TMA4212 Project 2

Stationary convection diffusion problems using finite difference methods

August Rolfsen, Ine Daiwei Zhao, and Matias Sunde Øiesvold

March 2023

1 Introduction

This paper presents a numerical solution of a 1d stationary convection diffusion problem using finite element methods. Finite element methods are a particular type of Galerkin methods, and is widely applied in industry. Here, we consider a model for a convective and diffusive substance in 1d in the stationary regime. The model is a boundary problem for a Poisson like equation with Dirichlet boundary conditions. Assume that $\alpha(x) \geq \alpha_0 > 0, c > 0, \|\alpha\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty} + \|f\|_{L^\infty} \leq K$. We have scaled the problem such that the domain is $(0,1)$, so our problem can be formulated as

$$\begin{cases} -\partial_x(\alpha(x)\partial_x u) + \partial_x(b(x)u) + c(x)u = f(x) \quad \text{in} \quad \Omega = (0,1) \\ u(0) = 0 = u(1) \end{cases} \tag{1}$$

where u is the concentration of the substance, $\alpha(x) > 0$ is the diffusion coefficient, $b(x)$ is the fluid velocity, $c(x) \geq 0$ is the decay rate of the substance, $f(x)$ is a source term.

2 Theory

2.1 Weak variational form, and the function space

Let $H_0^1(0,1)$ be the space of functions $v \in L^2$ where a weak derivative $v' \in L^2$ exists. By multiplying our problem (1) by $v \in H_0^1(0,1)$, and integrating over the domain, we get the following,

$$\int_0^1 -\partial_x(\alpha(x)\partial_x u)v + \partial_x(b(x)u)v + c(x)uv\,dx = \int_0^1 f(x)v\,dx, \quad v \in H_0^1(0,1)$$

Using integration by parts,

$$\implies \big[-\alpha(x)u_x v\big]_0^1 - \int_0^1 -\alpha u_x v_x dx + \big[b(x)uv\big]_0^1 - \int b(x)uv_x dx + \int_0^1 c(x)uv\,dx$$

Given the Dirichlet boundary conditions, problem (1) satisfies

$$a(u,v) := \int_0^1 \alpha(x)u_x v_x + b(x)uv_x + c(x)uv\,dx = \int_0^1 fv\,dx =: F(v) \quad \forall v \in H_0^1(0,1) \tag{2}$$

where a is a bilinear form, and F is a linear functional. With our notation above, we have $b \leq 0$.

2.2 Unique solution

We wish to show there exists a unique solution $u \in H_0^1(0,1)$ of (2). This can be proven by the Lax-Milgram theorem.

The Lax-Milgram theorem
Let V be a Hilbert space. Suppose that F is a bounded linear functional, and that a is a continuous, coercive bilinear form, then the variational problem admits a unique solution.

First, we must prove that our problem satisfies these conditions.

Claim 1: a is bilinear
Definition: a is bilinear if $a(c_1u_1 + c_2u_2, v) = c_1a(u_1, v) + c_2a(u_2, v)$, where c_1, c_2 are constants.
Proof:

$$\begin{aligned}
a(u,v) :&= \int_0^1 \alpha(x)u_xv_x + b(x)uv_x + c(x)uv\,dx \\
a(c_1u_1 + c_2u_2, v) &= \int_0^1 \alpha(x)(c_1u_1 + c_2u_2)_xv_x + b(x)(c_1u_1 + c_2u_2)v_x + c(x)(c_1u_1 + c_2u_2)v\,dx \\
&= c_1\int_0^1 \alpha(x)u_{1_x}v_x + b(x)u_1v_x + c(x)u_1v\,dx + c_2\int_0^1 \alpha(x)u_{2_x}v_x + b(x)u_2v_x + c(x)u_2v\,dx \\
&= c_1a(u_1, v) + c_2a(u_2, v)
\end{aligned}$$

In a similar way, it can be shown that $a(u, c_1v_1 + c_2v_2) = c_1a(u, v_1) + c_2a(u, v_2)$, using the linearity of the integral.

