in H^1 is defined as

$$\|u\|^2_{H^1} = \|u\|^2_{L^2} + \|u_x\|^2_{L^2}$$

This implies that

$$\|u\|_{L^2} \leq \|u\|_{H^1} \qquad \|u_x\|_{L^2} \leq \|u\|_{H^1} \tag{2}$$

Furthermore, we will also make use of the Poincaré inequality

$$\|u_x\|_{L^2} \geq C\|u\|_{H^1} \tag{3}$$

We also assume that

$$\|\alpha\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty} + \|f\|_{L^2} \leq K \tag{4}$$

and

$$\alpha(x) \geq \alpha_0 > 0 \qquad c(x) > 0 \tag{5}$$

1.2 Weak formulations

We now want to determine a bilinear form $a(u, v)$ and a function $F(v)$ such that any classical solution u of our PDE, satisfies the equation $a(u, v) = F(v)$ for all $v \in H^1_0(0, 1)$. We start with our problem from before and multiply by v and integrate both sides over our domain.

$$\int_0^1 (-\partial_x(\alpha(x)\partial_x u)v + \partial_x(b(x)u)v + c(x)uv)\mathrm{d}x = \int_0^1 f(x)v\mathrm{d}x$$

We start by looking at our left hand side and split up the integral in three different parts. Then, we can use integration by parts on he first two integrals and get the following result

$$\Big[-\alpha(x)\partial_x uv\Big]_0^1 + \int_0^1 \alpha(x)\partial_x u\partial_x v\mathrm{d}x + \Big[b(x)uv\Big]_0^1 - \int_0^1 b(x)u\partial_x v\mathrm{d}x + \int_0^1 c(x)uv\mathrm{d}x$$

Since $v \in H_0^1(0,1)$, we see that the non-integral terms are equal to zero and we are left with

$$a(u,v) := \int_0^1 (\alpha(x)\partial_x u \partial_x v - b(x)u\partial_x v + c(x)uv)\,\mathrm{d}x = \int_0^1 f(x)v\mathrm{d}x =: F(v)$$

The weak formulation is then

$$\text{Find } u \text{ in } H_0^1 \text{ such that } a(u,v) = F(v) \text{ for all } v \text{ in } H_0^1$$

1.3 Continuity of $F(v)$ and $a(u,v)$

We start by showing that $F(v)$ is continuous, by showing that the operator norm is bounded. By definition, we have

$$\|F\| = \max_{v \in H^1, v \neq 0} \frac{|F(v)|}{\|v\|_{H^1}}$$

by Hölder's inequality 1 we get

$$\|F\| \leq \max_v \frac{\|f\|_{L^2}\|v\|_{L^2}}{\|v\|_{H^1}}$$

we now use inequality 2, which yields

$$\|F\| \leq \max_v \frac{\|f\|_{L^2}\|v\|_{L^2}}{\|v\|_{L^2}} = \|f\|_{L^2} < \infty$$

By assumption 4, we know that $\|f\|_{L^2}$ is finite. $F(v)$ is then a bounded, and therefore also continuous, functional.

We now move on to $a(u,v)$, and aim to show that this is a continuous bilinear form. That it is bilinear holds since integration, differentiation and multiplication is linear, and for continuity we reason as follows

$$\begin{aligned} a(u,v) \leq |a(u,v)| &\leq \|\alpha u_x v_x\|_{L^1} + \|b u v_x\|_{L^1} + \|c u v\|_{L^1} \\ &\leq \|\alpha\|_{L^\infty}\|u_x v_x\|_{L^1} + \|b\|_{L^\infty}\|u v_x\|_{L^1} + \|c\|_{L^\infty}\|uv\|_{L^1} \end{aligned}$$

Again, we can use Hölder's inequality 1, to obtain

$$a(u,v) \leq \|\alpha\|_{L^\infty}\|u_x\|_{L^2}\|v_x\|_{L^2} + \|b\|_{L^\infty}\|u\|_{L^2}\|v_x\|_{L^2} + \|c\|_{L^\infty}\|u\|_{L^2}\|v\|_{L^2}$$

We can now use inequality 2 to obtain

$$a(u,v) \leq (\|\alpha\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty})\|u\|_{H^1}\|v\|_{H^1} < \infty$$

Which we know is finite from assumption 4 and $u, v \in H^1$. $a(u,v)$ is then continuous, as we have $|a(u,v)| \leq M\|u\|_{H^1}\|v\|_{H^1}$ with $M = \|\alpha\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty}$.

