Quasi-optimality of FEMs for the Helmholtz equation

T. van Beeck†, U. Zerbinati*

* *Mathematical Institute*
University of Oxford

† *Institute for Numerical and Applied Mathematics*
University of Göttingen

University of Oxford
Mathematical Institute

Oxford Mathematics

The wave equation

The **wave equation** is a second-order linear partial differential equation for the description of waves.

The wave equation

The **wave equation** is a second-order linear partial differential equation for the description of waves.

In particular, it describes the evolution of an **excess pressure** p_δ in a medium with speed of sound c.

The wave equation

University of Oxford
Mathematical Institute

The **wave equation** is a second-order linear partial differential equation for the description of waves.

In particular, it describes the evolution of an **excess pressure** p_δ in a medium with speed of sound c.

$$
\begin{aligned}
\partial_t^2 p_\delta - c^2 \Delta p_\delta &= f && \text{in} \quad \Omega \times (0, T), \\
p_\delta &= 0 && \text{on } \partial\Omega \times (0, T), \\
p_\delta(\cdot, 0) = p_0(\cdot), &\qquad \partial_t p_\delta(\cdot, 0) = \dot{p}_0(\cdot).
\end{aligned}
$$

Time harmonic solutions of the wave equation

The **time harmonic solutions** of the wave equation are of the form

$$p^{(\delta)}(\cdot, t) = \mathrm{Re}\left\{P(\cdot)e^{-i\omega t}\right\},$$

where $P : \Omega \to \mathbb{C}$ and ω is the **angular frequency**.

Time harmonic solutions of the wave equation

The **time harmonic solutions** of the wave equation are of the form

$$p^{(\delta)}(\cdot, t) = \mathrm{Re}\left\{P(\cdot)e^{-i\omega t}\right\},$$

where $P : \Omega \to \mathbb{C}$ and ω is the **angular frequency**.
Substituting this ansatz into the wave equation, we obtain the **Helmholtz equation**

$$\begin{aligned} -\Delta P - k^2 P = F \qquad & \text{in} \quad \Omega, \\ u = 0 \qquad & \text{on } \partial\Omega, \end{aligned}$$

where $k = \omega/c$ is the **wave number** and $F = \mathrm{Re}\left\{f(\cdot, t)e^{-i\omega t}\right\}$.

The weak formulation

The **weak formulation** of the Helmholtz equation is to find $u \in H_0^1(\Omega)$ such that

$$(\nabla P, \nabla Q)_{L^2} - k^2 (P, Q)_{L^2} = {}_{H^{-1}}\langle F, Q \rangle_{H_0^1} \quad \forall Q \in H_0^1(\Omega),$$

where $(\cdot, \cdot)_{L^2}$ denotes the L^2 inner product.

The weak formulation

The **weak formulation** of the Helmholtz equation is to find $u \in H_0^1(\Omega)$ such that

$$(\nabla P, \nabla Q)_{L^2} - k^2 (P, Q)_{L^2} = {}_{H^{-1}}\langle F, Q \rangle_{H_0^1} \quad \forall Q \in H_0^1(\Omega),$$

where $(\cdot, \cdot)_{L^2}$ denotes the L^2 inner product.

The weak formulation of the Helmholtz equation is **elliptic**, hence we can make use of **elliptic regularity theory**.

The weak formulation

UNIVERSITY OF OXFORD
Mathematical Institute

The **weak formulation** of the Helmholtz equation is to find $u \in H_0^1(\Omega)$ such that

$$(\nabla P, \nabla Q)_{L^2} - k^2 (P, Q)_{L^2} = {}_{H^{-1}}\langle F, Q \rangle_{H_0^1} \quad \forall Q \in H_0^1(\Omega),$$

where $(\cdot, \cdot)_{L^2}$ denotes the L^2 inner product.

The weak formulation of the Helmholtz equation is **elliptic**, hence we can make use of **elliptic regularity theory**.