Claim 2: a is continuous
Definition: a is continuous if $|a(u, v)| \leq M\|u\|_{H^1}\|v\|_{H^1}, \forall u, v \in H_0^1(0, 1)$, where $M > 0$ is a constant.
Proof:

$$\begin{aligned}
|a(u,v)| &\leq \|\alpha\|_{L^\infty}\int_0^1 |u_x||v_x|\,dx + \|b\|_{L^\infty}\int_0^1 u|v_x|\,dx + \|c\|_{L^\infty}\int_0^1 uv\,dx \\
&\leq \|\alpha\|_{L^\infty}\|u_x\|_{L^2}\|v_x + |_{L^2} + \|b\|_{L^\infty}\|u\|_{L^2}\|v_x\|_{L^2} + \|c\|_{L^\infty}\|u\|_{L^2}\|v\|_{L^2} \\
&\leq M\|u\|_{H^1}\|v\|_{H^1}
\end{aligned}$$

where $M = \|\alpha\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty}$

Claim 3: a is coercive
Definition: $a(u, v)$ is coercive if $a(u, u) \geq \gamma\|u\|_{H^1}^2 \quad \forall \quad u \in H_0^1(0, 1)$, where $\gamma > 0$ is a constant.
Proof: We can show that $a(u, v)$ is coercive when $c > \frac{|b|^2}{2\alpha}$ with Gårding and Young's inequality.

$$a(u,u) \geq (\alpha - \frac{\epsilon}{2}|b|)\int_0^1 u_x^2 dx + (c - \frac{1}{2\epsilon}|b|)\int_0^1 u^2 dx \quad \forall \quad \epsilon > 0 \tag{3}$$

As this holds for all ϵ, it will in particular hold for $\epsilon = \frac{\alpha}{|b|}$, we get

$$\begin{aligned}
a(u,u) &\geq \big(\alpha - \frac{\alpha}{2|b|}|b|\big)\int_0^1 u_x^2 dx + \big(c - \frac{|b|^2}{2\alpha}\big)\int_0^1 u^2 dx \\
&\geq \frac{\alpha}{2}\int_0^1 u_x^2 dx + \big(c - \frac{|b|^2}{2\alpha}\big)\int_0^1 u^2 dx \geq \gamma\big(\int_0^1 u_x^2 dx + \int_0^1 u^2 dx\big) = \gamma\|u\|_{H_0^1(0,1)}^2
\end{aligned}$$

where we define $\gamma = \min(\frac{\alpha}{2}, c - \frac{|b|^2}{2\alpha})$

Claim 4: F is linear
Definition: F is linear if $F(c_1v_1 + c_2v_2) = c_1F(v_1) + c_2F(v_2)$.
Proof:

$$\begin{aligned}
F(v) :&= \int_0^1 fv\,dx \\
F(c_1v_1 + c_2v_2) &= \int_0^1 f(c_1v_1 + c_2v_2)\,dx \\
&= c_1\int_0^1 fv_1\,dx + c_2\int_0^1 fv_2\,dx \\
&= c_1F(v_1) + c_2F(v_2)
\end{aligned}$$

Claim 5: F is bounded
Proof:
We have $f \in L^2$. From the definition of boundedness for Hilbert space, we get the following with the Cauchy-Schwartz inequality and by definition $\|v\|_{L^2} \leq \|v\|_{H^1}$:

$$\|F\| = \max_{v\neq 0, v\in H^1} \frac{|F(v)|}{\|v\|_{H^1}} \leq \max_{v\neq 0, v\in H^1} \frac{\|f\|_{L^2}\|v\|_{L^2}}{\|v\|_{H^1}} \leq \max_{v\neq 0, v\in H^1} \frac{\|f\|_{L^2}\|v\|_{H^1}}{\|v\|_{H^1}} = \|f\|_{L^2} < \infty$$

$\implies F$ is finite, and therefore bounded.