1.4 Coercivity

Now we want to check if the final criteria in the Lax-Milgram theorem is satisfied, i.e. if our bilinear form is coercive. In order to do this, we fix our coefficient terms as constant, and evaluate as follows

$$a(u,u) \geq \alpha \int_0^1 u_x^2 \mathrm{d}x - |b| \left| \int_0^1 u u_x \mathrm{d}x \right| + c \int_0^1 u^2 \mathrm{d}x$$

Now we investigate the middle term. Young's inequality states that $ab \leq \frac{1}{2\epsilon}a^2 + \frac{\epsilon}{2}b^2$ for all $\epsilon > 0$, meaning that we have the following inequality

$$\left| \int_0^1 u u_x \mathrm{d}x \right| \leq \int_0^1 \left(\frac{1}{2\epsilon} u^2 + \frac{\epsilon}{2} u_x^2 \right) \mathrm{d}x$$

If we insert this into the inequality presented for $a(u,u)$, we then get that

$$a(u,u) \geq (\alpha - \frac{\epsilon}{2}|b|) \int_0^1 u_x^2 \mathrm{d}x + (c - \frac{1}{2\epsilon}|b|) \int_0^1 u^2 \mathrm{d}x$$

Now we aim to utilize this inequality in order to show that $a(u,u)$ is coercive, as the Lax-Milgram theorem then guarantees a unique solution to our problem, since the bilinear form also is continuous. We start by replacing our integrals by appropriate norms, which yields that

$$a(u,u) \geq \left(\alpha - \frac{\epsilon}{2}|b|\right) ||u_x||_{L^2}^2 + \left(c - \frac{1}{2\epsilon}|b|\right) ||u||_{L^2}^2 \quad \geq \min\left\{\alpha - \frac{\epsilon}{2}|b|, c - \frac{1}{2\epsilon}|b|\right\} \|u\|_{H^1}^2$$

Where the last inequality follows from 2. We need $C := \min\left\{\alpha - \frac{\epsilon}{2}|b|, c - \frac{1}{2\epsilon}|b|\right\}$ to be positive for coercivity, so we have the following inequalities to consider:

$$\alpha - \frac{\epsilon}{2}|b| > 0 \qquad c - \frac{1}{2\epsilon}|b| > 0$$

The first inequality can be rewritten as $\epsilon < \frac{2\alpha}{|b|}$, while the second one can be written as $c > \frac{b}{2\epsilon}$. Combining these inequalities, we get

$$c > \frac{|b|^2}{4\alpha} \tag{6}$$

Thus, it clearly holds for $c > \frac{|b|^2}{2\alpha}$ as well. Hence, for these values of c, our bilinear form is coercive, and by the Lax-Milgram theorem, our problem $a(u,v) = F(v)$ admits a unique solution.

2 Numerical Approach

2.1 Implementation of a $\mathbb{P}_1$ FEM

In our numerical approach, we solve 1.2, but in the space $X_h^1 \cap H_0^1$ that consists of piecewise linear functions that have homogenous Dirichlet boundary conditions.

In order to work with this numerically, we have mainly followed the procedure presented in class and in the notebook "A P1 FEM for 1d Poisson". The main difference can be found in the stiffness matrix, in which we have needed more elemental matrices in order to accomodate the bilinear form derived earlier in this project. We have found the following elemental matrices, which are all added together in order to make one combined elemental matrix:

$$A_1^K = \frac{\alpha}{h_i}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad A_2^K = \frac{b}{2}\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, \quad A_3^K = \frac{ch_i}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad A^K = A_1^K + A_2^K + A_3^K$$

As an example we have produced the stiffness matrix for the case $\alpha = 1, b = 2, c = 3$, which is the combination of parameters we will mainly use throughout the project. We have also used $N = 7$ subintervals, with `np.random.seed(0)` for reproduceability, since we use random, variable step sizes. The matrix is presented in 8.