The Helmholtz problem is **ill-posed** if k^2 is an eigenvalue of the Laplacian, in the sense that the solution P is not uniquely determined by the data F.

Finite element discretization

The **finite element discretization** of the Helmholtz equation is to find $u_h \in X_h$ such that

$$(\nabla P_h, \nabla Q_h)_{L^2} - k^2 (P_h, Q_h)_{L^2} = (F, Q_h)_{L^2} \quad \forall Q_h \in X_h,$$

where $X_h \subset H_0^1(\Omega)$ is a finite-dimensional subspace.

Finite element discretization

The **finite element discretization** of the Helmholtz equation is to find $u_h \in X_h$ such that

$$(\nabla P_h, \nabla Q_h)_{L^2} - k^2(P_h, Q_h)_{L^2} = (F, Q_h)_{L^2} \quad \forall Q_h \in X_h,$$

where $X_h \subset H_0^1(\Omega)$ is a finite-dimensional subspace.

- We discretize the domain Ω into a mesh of elements $\{\mathcal{T}_h\}_{h>0}$.

Finite element discretization

The **finite element discretization** of the Helmholtz equation is to find $u_h \in X_h$ such that

$$(\nabla P_h, \nabla Q_h)_{L^2} - k^2 (P_h, Q_h)_{L^2} = (F, Q_h)_{L^2} \quad \forall Q_h \in X_h,$$

where $X_h \subset H_0^1(\Omega)$ is a finite-dimensional subspace.

- We discretize the domain Ω into a mesh of elements $\{\mathcal{T}_h\}_{h>0}$.
- We choose as X_h the space of piecewise linear polynomial functions on the mesh, i.e.

$$X_h := \{v \in L^2(\Omega) : v|_\tau \in \mathcal{P}^1(\tau)\ \forall \tau \in T_h\} \cap H_0^1(\Omega) \subset H_0^1(\Omega).$$

Finite element discretization

The **finite element discretization** of the Helmholtz equation is to find $u_h \in X_h$ such that

$$(\nabla P_h, \nabla Q_h)_{L^2} - k^2 (P_h, Q_h)_{L^2} = (F, Q_h)_{L^2} \quad \forall Q_h \in X_h,$$

where $X_h \subset H_0^1(\Omega)$ is a finite-dimensional subspace.

- We discretize the domain Ω into a mesh of elements $\{\mathcal{T}_h\}_{h>0}$.
- We choose as X_h the space of piecewise linear polynomial functions on the mesh, i.e.

$$X_h := \{v \in L^2(\Omega) : v|_\tau \in \mathcal{P}^1(\tau) \; \forall \tau \in T_h\} \cap H_0^1(\Omega) \subset H_0^1(\Omega).$$

- We obtain a **linear system** of the form $\underline{\underline{\mathcal{A}}}\,\underline{P}_h = \underline{F}$.