Given our 5 claims above, we can apply the Lax-Milgram theorem, and conclude that there exists a unique solution $u \in H_0^1(0,1)$ of our problem (2).

2.3 Galerkin approximation

From here, we assume that $\alpha > 0, b, c > 0$ be non-zero constants. We already have a weak formulation of our problem, which is (2). Now, we are interested in solving the problem (2) with a Galerkin method on a general grid in our domain Ω. This means that we want to restrict the functions u, v in (2) to be in a finite dimensional space $V_h \subset H_0^1(0,1)$. To achieve this, we need to choose a finite dimensional function space together with a basis. Then we can expand u and v with this basis, and set up a linear system of equations to be solved.

Finite element methods are a specific type of Galerkin methods where the functions spaces are piecewise continuous polynomials. In our case, we consider the space X_h^1 of linear piecewise continuous functions, together with the basis $\{\phi_i\}_{i=0}^{M+1}$. The basis is defined as

$$\phi_i = \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & , x_{i-1} < x < x_i \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} & , x_i < x < x_{i+1} \\ 0 & , \text{otherwise} \end{cases}$$

This results in our finite dimensional space $V_h = X_h^1 \cap H_0^1(0,1)$, and any $v \in V_h$ can be expressed as a linear combination of the basis functions, $v = \sum_{i=0}^{M+1} v_i \phi_i(x)$ where v_i are uniquely determined coefficients.

To assemble the linear system of equations, we replace the space V with V_h, and express $u, v \in V_h$ in terms of the basis. Then we have the following system,

$$A\vec{U} = \vec{F} \tag{4}$$

$$\begin{bmatrix} a(\phi_1,\phi_1) & a(\phi_1,\phi_2) & \cdots & a(\phi_1,\phi_M) \\ a(\phi_2,\phi_1) & a(\phi_2,\phi_2) & \cdots & a(\phi_2,\phi_M) \\ \vdots & & \ddots & \\ a(\phi_M,\phi_1) & a(\phi_M,\phi_2) & \cdots & a(\phi_M,\phi_M) \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_M \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_M \end{bmatrix}$$

where we want to solve for $\vec{U}$. For our problem (2) the stiffness matrix without including the boundary is

$$A = \text{tridiag}\Big[-\frac{\alpha}{h_{i-1}} - \frac{b}{2} + \frac{c}{6}h_{i-1}, \frac{\alpha}{h_i} + \frac{c}{3}h_i + \frac{\alpha}{h_{i+1}} + \frac{c}{3}h_{i+1}, -\frac{\alpha}{h_{i+1}} + \frac{b}{2} + \frac{c}{6}h_{i+1} \Big]$$

A printout of the matrix $\vec{U}$ can be found in (5). Naturally as a result, the Lax-Milgram theorem also holds for the solution $u_h \in X_h^1(0,1) \cap H_0^1(0,1)$.

2.4 Stability and convergence of Galerkin methods

To prove convergence of the Galerkin method of our problem, we need to show that following two lemmas below hold. Then, we can later find an H^1 error bound to see how close our approximation is to the solution.

Lemma 1: Galerkin orthogonality
Let $u \in V$, $u_h \in V_h$ be the solutions of the infinite and finite dimensional variational problems respectively. Then,

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

For our $\mathbb{P}_1$ FEM (V_h), we can show that this holds. Since a is bilinear and with the definition of our problem from (2) and (4),

$$a(u - u_h, v_h) = a(u, v_h) - a(u_h, v_h) = F(v_h) - F(v_h) = 0 \tag{5}$$

Thus, showing that Galerkin orthogonality holds for our $\mathbb{P}_1$ FEM (V_h).