Examples of our numerical solutions can be found in figures 1 and 2, alongside with log-log plots and error bounds which we will discuss later. We see that the numerical approximation lies close to our exact solution, with only 50 nodes, which are not even equispaced.

2.2 Numerical Testing and Convergence Results

Although the previously presented numerical solutions qualitatively look pretty similar to the analytical, we now aim to investigate this further, and try to quantify the numerical error. One way of doing this is considering the convergence rate. In order to do this, we have considered a sequence of equispaced grids, because this simplifies the theory compared to doing it for general grids, as we do not have the same direct opportunity to control the step size.

We will look at the convergence rate in the L^2 and H^1 norm. These are norms on functions that involve integrals, so we will have to approximate these integrals with a quadrature when computing the norms. This introduces a small error, so we have to make sure that the error is smaller than what we get by FEM. We have chosen to use the trapezoidal rule, since this is of second order. The integral can then be considered as the sum of areas of triangles. If we let U and U_h denote the function values of u and u_h at the nodes, we have

$$\|u - u_h\|_{L^2} = \sqrt{\int_0^1 (u - u_h)^2 \, \mathrm{d}x} \approx \sqrt{h \sum_{i=1}^{M} (U - U_h)_i^2}$$

For the H^1 norm, we also need to consider the term $\|(u - u_h)_x\|_{L^2}$. The derivative of the hat functions are constant, so by Trapezoidal rule we can simply approximate the integral to be the area of rectangles:

$$\|(u - u_h)_x\|_{L^2} \approx \sqrt{2h \sum_{i=1}^{M} \left((U - U_h)_i \frac{1}{h} \right)^2} = \sqrt{\frac{2}{h} \sum_{i=1}^{M} (U - U_h)_i}$$

We have then tested the convergence rate of our implemented method with the functions

$$u_1(x) = x(1 - x) \qquad u_2(x) = \sin(3\pi x)$$

For the H^1 norm, we get order of convergence $p \approx 1.0$. This is in line with the theory, which states that $\|u - u_h\|_{H^1} \leq Ch|u - u_h|_{H^2}$

For the L^2 norm we observe that the order is estimated to be 2, as we also experienced in our previous project, working with finite difference methods. This also makes sense when we consider that the L^2 norm evaluates the error made in u, while the H^1 norm also takes the derivatives into consideration.

Compared to the methods we have previously utilized, this approach has been easier to implement, and it has also been fascinating to see how the approach also tackles general grids so effortlessly.

2.3 H^1 Error Bound and More Theoretical Results

In order to derive an error bound in the H^1 norm, we want to utilize the result of Cea's lemma, i.e. that for the solution of our infinite dimensional problem, u, and the solution for our finite dimensional problem, u_h, we have that

$$||u - u_h||_{H^1} \leq \frac{M}{C}||u - v_h||_{H^1}$$

Where M is the constant we found when we considered the continuity of $a(u, v)$ previously, and C is the coercivity constant. We know that Cea's lemma holds when the assumptions of the Lax-Milgram theorem holds, which we have already shown. From the proof of Cea's lemma we know that the lemma holds when the problem has Galerkin orthogonality, so we first focus on showing this. Galerkin orthogonality means that for the previously mentioned solutions u and u_h, we have that

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

And since we have a bilinear a, we here have that for these solutions:

$$a(u - u_h, v_h) = a(u, v_h) - a(u_h, v_h) = F(v_h) - F(v_h) = 0$$

Where the second to last equality follows from the fact that both functions are solutions to our problems in the spaces H_0^1 and $X_h^1 \cap H_0^1$ respectively.

