Modelling stationary convection diffusion problem using FEM

March 2023

1 Introduction

The convection-diffusion equation describes the phenomena where physical quantities are transferred inside a system due to diffusion and convection processes. Thus, an area of study with many applications in engineering, as well as in applied physics. This paper will focus on the 1d case governing the transport/mixing/decay of a chemical in a fluid moving in a finite tube. The boundary value problem can be modeled in the following way:

$$-\mathcal{L}u = -\partial_x\left(\alpha(x)\partial_x u\right) + \partial_x(b(x)u) + c(x)u = f(x), \qquad \text{in} \qquad \Omega = (0,1) \tag{1}$$

This PDE arises from some well known physical principles and results, i.e. conservation of mass, Fick's law for diffusion and Reynold's transport theorem.

2 Theoretical background

Assume $\alpha(x) \geq \alpha_0 > 0, c > 0, \|\alpha\|_{L^\infty} + \|b\|_{L^\infty} + \|c\|_{L^\infty} + \|f\|_{L^2} \leq K$, and that we have Dirichlet boundary conditions $u(0) = 0 = u(1)$.

Claim: Any classical solution of (1) can be expressed as $a(u,v) = F(v)\ \forall v \in H_0^1(0,1)$, for some form $a : H^1 \times H^1 \to \mathbb{R}$, and operator $F : H^1 \to \mathbb{R}$.

Proof: Consider $(-\mathcal{L}u)v = f(x)v$, for some $v \in H_0^1(0,1)$.

$$-\partial_x\left(\alpha(x)\partial_x u\right)v + \partial_x(b(x)u)v + c(x)uv = f(x)v$$
$$\implies \int_0^1 (-\partial_x\left(\alpha(x)\partial_x u\right))v\,dx + \int_0^1 (\partial_x(b(x)u))v\,dx + \int_0^1 (c(x)u)v\,dx = \int_0^1 f(x)v\,dx$$

Preforming integration by parts on the first two terms:

$$-\left[(\alpha(x)\partial_x u))v\right]_0^1 + \int_0^1 \left(\alpha(x)\partial_x u\right)\partial_x v\,dx + \left[(b(x)u)v\right]_0^1 - \int_0^1 (b(x)u)\partial_x v\,dx + \int_0^1 (c(x)u)v\,dx = \int_0^1 f(x)v\,dx$$
$$\overset{v(0)=v(1)=0}{\implies} \int_0^1 \left(\alpha(x)\partial_x u\right)\partial_x v\,dx - \int_0^1 (b(x)u)\partial_x v\,dx + \int_0^1 (c(x)u)v\,dx = \int_0^1 f(x)v\,dx$$

Define a and F:

$$a(u,v) \triangleq \int_0^1 \left(\alpha(x)\partial_x u\right)\partial_x v - (b(x)u)\partial_x v + (c(x)u)v\,dx, \quad F(v) \triangleq \int_0^1 f(x)v\,dx \tag{2}$$

□

Claim: $F : H^1 \to \mathbb{R}$, where $\mathbb{R}$ is equipped with the usual norm, is a bounded linear functional.

Proof: Consider the operator norm of F:

$$\|F\|_{\text{op}} \triangleq \sup_{v\in H_0^1, v\neq 0} \frac{\|F(v)\|_{\mathbb{R}}}{\|v\|_{H^1}} = \sup_{v\in H_0^1, v\neq 0} \frac{\left|\int_0^1 f(x)v(x)dx\right|}{\|v\|_{H^1}} \overset{\text{C.S.}}{\implies} \|F\|_{\text{op}} \leq \frac{\|f\|_{L^2}\|v\|_{L^2}}{\|v\|_{H^1}} \leq \|f\|_{L^2} \overset{f\in L^2}{<} \infty$$

For linearity, observe: $F(v+u) = \int_0^1 f(x)(v(x)+u(x))dx = \int_0^1 f(x)v(x)dx + \int_0^1 f(x)u(x)dx = F(v) + F(u)$. The claim follows. □

Claim: $a : H^1 \times H^1 \to \mathbb{R}$ is a bilinear continuos form.

Proof: Linearity in the first argument follows directly from linearity of integration and differentiation. Assume $\xi \in \mathbb{R}$ and $u, v, t \in H^1$

$$\begin{aligned}
a(\xi t+u,v) &= \int_0^1 (\alpha(x)\partial_x(\xi t+u))\,\partial_x v-(b(x)(\xi t+u))\partial_x v+(c(x)(\xi t+u))v dx\\
&= \int_0^1 (\alpha(x)\xi\partial_x t+\alpha(x)\partial_x u)\,\partial_x v(b(x)\xi t+b(x)u)(\partial_x v)+(c(x)\xi t+c(x)u)v dx\\
&= \xi\int_0^1 (\alpha(x)\partial_x t)\partial_x v(b(x)t)\partial_x v+(c(x)t)v dx+\int_0^1 (\alpha(x)\partial_x u)\partial_x v(b(x)u)\partial_x v+(c(x)u)v dx\\
&= \xi a(t,v)+a(u,v)
\end{aligned}$$

This shows a is linear in the first argument. An analogous proof may be preformed for the second argument.