Lemma 2: Cea's lemma
Let $u \in V$, $u_h \in V_h$ be the solutions of the infinite and finite dimensional variational problems respectively, and suppose that the hypotheses of Lax-Milgram theorem are satisfied. Notably, we assume that a is continuous and coercive with constants M and γ. Then,

$$\|u - u_h\|_V \leq \frac{M}{\gamma}\|u - v_h\|_V, \quad \forall v_h \in V_h \tag{6}$$

For our problem, we can show with Claim 3 (coercivity),

$$\gamma\|u - u_h\|_{H^1}^2 \leq |a(u - u_h, u - u_h)| = |a(u - u_h, u - v_h + v_h - u_h)|$$

With Claim 2 of bilinearity, we have that,

$$|a(u - u_h, u - v_h + v_h - u_h)| = |a(u - u_h, u - v_h)| + |a(u - u_h, v_h - u_h)|$$

Due to (5), the second term on the right hand side equals zero, and with Claim 2 it follows,

$$\gamma\|u - u_h\|_{H^1}^2 \leq |a(u - u_h, u - v_h + v_h - u_h)| = |a(u - u_h, u - v_h)| \leq M\|u - u_h\|_{H^1}\|u - v_h\|_{H^1}$$

$$\implies \|u - u_h\|_{H^1} \leq \frac{M}{\gamma}\|u - v_h\|_{H^1} \quad \forall v_h \in X_h^1 \cap H_0^1(0,1)$$

Thus, showing that Cea's lemma holds for our problem.

2.5 H^1 Error bound

It is known that $\|u - I_h u\|_{H^1} \leq 2h\|u''\|_{L^2}$ and $I_h u \in X_h^1 \cap H_0^1$. Therefore we can show the following error bound,

$$\|u - u_h\|_{H^1} \overset{(6)}{\leq} \frac{M}{\gamma} \inf_{v_h \in X_h^1 \cap H_0^1} \|u - v_h\|_{H^1} \leq \frac{M}{\gamma}\|u - I_h u\|_{H^1} \leq \frac{2M}{\gamma} h\|u''\|_{L^2}$$

$$\implies \|u - u_h\|_{H^1} \leq \frac{2M}{\gamma} h\|u''\|_{L^2} \tag{7}$$

3 Numerical results

3.1 Smooth functions

We consider two potential exact solutions satisfying the Dirichlet boundary conditions, and their convergence rates.

$$u_1(x) := x(1-x) \tag{8}$$

$$u_2(x) := \sin(3\pi x) \tag{9}$$

In figure (3.1) we plot $u_1(x)$ using both unevenly and evenly spaced nodes:

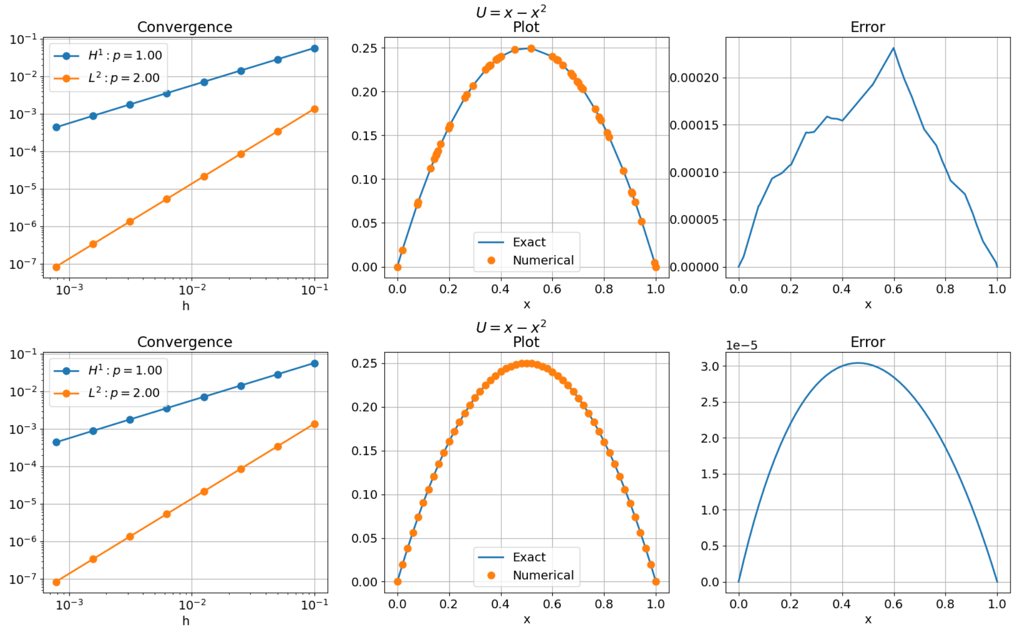

Figure 1: Plots of unevenly spaced and equidistant nodes of the function $u_1(x)$. Plots for $u_2(x)$ are shown in the appendix

As shown above, the order of convergence in H^1-norm is 1, which fits perfectly with the calculated error bound in (7). Using the L^2-norm we get quadratic convergence, which agrees with theory: We expect order 2 convergence for functions in $H_0^2(0,1)$, and both u_1 and u_2 are in $H_0^2(0,1)$.

3.2 Non-smooth functions

Now we will look at two non-smooth potential exact solutions.

$$w_1(x) = \begin{cases} 2x & , x \in (0, \frac{1}{2}) \\ 2(1-x) & , x \in (\frac{1}{2}, 1) \end{cases} \tag{10}$$

$$w_2(x) = x - |x|^{\frac{2}{3}} \tag{11}$$

w_1 is not differentiable in $x = \frac{1}{2}$ and w_2 is not differentiable in $x = 0$, both due todifferent left and right derivatives. However, they both have weak derivatives, and are therefore in H^1. For w_1 using the standard method of multiplying with a test function v, integrating over Ω, splitting the integral into the intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ and using integration by parts, one can show that

$$w_1'(x) = \begin{cases} 2 & , x \in (0, \frac{1}{2}) \\ -2 & , x \in (\frac{1}{2}, 1) \end{cases} = 2 - 4\theta(x - \frac{1}{2})$$

where $\theta(x)$ is the Heaviside step function. $w_1'(x)$ is obviously in L^2, which means $w_1 \in H^1$. However, the weak derivative of the Heaviside step function is the Dirac Delta "function", which is not an element in L^2.

As for w_2, we note that our domain is $(0,1)$ and as such we can drop the absolute value:

$$w_2'(x) = 1 - \frac{2}{3}x^{-\frac{1}{3}} \implies \|w_2'(x)\|_{L^2}^2 = \int_0^1 (1 - \frac{2}{3}x^{-\frac{1}{3}})^2 dx = \frac{1}{3} < \infty$$

which means that the $w_2' \in L^2(0,1)$ and $w_2 \in H_0^1(0,1)$. However, w_2'' is not in $L^2(0,1)$:

$$\|w_2''(x)\|_{L^2}^2 = \int_0^1 (w_2'')^2 dx = \lim_{k\to 0}\int_k^1 (\frac{2}{9}x^{-\frac{4}{3}})^2 = \lim_{k\to 0}\int_k^1 (\frac{4}{81}x^{-\frac{8}{3}})dx = \infty$$

which shows that $w_2 \notin H_0^2(0,1)$.