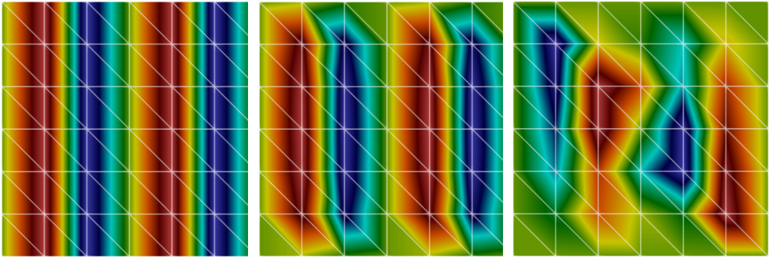
A motivating example

$$P(\cdot) = \sin(\pi\omega\cdot), \qquad \omega = 16$$

A motivating example

UNIVERSITY OF OXFORD
Mathematical Institute

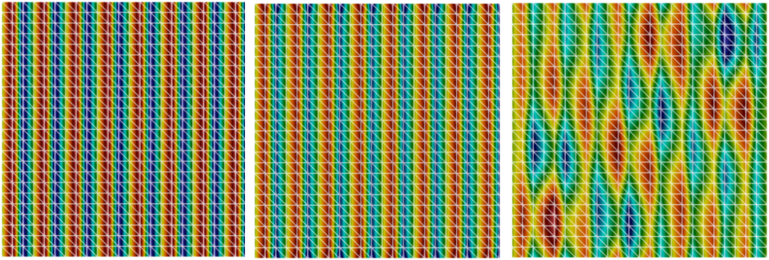
$$P(\cdot) = \sin(\pi\omega\cdot), \qquad \omega = 16$$



L^2 projection $\qquad k = 0 \qquad k = 256$

A motivating example

$$P(\cdot) = \sin(\pi\omega\cdot), \qquad \omega = 16$$

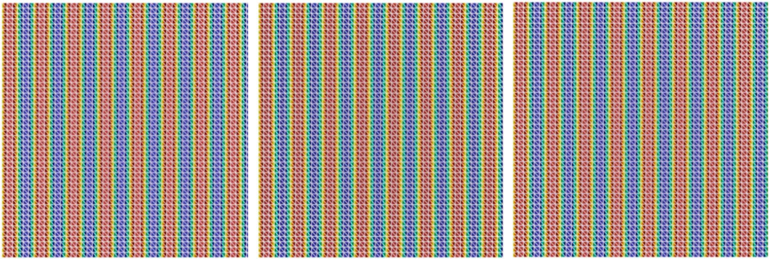


L^2 projection $\qquad k = 0 \qquad k = 256$

A motivating example

UNIVERSITY OF OXFORD
Mathematical Institute

$$P(\cdot) = \sin(\pi\omega\cdot), \qquad \omega = 16$$



L^2 projection $\qquad$ $k = 0$ $\qquad$ $k = 256$

Coercive elliptic problems

University of Oxford
Mathematical Institute

Definition (Coercivity)

We call a sesquilinear form $\mathcal{A} : X \times X \to \mathbb{C}$ **coercive** on X if there exists a constant $\alpha > 0$ such that

$$\mathcal{A}(P, P) \geq \alpha \|P\|_X^2 \qquad \forall P \in X.$$

Coercive elliptic problems

Definition (Coercivity)

We call a sesquilinear form $\mathcal{A} : X \times X \to \mathbb{C}$ **coercive** on X if there exists a constant $\alpha > 0$ such that

$$\mathcal{A}(P, P) \geq \alpha \|P\|_X^2 \qquad \forall P \in X.$$

Theorem (Lax-Milgram)

If $\mathcal{A} : X \times X \to \mathbb{C}$ is coercive on X, then the variational problem

$$\text{find } P \in X \text{ such that } \mathcal{A}(P, Q) = {}_{X'}\langle f, Q \rangle_X \quad \forall Q \in X$$

is **well-posed** *for all $f \in X'$.*

Discrete coercive problems

- Via Lax-Milgram, we also know that the discrete problem is well-posed if the bilinear form is coercive.

Discrete coercive problems

- Via Lax-Milgram, we also know that the discrete problem is well-posed if the bilinear form is coercive.

Cea's lemma

Let $\mathcal{A} : X \times X \to \mathbb{C}$ be coercive on X and let $P \in X$ be the solution of the continuous variational problem. Then the solution $P_h \in X_h$ of the discrete variational problem satisfies

$$\|P - P_h\|_X \leq \frac{M}{\alpha} \inf_{Q_h \in X_h} \|P - Q_h\|_X,$$

where M is the continuity constant of $\mathcal{A}$.