Recall that for bilinear forms, boundedness implies continuity.

Claim: a is bounded.
Proof:

$$\begin{aligned}
|a(u,v)| &= \left|\int_0^1 (\alpha(x)\partial_x u)\,\partial_x v-(b(x)u)\partial_x v+(c(x)u)v dx\right| \le \left|\langle \alpha\partial_x u,\partial_x v\rangle_{L^2(0,1)}\right|+\left|\langle bu,\partial_x v\rangle_{L^2(0,1)}\right|+\left|\langle cu,v\rangle_{L^2(0,1)}\right|\\
&\overset{\text{C.S}}{\le} \|\alpha\|_{L^\infty}\|\partial_x u\|_{L^2}\|\partial_x v\|_{L^2}+\|b\|_{L^\infty}\|u\|_{L^2}\|\partial_x v\|_{L^2}+\|c\|_{L^\infty}\|u\|_{L^2}\|v\|_{L^2}\\
&\overset{\text{P.I}}{\le} \|\alpha\|_{L^\infty}\|\partial_x u\|_{L^2}\|\partial_x v\|_{L^2}+2\|b\|_{L^\infty}\|\partial_x u\|_{L^2}\|\partial_x v\|_{L^2}+\|c\|_{L^\infty}\|u\|_{L^2}\|v\|_{L^2}\\
&= (\|\alpha\|_{L^\infty}+2\|b\|_{L^\infty})\|\partial_x u\|_{L^2}\|\partial_x v\|_{L^2}+\|c\|_{L^\infty}\|u\|_{L^2}\|v\|_{L^2} \le (\|\alpha\|_{L^\infty}+2\|b\|_{L^\infty})\|u\|_{H^1}\|v\|_{H1}+\|c\|_{L^\infty}\|u\|_{H^1}\|v\|_{H^1}\\
&= (\|\alpha\|_{L^\infty}+2\|b\|_{L^\infty}+\|c\|_{L^\infty})\|u\|_{H^1}\|v\|_{H^1} = \varpi\|u\|_{H^1}\|v\|_{H^1}
\end{aligned}$$

Where we have used that $\|f\|_{H^1} = \|f\|_{L^2}+\|\partial_x f\|_{L^2}$, meaning $\|f\|_{L^2}, \|\partial_x f\|_{L^2} \le \|f\|_{H^1}$. □

Claim: Suppose $\alpha, b, c \in \mathbb{R}\setminus\{0\}$, then a satisfies Gårding's inequality.

Proof: This can be shown in a direct manner:

$$\begin{aligned}
a(u,u) &= \int_0^1 \alpha(\partial_x u)^2-(bu)\partial_x u+cu^2 dx = \alpha\int_0^1 (\partial_x u)^2 dx-b\int_0^1 u\partial_x u dx+c\int_0^1 u^2 dx\\
&\overset{\text{Y. I.}}{\ge} \alpha\int_0^1 (\partial_x u)^2 dx-\frac{\varepsilon}{2}|b|\int_0^1 (\partial_x u)^2-\frac{1}{2\varepsilon}|b|\int_0^1 u^2 dx+c\int_0^1 u^2 dx\\
&= \left(\alpha-\frac{\varepsilon}{2}|b|\right)\int_0^1 u_x^2 dx+\left(c-\frac{1}{2\varepsilon}|b|\right)\int_0^1 u^2 dx \quad \text{for all} \quad \varepsilon > 0 \quad \square
\end{aligned}$$

The above inequality can further be used to show coerciveness of a.

Claim: Suppose $c > \frac{|b|^2}{2\alpha}$, then a is coercive, i.e. $\exists\ \varrho \in \mathbb{R}^{>0}$ such that $a(u,u) \ge \varrho\|u\|_{H^1}^2 \forall\ h \in H^1$.

Proof: Firstly, observe the following

$$\begin{aligned}
a(u,u) &= \left(\alpha-\frac{\varepsilon}{2}|b|\right)\int_0^1 u_x^2 dx+\left(c-\frac{1}{2\varepsilon}|b|\right)\int_0^1 u^2 dx = \left(\alpha-\frac{\varepsilon}{2}|b|\right)\|u_x\|_{L^2}^2+\left(c-\frac{1}{2\varepsilon}|b|\right)\|u\|_{L^2}^2\\
&\overset{c>\frac{|b|^2}{2\alpha}}{\ge} \left(\alpha-\frac{\varepsilon}{2}|b|\right)\|u_x\|_{L^2}^2+\left(\frac{|b|^2}{2\alpha}-\frac{1}{2\varepsilon}|b|\right)\|u\|_{L^2}^2
\end{aligned}$$

Recall that Gårding´s inequality holds for all $\varepsilon \in \mathbb{R}^{>0}$. Notably, ε can be chosen such that the coefficients in the first and second terms are positive, i.e.:

$$\alpha-\frac{\varepsilon}{2}|b| > 0 \quad\wedge\quad \frac{|b|^2}{2\alpha}-\frac{1}{2\varepsilon}|b| > 0 \quad\Longrightarrow\quad \left.\begin{array}{l} \alpha-\frac{\varepsilon}{2}|b| > 0 \implies \varepsilon < \frac{2\alpha}{|b|}\\ \frac{|b|^2}{2\alpha}-\frac{1}{2\varepsilon}|b| > 0 \implies \varepsilon > \frac{\alpha}{|b|}\end{array}\right\} \Longrightarrow \frac{\alpha}{|b|} < \varepsilon < \frac{2\alpha}{|b|}$$