In order to implement this we must determine $\vec{F}$ analytically.

$$\int_0^1 v_i(-\alpha u'' + bu' + cu)dx = \int_0^1 f v_i dx = F(v_i) = F_i$$

We show how to calculate the term containing u'', the other terms are calculated in a similar manner applying integration by parts and the Fundamental theorem of calculus.

$$-\int_0^1 v_i \alpha u'' dx = -\left[u'\phi_i\right]_0^1 + \int_0^1 v_1 \alpha \phi_i' dx = \int_{K_i} \alpha v' \phi_i' dx + \int_{K_{i+1}} \alpha v' \phi_{i+1}' dx$$

Since $\phi_i(0) = \phi_i(1) = 0$. Now we use the fundamental theorem of calculus to get the following expression:

$$= \alpha\left(\frac{u(x_i) - u(x_{i+1})}{h_i} + \frac{u(x_i) - u(x_{i-1})}{h_{i+1}}\right)$$

Which we can use in the implementation. We plot the functions, errors, and convergence:

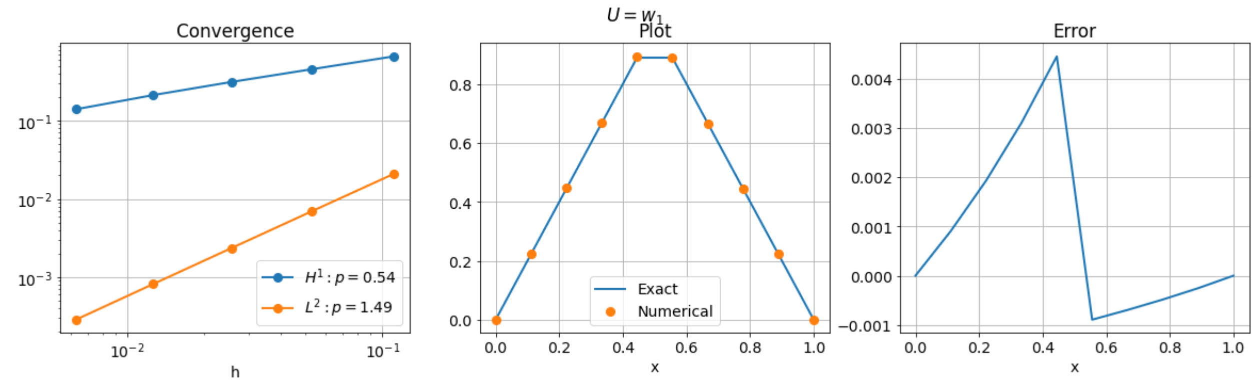

Figure 2:

We see that we still appear to have convergence for w_1. Convergence is not something one can expect in general, for a function that is not in $H_0^2(0,1)$. We still have a guaranteed existence of solution since we have $w_1 \in H_0^1(0,1)$. We see that the H^1-convergence is approximately 0.5, which is less than the guaranteed rate of 1 for H^2-functions. Similarly, we see that the L^2-convergence is approximately 1.5, which is less than 2, which is the guaranteed minimum convergence rate of in L^2.

We also made the observation that if we include the point $x = \frac{1}{2}$, we would have machine precision. This is due to the fact that this is exactly the same as running the method over the intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ separately. The error is then less than $\|u''\|_{L^2} = 0$ evaluated on each of these open intervals seperately. A plot of this can be seen in the appendix, figure (7).

For w_2, we see in figure (3) see the consequences of the function not being in $H_0^2(0,1)$, as the convergence rates are smaller than what a $H_0^2(0,1)$ function would have. We see that the largest error is near $x = 0$, where $\|u''\|$ goes to ∞.

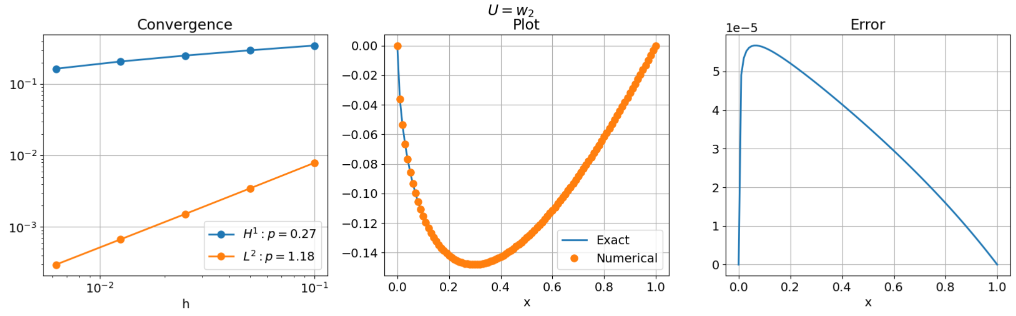

Figure 3:

3.3 Adjusting node placement

Finally we want to observe the numerical solution when

$$f(x) = x^{-\frac{1}{4}} \quad , x \in (0,1) \tag{12}$$

when we alter the node placement.

It's clear that $f(x) \in L^2(0,1)$: $\|f(x)\|^2 = \int_0^1 (x^{-\frac{1}{4}})^2 = \left[2\sqrt{x}\right]_0^1 = 2 < \infty$ which means that the Lax-Milgram theorem holds, and we have a unique solution.

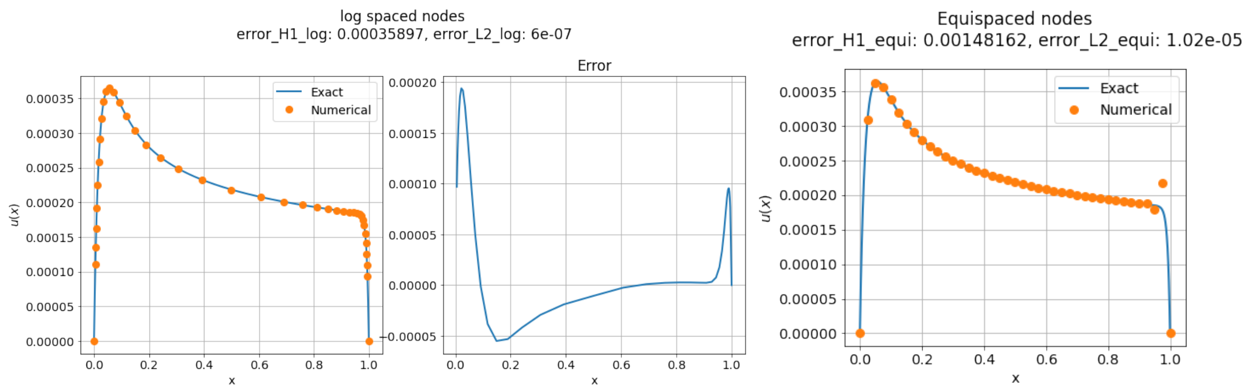

Figure 4: Left: More nodes near $x = 0$ and $x = 1$. Middle: Error plot corresponding to the leftmost plot. Right: Equispaced nodes.

In the leftmost plot we have concatinated two logspaces together in order to have a denser node placement near $x = 0$ and $x = 1$, as this is where we have the largest error.

We see that node placement is critical for the error. In the H^1-norm, we get an error of 0.003 in the leftmost plot, whereas with equispaced nodes we get an error of 0.00148. For the L^2-norm, the log-spaced plot gives an error of $6 \cdot 10^{-7}$, whereas equispaced nodes give $1 \cdot 10^{-5}$.

4 Conclusion

In this paper we have implemented 1-dimentional finite element methods, and experimented with various parameters such as comparing random node placement equispaced nodes, as well as analysing functions with different smoothness properties, and verifying that the our code agrees with the results expected from theory. It has been quite successful, as all the results we obtained were reasonable, expected results.