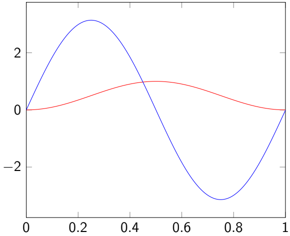
Lack of coercivity of the Helmholtz problem

UNIVERSITY OF OXFORD
Mathematical Institute

Consider the function $P(x) = \sin(\pi\omega x)$.

$$\|P\|^2_{L^2([0,1])} = \int_0^1 \sin^2(\pi\omega x) = \frac{1}{2}$$

$$\|P'\|^2_{L^2([0,1])} = \int_0^1 2\pi \sin(\pi\omega x)\cos(\pi\omega x)$$



notice that the last integral vanishes for $\omega \in \mathbb{N}$, on the interval $[0,1]$.

T-coercive elliptic problems

Definition (T-coercivity)

We call a sesquilinear form $\mathcal{A}(\cdot,\cdot)$ **T-coercive** on X if there exists a bijective operator $T \in L(X)$ and a constant $\alpha > 0$ s.t.

$$\mathcal{A}(Tu, u) \geq \alpha \|u\|_X^2 \qquad \forall u \in X.$$

T-coercive elliptic problems

Definition (T-coercivity)

We call a sesquilinear form $\mathcal{A}(\cdot,\cdot)$ **T-coercive** on X if there exists a bijective operator $T \in L(X)$ and a constant $\alpha > 0$ s.t.

$$\mathcal{A}(Tu, u) \geq \alpha \|u\|_X^2 \qquad \forall u \in X.$$

Theorem (Ciarlet[1])

If $\mathcal{A}(\cdot,\cdot)$ is T-coercive on X, then the corresponding variational problem is well-posed.

1. see e.g., P. Ciarlet Jr., "T-coercivity: Application to the discretization of Helmholtz-like problems", 2012.

Discrete T-coercivity

- T-coercivity is a **sufficient and necessary** condition for well-posedness of the weak formulation.

Discrete T-coercivity

- T-coercivity is a **sufficient and necessary** condition for well-posedness of the weak formulation.
- T-coercivity is **not** automatically inherited onto the discrete level.

Discrete T-coercivity

- T-coercivity is a **sufficient and necessary** condition for well-posedness of the weak formulation.
- T-coercivity is **not** automatically inherited onto the discrete level.
- T-coercivity, at the discrete level, is a equivalent to **uniform inf-sup stability**, i.e.

$$\inf_{v_h \in X_h} \sup_{w_h \in X_h} \frac{|a_h(v_h, w_h)|}{\|v_h\|_X \|w_h\|_X} \geq \beta > 0,$$

which guarantees the **well-posedness** of the discrete problem.

Compact Eigenvalue problems

Considering the eigenvalue problem, find $(\lambda, P) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$(\nabla P, \nabla Q)_{L^2} = \lambda (P, Q)_{L^2} \quad \forall Q \in H_0^1(\Omega).$$

Compact Eigenvalue problems

Considering the eigenvalue problem, find $(\lambda, P) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$(\nabla P, \nabla Q)_{L^2} = \lambda (P, Q)_{L^2} \quad \forall Q \in H_0^1(\Omega).$$

We can rewrite this as an eigenvalue problem associated with an operator $\mathcal{S} : H_0^1(\Omega) \to H_0^1(\Omega)$, i.e. find $(\lambda, P) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$(\nabla \mathcal{S} f, \nabla Q) = \lambda (f, Q) \quad \forall Q \in H_0^1(\Omega).$$

Compact Eigenvalue problems

Considering the eigenvalue problem, find $(\lambda, P) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$(\nabla P, \nabla Q)_{L^2} = \lambda (P, Q)_{L^2} \quad \forall Q \in H_0^1(\Omega).$$

We can rewrite this as an eigenvalue problem associated with an operator $\mathcal{S} : H_0^1(\Omega) \to H_0^1(\Omega)$, i.e. find $(\lambda, P) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$(\nabla \mathcal{S} f, \nabla Q) = \lambda (f, Q) \quad \forall Q \in H_0^1(\Omega).$$

Elliptic regularity

Via elliptic regularity theory, we know that the operator $\mathcal{S}$ is compact. In fact, by Rellich-Kondrachov we know that $H^2(\Omega) \subset\subset H_0^1(\Omega)$.