Thus, for any ε on this interval we have:

$$a(u,u) \ge \left(\alpha-\frac{\varepsilon}{2}|b|\right)\|u_x\|_{L^2}^2+\left(\frac{|b|^2}{2\alpha}-\frac{1}{2\varepsilon}|b|\right)\|u\|_{L^2}^2 \ge \varrho\|u\|_{H^1}^2$$

Where $\varrho \triangleq \min\{\alpha - \frac{\varepsilon}{2}|b|, \frac{|b|^2}{2\alpha} - \frac{1}{2\varepsilon}|b|\}$. This proves the claim. □

In conclusion of the above discussion; a is both continuous and coercive, which by Lax-Milgram Theorem implies the existence of a unique solution $u \in H_0^1(0,1)$ for $a(u,v) = F(v)\ \forall\ v \in H_0^1(0,1)$.

Claim (Galerkin orthogonality): Let u and u_h be the solutions of the infinite and finite dimensional variational problems respectively. Then $a(u-u_h, v) = 0\ \forall v_h \in V_h$.
Proof: Choose any $v_h \in V_h$, and use the fact that a is bilinear.

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

Claim (Cea's lemma): Let u and u_h be the solutions of the infinite and finite dimensional variational problems respectively. Suppose that F is a continuous linear functional, and that a is a continuous, coercive bilinear form. Notably, assume that a is continuous and coercive with constants M and μ respectively. Then

$$\|u-u_h\|_{H^1} \leq \frac{M}{\mu} \inf_{v_h \in V_h} \|u-v_h\|_{H^1}$$

Proof: Due to a being coercive, the following inequality holds:

$$\|u-u_h\|_{H^1}^2 \leq \frac{1}{\mu}|a(u-u_h, u-u_h)|$$

Resulting from bilinearity, the term $|a(u-u_h, u-u_h)|$ can be expanded as

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

were the last equality holds due to the Galerkin orthogonality. Since a is continous with a constant M depending on the coefficients α, b and c, it can be obtained that

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

Inserting this back into the previous inequality, and dividing by $\|u-u_h\|_{H^1}$ on both sides yields

$$\|u-u_h\|_{H^1} \leq \frac{M}{\mu}\|u-v_h\|_{H^1}, \forall v_h \in V_h \quad \Leftrightarrow \quad \|u-u_h\|_{H^1} \leq \frac{M}{\mu} \inf_{v_h \in V_h} \|u-v_h\|_{H^1}$$

where μ and M are constants. Thus, Cea's lemma holds for this problem. □

Claim: The global error is bounded as

$$\|u-u_h\|_{H^1} \leq \frac{2M}{\mu}\|u''\|_{L^2} h \quad \text{and} \quad \|u-u_h\|_{L^2} \leq C\|u''\|_{L^2} h^2$$

where M, μ and C are constants.
Proof: From Cea's lemma it follows that

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

where $v_h \in V_h = X_h^1 \bigcap H_0^1(0,1) = \text{span}\{\phi_0, \dots, \phi_M\}$. ϕ_i being the hat function of height 1 on the interval $[x_{i-1}, x_{i+1}]$. Let $I_h u \in V_h$ denote the continuous piecewise linear function on the subdivision $\{x_0, ..., x_N\}$ which coincides with u at the mesh-points x_i for $i = 0, ..., N$. Thus

$$I_h u(x) = \sum_{i=0}^{N} u(x_i)\phi_i(x)$$

The function $I_h u$ is called the interpolant of u in the finite space space V_h. Choosing $v_h = I_h u$ to get that

$$\|u-u_h\|_{H^1} \leq \frac{M}{\mu}\|u-I_h u\|_{H^1}$$

To determine an H^1 error bound on the global error $u-u_h$, one can therefore instead seek a bound on the interpolation error $u-I_h u$ in the same norm. Claim without justification that the interpolation error $u-I_h u$ is bounded in the H^1 and L^2-norms as

$$\|u-I_h u\|_{H^1} \leq 2h\|u''\|_{L^2} \quad \text{and} \quad \|u-I_h u\|_{L^2} \leq h^2\|u''\|_{L^2}$$

Inserting this into the global errors to obtain obtain the error bounds

$$\|u-u_h\|_{H^1} \leq \frac{2M}{\mu}\|u''\|_{L^2} h \quad \text{and} \quad \|u-u_h\|_{L^2} \leq C\|u''\|_{L^2} h^2$$

□

Definition 1.1: Let $u \in L^1(\Omega)$, then $\tilde{u} \in L^1(\Omega)$ is a weak derivative of u if $\int_\Omega u\varphi' = -\int_\Omega \tilde{u}\varphi \ \forall \ \varphi \in C_c^\infty(\Omega)$. Note that any strong derivative u' is also a weak derivative of u.