5 Appendix

```
A:
 [[  7.038  -2.829   0.      0.      0.      0.    ]
 [ -1.829  21.944 -19.442   0.      0.      0.    ]
 [  0.    -18.442  57.328 -38.768   0.      0.    ]
 [  0.      0.    -37.768  45.662  -7.647   0.    ]
 [  0.      0.      0.     -6.647  42.25  -35.352]
 [  0.      0.      0.      0.    -34.352  45.112]]
F:
 [0.718 0.627 0.119 0.253 0.258 0.195]

Xk: [0.    0.259 0.655 0.708 0.734 0.872 0.901 1.   ]
```

Figure 5: Stiffness matrix A from x_1 to x_6 without boundary(x_0 and x_7)

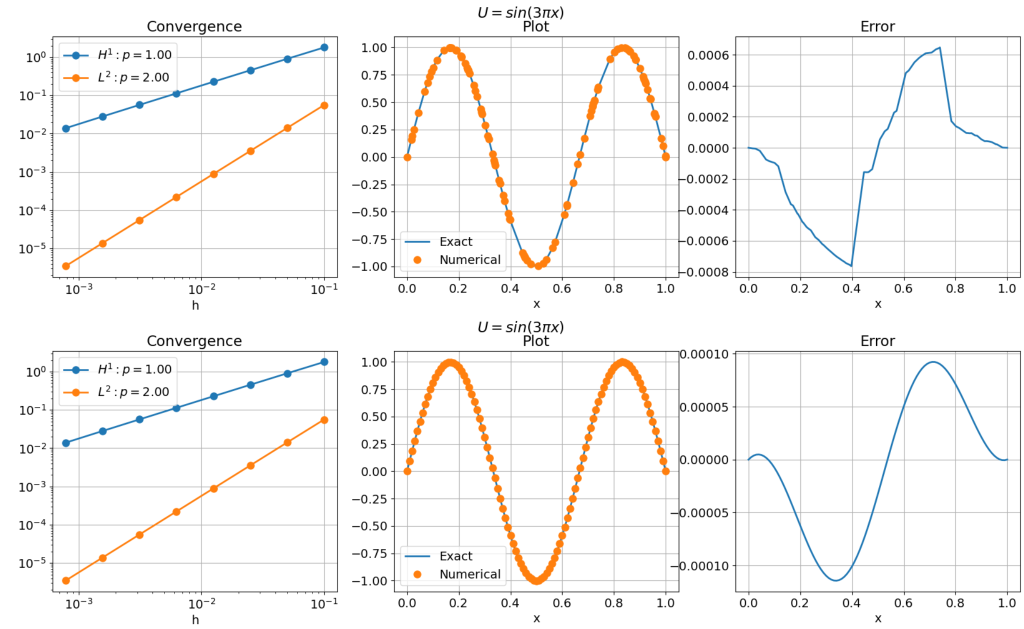

Figure 6: Plots of unevenly spaced and equidistant nodes of the function $u_2(x)$

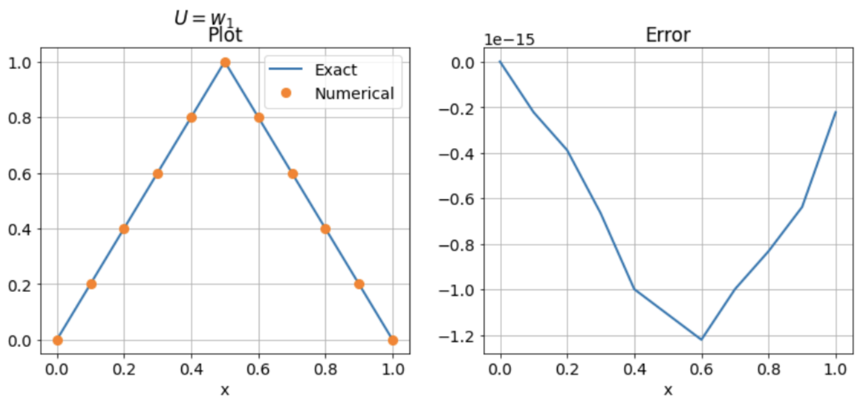

Figure 7: convergence plot of w_1 when including a node at $x = \frac{1}{2}$