Now, by starting from the result of Cea's lemma, we have derived the following interpolation error estimate in class

$$||e||_{H^1} \leq 2h||u''||_{L^2(0,1)} \tag{7}$$

These error bounds have also been included in our previously presented log-log plots of the errors, and can be found in figures 1 and 2. By evaluating 7 for $u(x) = x(1 - x)$, we get an error bound of $||e||_{H^1} \leq 4h$. For $u(x) = \sin(3\pi x)$, we get a bound of $||e||_{H^1} \leq 9\sqrt{2}\pi^2 h$. We observe that our results from the part about Numerical testing and convergence results are consistent with these bounds for the H^1 norm.

2.4 Non-Smooth Functions

Up until this point, we have mainly focused on smooth functions, as we also have in the finite differences part of the course. Now, however, we aim to solve our problem for the two following non-smooth exact solutions:

$$w_1(x) = \begin{cases} 2x, & x \in (0, \frac{1}{2}) \\ 2(1 - x), & x \in (\frac{1}{2}, 1) \end{cases} \qquad w_2(x) = x - |x|^{2/3}$$

We consider these functions' piecewise derivatives

$$w_1'(x) = \begin{cases} 2, & x \in (0, \frac{1}{2}) \\ -2, & x \in (\frac{1}{2}, 1) \end{cases} \qquad w_2'(x) = \frac{-2|x| + 3x\sqrt[3]{|x|}}{3x\sqrt[3]{|x|}}$$

We observe that $w_1'(x)$ is discontinuous at $x = \frac{1}{2}$, and the same holds for $w_2'(x)$ at $x = 0$, so neither w_1 nor w_2 is differentiable at all points in $[0, 1]$. Although these functions are not differentiable on our domain, we now aim to show that they are still elements in H^1, but not H^2. Thus our FEM will need some alterations before we tackle the problem of solving our PDE for these functions, but it will still be able to do the task. For a function to be an element of H^1, we need that its weak derivative exists, and analogously for a function to be an element of H^2, we need that its weak derivative also has a weak derivative. We first check this for the function $w_1(x)$ below

$$\begin{aligned}\int_0^1 w_1\phi' \mathrm{d}x &= 2\left(\left[x\phi\right]_0^{\frac{1}{2}} - \int_0^{\frac{1}{2}} \phi \mathrm{d}x + \left[(1-x)\phi\right]_{\frac{1}{2}}^1 + \int_{\frac{1}{2}}^1 \phi \mathrm{d}x\right) \\ &\overset{\phi(0)=0=\phi(1)}{=} \int_0^{\frac{1}{2}} (-2)\phi \mathrm{d}x + \int_{\frac{1}{2}}^1 2\phi \mathrm{d}x = -\int_0^1 v_1\phi \mathrm{d}x\end{aligned}$$

Meaning that the function has a weak derivative, v_1, which in this case coincides with the piecewise derivative, $w_1'(x)$, presented earlier. Now, we consider the existence of a weak derivative for v_1, and apply the fundamental theorem of calculus to get that

$$\int_0^1 v_1\phi' \mathrm{d}x = 2\left(\phi\left(\frac{1}{2}\right) - \phi(0) - \phi(1) + \phi\left(\frac{1}{2}\right)\right) \overset{\phi(0)=0=\phi(1)}{=} 4\phi\left(\frac{1}{2}\right) \neq -\int_0^1 v_1'\phi \mathrm{d}x$$

Where we cannot have equality above because there exists no function v_1' which could make the relation hold for all test functions ϕ. This means that $w_1(x) \notin H^2$.