Proposition 1: Let u' be the weak derivative of u on (a,b). Then for all intervals $(\alpha,\beta) \subset (a,b)$ it holds that $u'|_{(\alpha,\beta)}$ is also the weak derivative of $u|_{(\alpha,\beta)}$ on (α,β).

Proof: Let $(\alpha,\beta) \subset (a,b)$ and $\phi \in C_c^\infty(\alpha,\beta)$ and define the trivial extension of ϕ by $\tilde{\phi} \in C_c^\infty(a,b)$. Then, we conclude

$$\int_\alpha^\beta u(x)\phi'(x)dx = \int_a^b u(x)\tilde{\phi}'(x)dx = -\int_a^b u'(x)\tilde{\phi}(x)dx = -\int_\alpha^\beta u'(x)\phi(x)dx,$$

which implies the proposition. □

Define the following functions:

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

Claim: w_1 and w_2 is in $H^1(0,1)$ but not in $H^2(0,1)$.

Proof: First note that $w_1, w_2 \in L^2(0,1)$. Moreover, there derivatives are not defined in $x = \frac{1}{2}$ and $x = 0$ respectively. Recall that the functions in $H^1(0,1)$, are those where both the function and its weak derivative are in $L^2(0,1)$.

Let

$$g_\lambda(x) = \begin{cases} 2, & x \in \left(0, \frac{1}{2}\right) \\ \lambda \in \mathbb{R}, & x = 1/2 \\ -2, & x \in \left(\frac{1}{2}, 1\right) \end{cases}$$

Observe that $g_\lambda|_{(0,\frac{1}{2})} = 2$ and $g_\lambda|_{(\frac{1}{2},1)} = -2$ are strong derivatives of w_1 on their respective intervals. In particular they are weak derivatives. Recall that all weak derivatives are equivalent up to a set of measure zero. Then by proposition 1 we know that any weak derivative of w_1 must coincide with g on the intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$. Thus, making g_λ the only candidate for a weak derivative of w_1.

$$\begin{aligned} -\int_0^1 w_1(x)\varphi'(x)dx &= -2\int_0^{\frac{1}{2}} x\varphi' dx - 2\int_{\frac{1}{2}}^1 \varphi' dx + 2\int_{\frac{1}{2}}^1 x\varphi' dx \overset{I.B.P.}{=} -\varphi(\frac{1}{2}) + 2\int_0^{\frac{1}{2}} \varphi(x)dx + 2\varphi(\frac{1}{2}) - \varphi(\frac{1}{2}) - 2\int_{\frac{1}{2}}^1 \varphi(x)dx \\ &= 2\int_0^{\frac{1}{2}} \varphi(x)dx - 2\int_{\frac{1}{2}}^1 \varphi(x)dx = \int_0^1 g_\lambda(x)\varphi(x)dx \end{aligned}$$

Making g_λ the weak derivative of w_1. Notably, $g_\lambda \in L^2(0,1)$, implying $w_1 \in H_0^1(0,1)$. Again by proposition 1 and uniqueness up to a set of meassure zero, any weak derivative of g_λ must coincide with $\tilde{g}_\lambda(x) = 0 \ \forall \ x \in (0,1)$. One can observe that the definition is not fulfilled:

$$\begin{aligned} -\int_0^1 g(x)\varphi'(x)dx = -2\int_0^{\frac{1}{2}} \varphi'(x)dx + 2\int_{\frac{1}{2}}^1 \varphi'(x)dx = -2(\varphi(\frac{1}{2}) - \varphi(0)) + 2(\varphi(1) - \varphi(\frac{1}{2})) = -4\varphi(\frac{1}{2}) \\ \neq 0 = \int_0^1 \tilde{g}_\lambda(x)\varphi(x), \quad \varphi \in C_c^\infty(0,1) \end{aligned}$$

In conclusion w_1 does not have any weak derivative, implying $w_1 \notin H^2(0,1)$

Observe that $w_2 = x - x^{\frac{2}{3}} \ \forall x \in (0,1)$. Then by proposition 1 and uniqueness of the weak derivative the first and second order weak derivatives of w_2 must coincide with the first and second order strong derivatives, $w_2' = 1 - \frac{2}{3}x^{-\frac{1}{3}}$ and $w_2'' = \frac{2}{9}x^{-\frac{4}{3}}$, respectively. Consider the integrals:

$$\lim_{a \to 0^+} \int_a^1 w_1'(x)^2 dx = \lim_{a \to 0^+} \int_a^1 \left(1 - \frac{2}{3}x^{-\frac{1}{3}}\right)^2 dx = \lim_{a \to 0^+} \left[-2x^{2/3} + x + \frac{4x^{\frac{1}{3}}}{3} \right]_a^1 = \frac{1}{3} < \infty$$

$$\lim_{a \to 0^+} \int_a^1 w_1''(x)^2 dx = \lim_{a \to 0^+} \int_a^1 \left(\frac{2}{9}x^{-\frac{4}{3}}\right)^2 dx = \lim_{a \to 0^+} \left[-\frac{2}{3}x^{-\frac{1}{3}} \right]_a^1 = \infty$$

Therefore, $w_2' \in L^2(0,1)$ and $w_2'' \notin L^2(0,1)$. Hence, w_2 is in $H_0^1(0,1)$ and not in $H_0^2(0,1)$. □

Claim: Let $f(x) = x^{-\frac{1}{4}}$, then $f \in L^2(0,1)$.
Proof: For $f(x) = x^{-\frac{1}{4}}$ to be in $L^2(0,1)$, we simply check that it is square-integrable on $x \in (0,1)$. Notice that $f(x) > 0$ for all $x \in (0,1)$, which means that $|f(x)|^2 = f^2(x)$ for any $x \in (0,1)$.

$$\int_0^1 |f(x)|^2 dx = \int_0^1 f^2(x) dx = \int_0^1 x^{-\frac{1}{2}} dx = \left[2\sqrt{x}\right]_0^1 = 2 < \infty$$

Thus $f \in L^2(0,1)$. □

Numerical method and testing

By the finite elements method, we must construct our linear system to solve. The matrix A is given by $A = [a(\phi_i, \phi_j)]_{ij}$ where ϕ_i is the hat function. By integration it can be shown that the entries is given by:

$$\text{superdiagonal} = -\frac{\alpha}{h_{i+1}} + \frac{b}{2} + \frac{h_{i+1}}{6}c, \quad \text{diagonal} = \frac{\alpha}{h_i} + \frac{\alpha}{h_{i+1}} + \frac{h_i + h_{i+1}}{6}c, \quad \text{subperdiagonal} = -\frac{\alpha}{h_i} - \frac{b}{2} + \frac{h_i}{6}c$$

The stiffness matrix of size 6 × 6 is displayed in appendix A.

Furthermore we use Simpsons method for integrals to approximate $F(\phi_i) = \int_0^1 f(x)\phi_i(x)dx$. The node distribution is inconsequential to the numerical solver. Therefore, the nodes can be positioned arbitrarily as well as equidistantly. Importantly, the distribution of nodes can be focused or clustered in certain areas of the interval.

To impose w as an exact solution, one must determine

$$F_i = \int_0^1 (-\mathcal{L}w)(x)\phi_i(x) = \int_0^1 -\alpha\partial_x^2 w(x)\phi_i(x) + b\partial_x w(x)\phi_i(x) + cw(x)\phi_i(x)dx$$

When imposing an exact solution $w \in H_0^1(0,1) \setminus H_0^2(0,1)$ further steps is required to obtain the F vector. An immediate issue with such an exact solution is the non existence of $\partial_x^2 w$. However, this impediment can be resolved using integration by parts. Consider:

$$\begin{aligned}\int_0^1 \partial_x^2 w(x)\phi_i(x)dx &= \int_{x_{i-1}}^{x_i} \partial_x^2 w(x)\phi_i(x)dx + \int_{x_i}^{x_{i+1}} \partial_x^2 w(x)\phi_i(x)dx \\ &= \left[\phi_i(x)\partial_x w(x)\right]_{x_{i-1}}^{x_i} - \int_{x_{i-1}}^{x_i} \partial_x w(x)\partial_x\phi_i(x)dx + \left[\phi_i(x)\partial_x w(x)\right]_{x_i}^{x_{i+1}} - \int_{x_i}^{x_{i+1}} \partial_x w(x)\partial_x\phi_i(x)dx\end{aligned}$$

Recall, $\phi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_i}, & x \in (x_{i-1}, x_i) \\ \frac{x_{i+1} - x}{h_{i+1}}, & x \in (x_i, x_{i+1}) \end{cases}$. Hence, the integral can be resolved

$$-\frac{1}{h_i}\int_{x_{i-1}}^{x_i} \partial_x w(x)dx - \frac{1}{h_{i+1}}\int_{x_i}^{x_{i+1}} \partial_x w(x)dx = -\frac{1}{h_i}(w(x_i) - w(x_{i-1})) + \frac{1}{h_{i+1}}(w(x_{i+1}) - w(x_i)) = \frac{w(x_{i+1})}{h_{i+1}} + \frac{w(x_{i-1})}{h_i} - w(x_i)\Big(\frac{1}{h_i} + \frac{1}{h_{i+1}}\Big)$$

This approach was used when w_1 and w_2 were chosen as exact solutions.

We can see that numerical solution coincides with the exact solution. Both for equidistributed nodes like u_1 and for randomly distributed nodes like u_2 and u_3.

Choosing certain arbitrary values for α, b, c can lead to inadequate numerical outcomes, where the approximation fails to capture the analytic solution. The Lax-Milgram theorem provides insight into this issue, as when $\alpha = 1, b = 10, c = 1$, the bilinear form is noncoercive, meaning that Lax-Milgram cannot guarantee a unique solution. As a result, the FEM may yield a misleading numerical solution, as illustrated in the plot below. To avoid these issues, it is important to carefully select appropriate values for α, b, c in order to obtain an accurate numerical solution.