Similarly for $w_2(x)$ we first determine its weak derivative as follows, where we ignore the absolute value sign because we consider an interval where x only takes positive values

$$\int_0^1 w_2\phi' \mathrm{d}x = \left[\left(x - x^{\frac{2}{3}}\right)\phi\right]_0^1 - \int_0^1 \left(1 - \frac{2}{3}x^{-\frac{1}{3}}\right)\phi \mathrm{d}x = -\int_0^1 v_2\phi \mathrm{d}x$$

Meaning that the weak derivative exists, and is equal to $1 - \frac{2}{3}x^{-\frac{1}{3}}$, and $w_2(x) \in H^1$. Now we consider the existence of a weak derivative for the weak derivative

$$\int_0^1 v_2\phi' \mathrm{d}x = \lim_{\epsilon\to 0}\left(\left[\left(1 - \frac{2}{3}x^{-\frac{1}{3}}\right)\phi\right]_\epsilon^1 - \frac{2}{9}\int_\epsilon^1 x^{-\frac{4}{3}}\phi \mathrm{d}x\right) \overset{\phi(1)=0}{=} \lim_{\epsilon\to 0}\left(\left(\frac{2}{3}\epsilon^{-\frac{1}{3}} - 1\right)\phi(\epsilon) - \frac{2}{9}\int_\epsilon^1 x^{-\frac{4}{3}}\phi \mathrm{d}x\right)$$

The first term in our last expression will now generally approach infinity as ϵ approaches zero, and thus the integral diverges. This means that the weak derivative of w_2 does not have a weak derivative itself, and $w_2 \notin H^2(0, 1)$.

Since our functions are now no longer elements of H^2, we have to adjust our functional F accordingly. In order to do this, we started with our weak formulation from earlier, namely:

$$\int_0^1 \left(\alpha(x)\partial_x u\partial_x v - b(x)u\partial_x v + c(x)uv\right) \mathrm{d}x = \int_0^1 f(x)v \mathrm{d}x$$

Since we know our exact solution is either w_1 or w_2, that $v = \phi$ and α, b and c are constants, the left hand side can be computed exactly. The result of the integral is then our new load vector.

It is clear from figures 3 and 4 that the convergence for these non-smooth functions are different to functions in H^2. For w_1 the error is close to machine precision regardless of our step length, which makes sense as the FEM should approximate a linear function exactly. However, if we do not have a node at 0.5 the situation becomes more complex as we then need to consider both $2x$ and $2(1-x)$ in the same sub-interval. For w_2 we observe the same general trends as for functions in H^2, just slightly worse convergence rates. This stems from the fact the w_2 is non-smooth at $x = 0$, which slightly worsens the convergence rates for both norms.

2.5 Distribution Dependency of the Nodes

Now we want to take a different approach to our problem, and aim to solve it for a given source term, rather than to check how good our numerical solution reconstructs an exact solution. Our chosen source term is $f(x) = 1$, which is clearly an element of the space $L^2(0, 1)$. We have also used $a = 1$, $b = 100$, $c = 0$.

The main point of interest for this function is to analyze how the error behaves depending on the location of the nodes. This choice of source term provides a boundary layer near $x = 1$, for which the solutions decreases rapidly. In order to investigate this further, we have used a fixed amount of points on two different grids, one with equispaced nodes, and one with uneven intervals that get progressively smaller towards $x = 1$. In figure 5 we present our numerical solution for the two types of grids, along with error plots

We note that increasing the number of nodes close to our singularity decreases the error significantly. If we were to do this for a finite difference method, this would have been more difficult, as we would have to implement it with variable step lengths. A different approach would be to increase the number of points on our whole domain, which could be superfluous on the opposite part of the domain.

The original problem was to consider the source term $f(x) = x^{-\frac{1}{4}}$. Here we tried to consider nodes with smaller sub-intervals near $x = 0$, as we have a singularity here. The result is presented in figure 6. Here we see that the error clearly decreased close to the singularity, however it also increases toward $x = 1$ as we have fewer nodes there. An alternative solution could be to simply add more nodes closer to $x = 0$.

3 Conclusion

Throughout this project we have derived several theoretical results regarding existence and uniqueness of a solution for our chosen problem. Having done this, we implemented a numerical approach to solve our problem for different densities and source terms. This was rather painfree for the smooth functions, and we also saw that we were able to tackle problems we have not previously been able to solve with our finite differences approach, although it was at the cost of our convergence rate. Furthermore, we have experienced the power of being able to place our nodes wherever is most beneficial. This helped with singularities, either inside our domain or on the boundary.