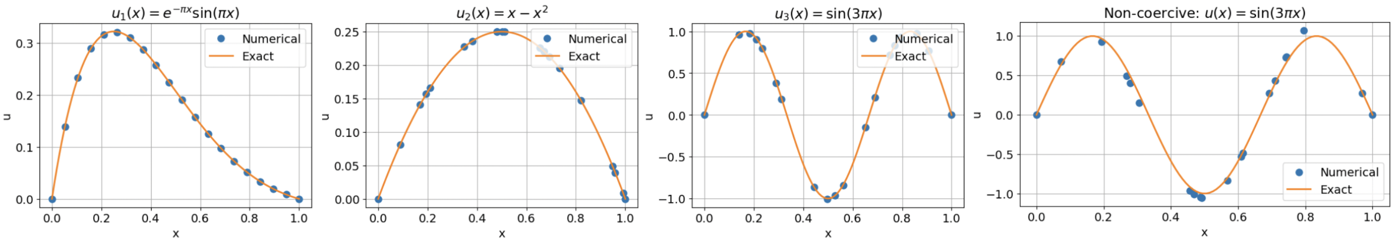


Figure 1: Numerical and exact solutions for different functions, the last of which being non-coercive.

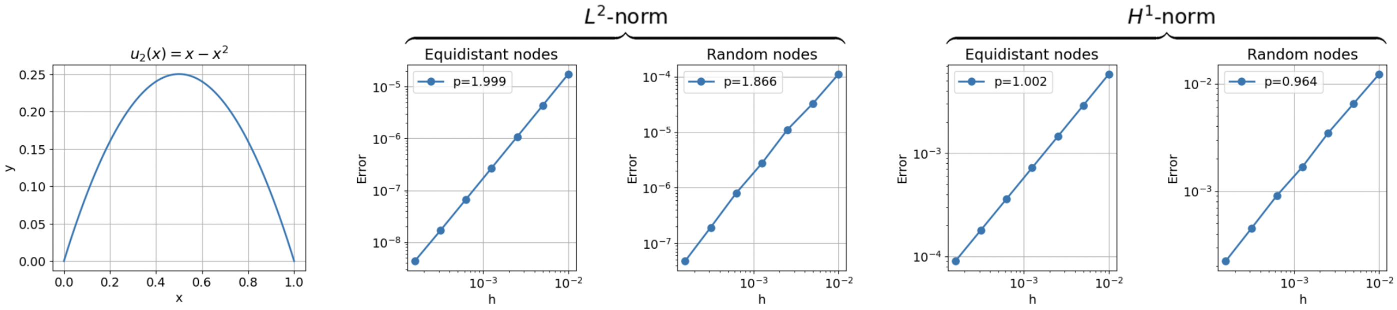


Figure 2: Numerical solution and convergence in L^2 and H_0^1 for $u_2(x) = x - x^2$ using $\alpha = b = c = 1$

Results and discussion

We see that finite elements method for this function has quadratic convergence in L^2 and linear in H^1. This corresponds to the upper bounds derived in the theory section.

Random node distributions brings about some deviations from the expected linear trend in the loglog plot, but the data clearly shows that the average slope in L^2 and H^1 is 2 and 1 respectively. When the number of nodes is reduced, the distribution can have a larger impact on the accuracy of the numerical solution. Certain configurations may result in large gaps along the interval, which can significantly reduce the resolution and accuracy of the model on the corresponding intervals. As a result, the numerical approximation may not perform as well in these cases, which might explain a ”more linear” error plot for the largest values of N.

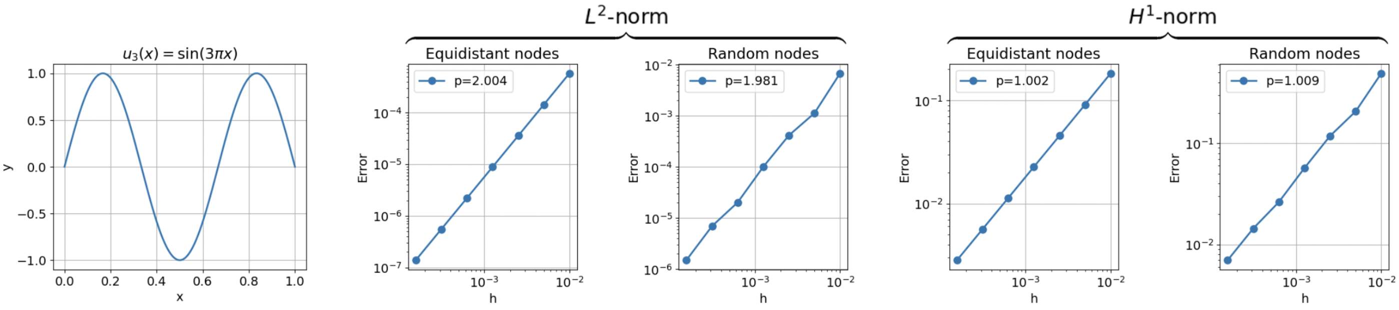


Figure 3: Numerical solution and convergence in L^2 and H_0^1 respectively for $u_3(x) = \sin(3\pi x)$ using $\alpha = b = c = 1$

For the hat function, we see linear convergence in L^2 and H^1. Contrary to u_1, u_2, the function, w_1, is not in H^2 which may explain the difference in convergence rate in L^2. The error bound which states quadratic convergence in L^2 assumes $w_1 \in H^2$ which is not the case.