4 Appendix

Example of stiffnes matrix

$$\begin{bmatrix} 1.784 & -1.149 & 0. & 0. & 0. & 0. & 0. & 0. \\ -3.149 & 11.154 & -7.188 & 0. & 0. & 0. & 0. & 0. \\ 0. & -9.188 & 262.806 & -253.43 & 0. & 0. & 0. & 0. \\ 0. & 0. & -255.43 & 273.026 & -17.509 & 0. & 0. & 0. \\ 0. & 0. & 0. & -19.509 & 41.818 & -22.164 & 0. & 0. \\ 0. & 0. & 0. & 0. & -24.164 & 37.729 & -13.396 & 0. \\ 0. & 0. & 0. & 0. & 0. & -15.396 & 18.296 & -2.369 \\ 0. & 0. & 0. & 0. & 0. & 0. & -4.369 & 4.796 \end{bmatrix} \tag{8}$$

Figure 1: $u(x) = x(1-x)$

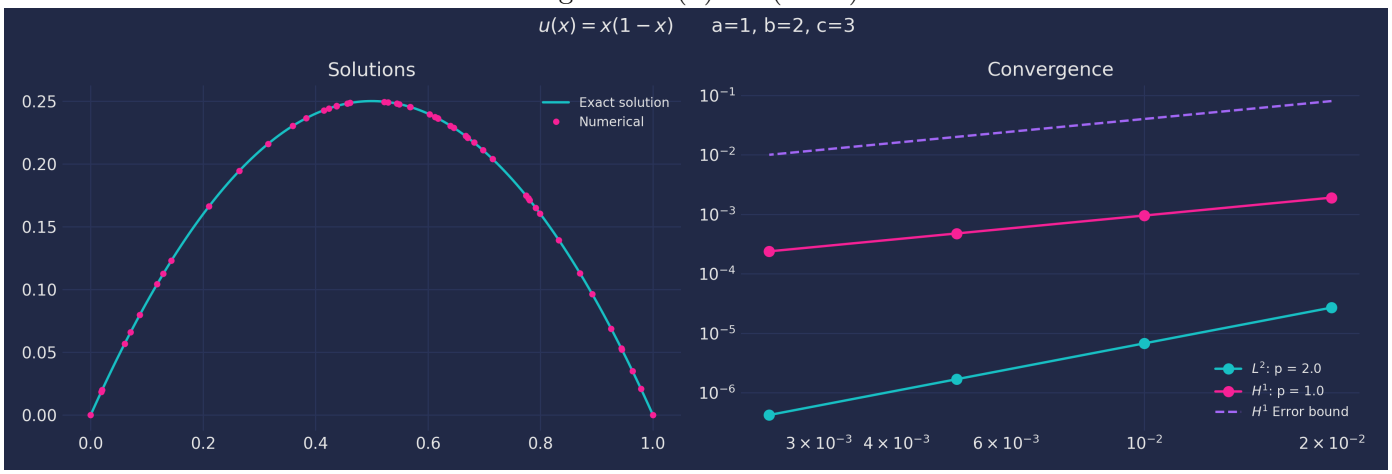

Figure 2: $u(x) = sin(3\pi x)$

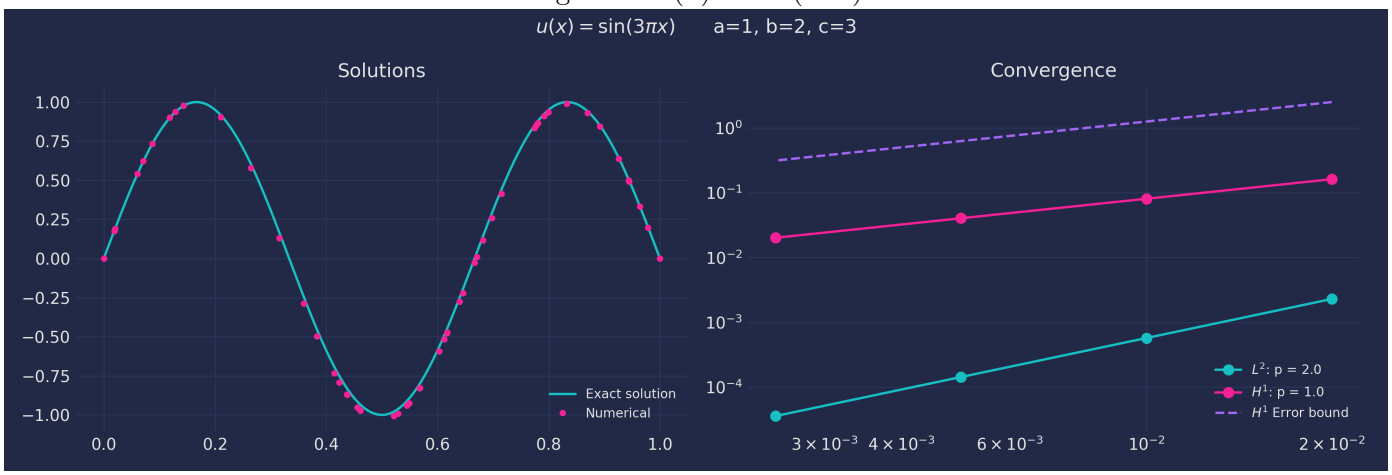

Figure 3: $w_1(x)$

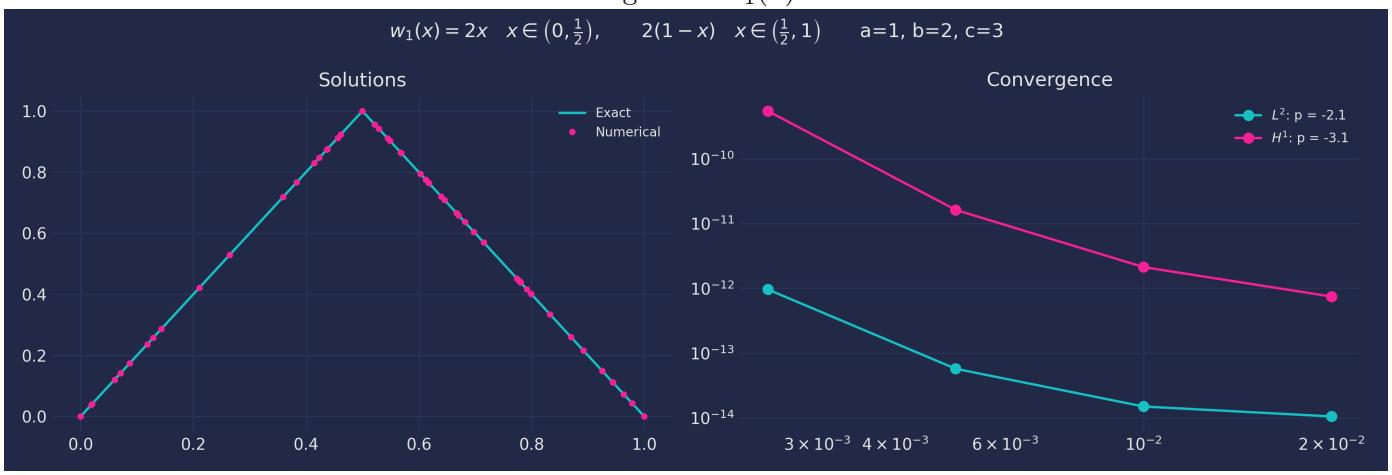

Figure 4: $w_2(x)$

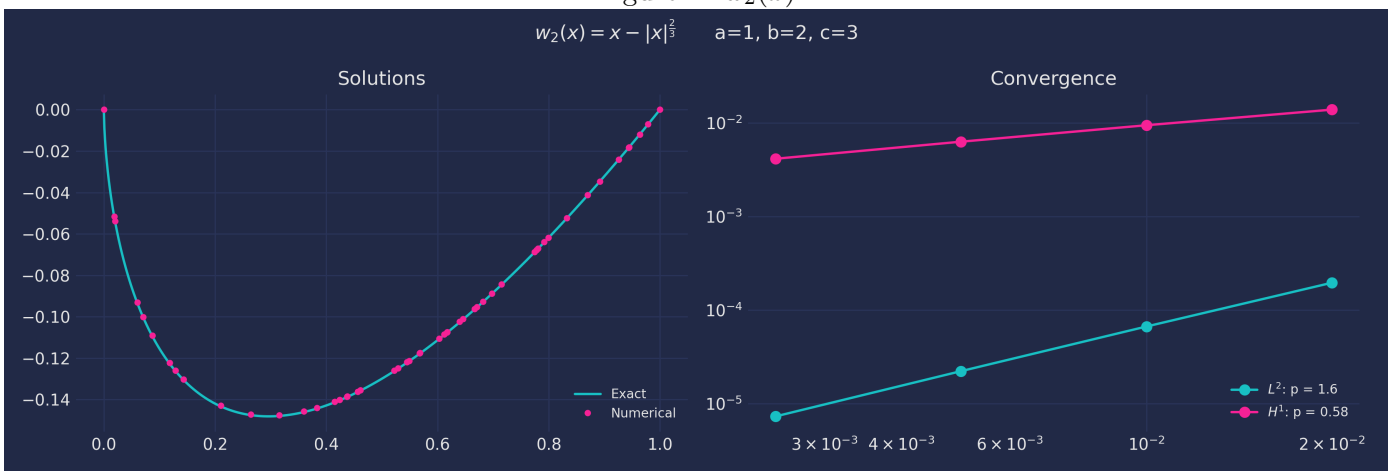

Figure 5: Distribution Of Nodes

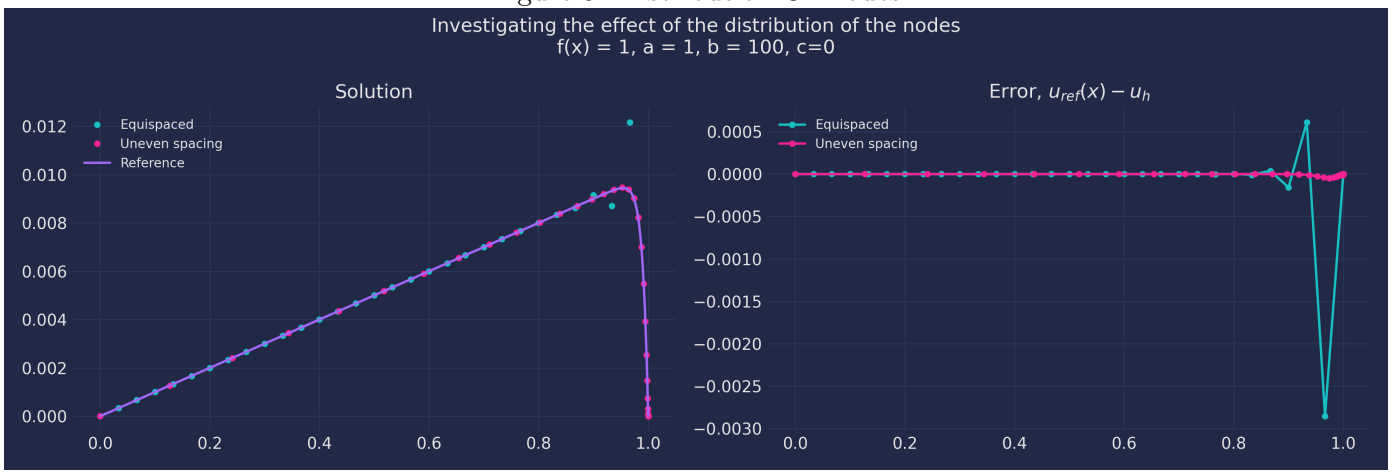

Figure 6: Original problem, with $f(x) = x^{-\frac{1}{4}}$