The error plots might suggest a relationship between pointwise error and whether the germ of ω_i at point p is in H^2. We have seen that w_1 fails to be in H^2 since it's double weak derivative is nonmeassurable at $x = \frac{1}{2}$, which is the sole point responsible for $\omega_1 \notin H^2$ and where the largest error occurs.

Considering w_2, the convergence rate for L^2 is 1.17 and for H^1 is 0.176. Although these rates are unusual, the convergence in H^1 is poorer than in L^2, consistent with both theoretical predictions and our findings from w_1.

We suspect that the grid points near $x = 0$ are causing some issues, due to boundary effects. Since the function does not belong to H^2, the error bound derived in the theory section cannot be relied upon. Specifically, the nodes in the vicinity of 0 may have low or

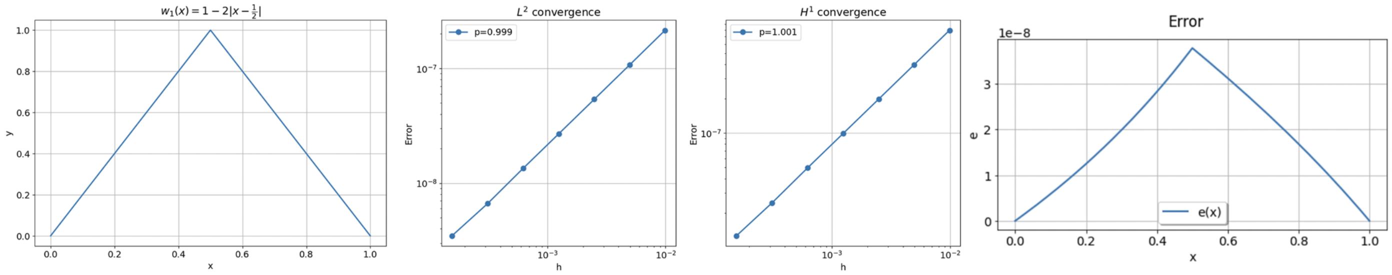


Figure 4: Numerical solution, error and convergence in L^2 and H_0^1 respectively for $\omega_1(x) = 1-2|x-\frac{1}{2}|$ using $\alpha = b = c = 1$

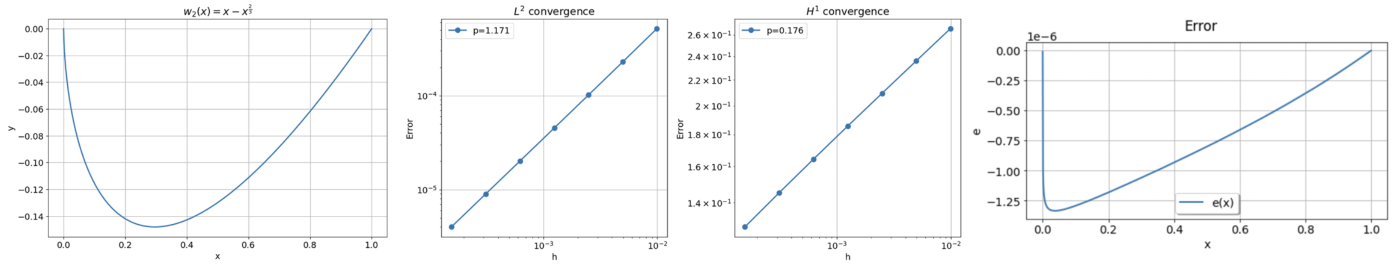


Figure 5: Numerical solution, error and convergence in L^2 and H_0^1 respectively for $\omega_2(x) = x - x^{\frac{2}{3}}$ using $\alpha = b = c = 1$

no convergence, due to the nature of the solution, thereby lowering the global convergence rate. The error plot illustrates how the error spikes at the nodes adjacent to $x = 0$, further implying a relationship between pointwise error and whether the germ is in H^2

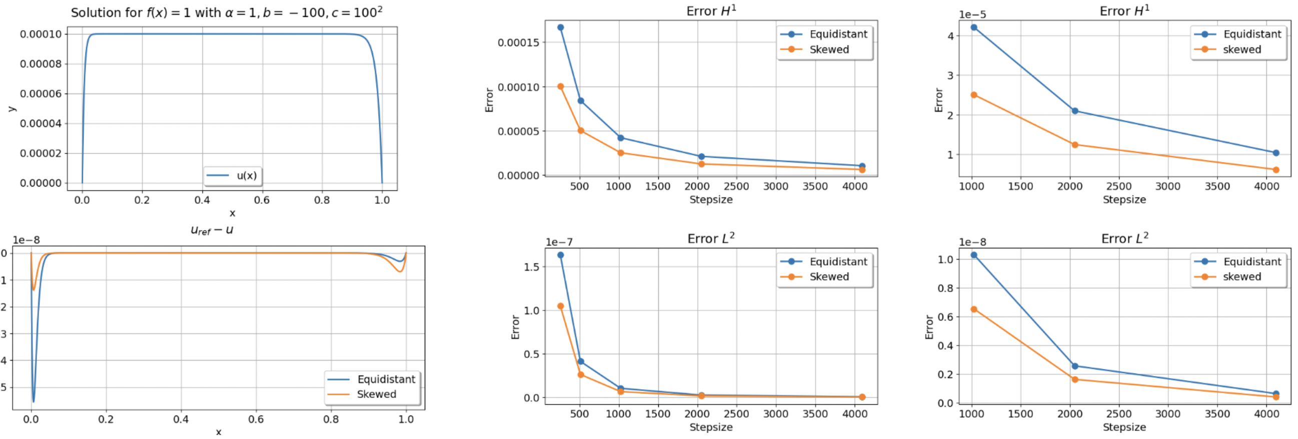


Figure 6: Numerical solution, error and convergence in L^2 and H_0^1 respectively for $f(x) = x^{-\frac{1}{4}}$ using $\alpha = 1, b = -100, c = 100^2$

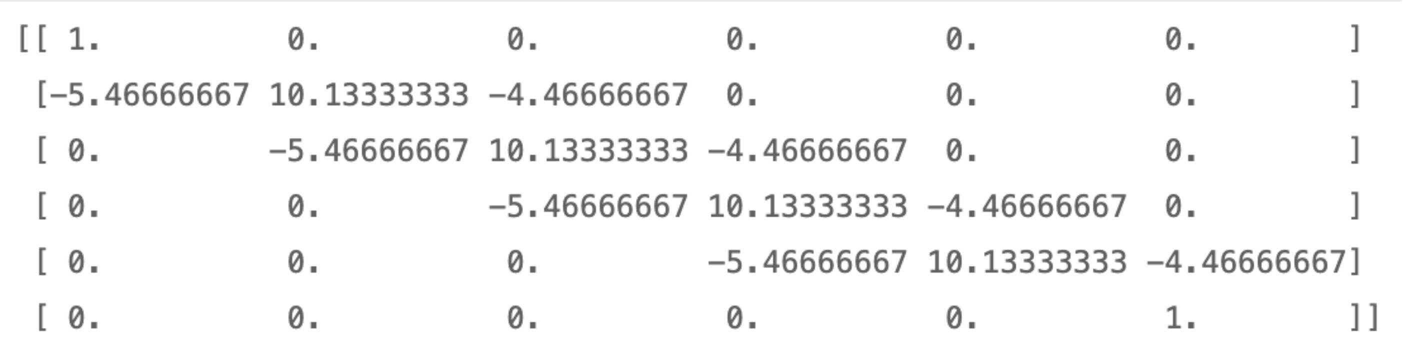
Using these specific values for α, b, and c, the resulting solution u exhibits a sharp rise in a neighborhood of $x = 0$, followed by a rapid but smoother decline as x approaches 1.

The error plot showcases how the inaccuracy is distributed. In both cases we see how the error is largest for earlier x. The skewed nodes are considerably better in this interval and slightly worse in the latter part. The general trend indicates that skewing the node distribution towards the steepest regions gives a better approximation of the reference solution. Both in L^2 and H^1-norm the trend seems to hold. The "skewed error" is slightly below "equidistant error" for all the test values of N. These results suggests it is more important to capture abrupt changes.

Conclusion

Having a continuous coercive bilinear form we get a unique solution in V_h which we can find by using the finite elements method. The convergence rate in L^2 and H^1 seemed to be 2 and 1 respectively for function in H^2, and 1 and 1 respectively for function in H^1. The error of the numerical solution was more significant in neighborhoods having a nonsmooth solution. Finally, having more nodes around regions with rapid decay/increase gave a better approximation.

A Figures



```
[[ 1.          0.          0.          0.          0.          0.        ]
 [-5.46666667 10.13333333 -4.46666667  0.          0.          0.        ]
 [ 0.         -5.46666667 10.13333333 -4.46666667  0.          0.        ]
 [ 0.          0.         -5.46666667 10.13333333 -4.46666667  0.        ]
 [ 0.          0.          0.         -5.46666667 10.13333333 -4.46666667]
 [ 0.          0.          0.          0.          0.          1.        ]]
```

Figure 7: Stiffness matrix for equidistant nodes

B Abbreviations and Theorems

C.S.: Cauchy Schwarz, i.e $|\langle u, v\rangle| \leq \|u\|\|v\|$
Y.I.: Youngs inequality, i.e. $ab \leq \frac{1}{2\varepsilon}a^2 + \frac{\varepsilon}{2}b^2 \ \forall \ \varepsilon > 0$
P.I.: Poincaré inequality, i.e. $\|u\|_{L^p(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)} \ \forall \ u \in H_0^1(\Omega)$
I.B.P.: Integration by parts

Lax-Milgram´s theorem: Let V be a Hilbert space. Suppose that F is a continuous linear functional (it suffices that F be bounded), and that a is a continuous, coercive bilinear form, i.e. there exist M, α such that
1. (Continuous) $a(u, v) \leq M\|u\|_V\|v\|_V$ for all $u, v \in V$
2. (Coercive) $a(v, v) \geq \alpha\|u\|_V^2$ for all $v \in V$ The the variational problem: find $u \in V$ such that, for all $v \in V$,

$$a(u, v) = F(v)$$

admits a unique solution.