Hilbert basis

- Let $(\lambda^{(i)}, \Phi^{(i)})_{i\in\mathbb{N}}$ be the eigenpairs of the operator $\mathcal{S}$.

Hilbert basis

- Let $(\lambda^{(i)}, \Phi^{(i)})_{i\in\mathbb{N}}$ be the eigenpairs of the operator $\mathcal{S}$.
- The eigenfunctions $(\Phi^{(i)})_{i\in\mathbb{N}}$ form a **Hilbert basis** of $H_0^1(\Omega)$.

Hilbert basis

- Let $(\lambda^{(i)}, \Phi^{(i)})_{i\in\mathbb{N}}$ be the eigenpairs of the operator $\mathcal{S}$.
- The eigenfunctions $(\Phi^{(i)})_{i\in\mathbb{N}}$ form a **Hilbert basis** of $H_0^1(\Omega)$.
- We expand any $\Phi^{(i)} \in H_0^1(\Omega)$ as $P = \sum_{i\in\mathbb{N}} \Phi^{(i)}(P, \Phi^{(i)})_{H^1}$.

Hilbert basis

- Let $(\lambda^{(i)}, \Phi^{(i)})_{i\in\mathbb{N}}$ be the eigenpairs of the operator $\mathcal{S}$.
- The eigenfunctions $(\Phi^{(i)})_{i\in\mathbb{N}}$ form a **Hilbert basis** of $H^1_0(\Omega)$.
- We expand any $\Phi^{(i)} \in H^1_0(\Omega)$ as $P = \sum_{i\in\mathbb{N}} \Phi^{(i)}(P, \Phi^{(i)})_{H^1}$.
- We will adopt the convection that the $\left\|\Phi^{(i)}\right\|_{H^1}$ is unitary.

Hilbert basis

- Let $(\lambda^{(i)}, \Phi^{(i)})_{i\in\mathbb{N}}$ be the eigenpairs of the operator $\mathcal{S}$.
- The eigenfunctions $(\Phi^{(i)})_{i\in\mathbb{N}}$ form a **Hilbert basis** of $H^1_0(\Omega)$.
- We expand any $\Phi^{(i)} \in H^1_0(\Omega)$ as $P = \sum_{i\in\mathbb{N}} \Phi^{(i)}(P, \Phi^{(i)})_{H^1}$.
- We will adopt the convection that the $\left\|\Phi^{(i)}\right\|_{H^1}$ is unitary.

Spectral decomposition

S is a compact operator on $H^1_0(\Omega)$, hence the eigenvalues $\lambda^{(i)}$ are **discrete** and **tend to zero**. Furthermore, the eigenfunctions $P^{(i)}$ form a **Hilbert basis** of $H^1_0(\Omega)$.

T-coercivity of the Helmholtz problem

- $\|\Phi^{(i)}\|_{H^1} = 1$ implies $\|\Phi^{(i)}\|_{L^2} = (1 + \lambda^{(i)})^{-1/2}$.

T-coercivity of the Helmholtz problem

- $\|\Phi^{(i)}\|_{H^1} = 1$ implies $\|\Phi^{(i)}\|_{L^2} = (1 + \lambda^{(i)})^{-1/2}$.

$$
\begin{aligned}
\mathcal{A}(P, P) &= (\nabla P, \nabla P)_{L^2} - k^2 (P, P)_{L^2(\Omega)} \\
&= \sum_{i \in \mathbb{N}} P^{(i)} (\nabla \Phi^{(i)}, \nabla P)_{L^2} - k^2 P^{(i)} (\Phi^{(i)}, P)_{L^2} \\
&= \sum_{i \in \mathbb{N}} \lambda^{(i)} P^{(i)} (\Phi^{(i)}, P)_{L^2} - k^2 P^{(i)} (\Phi^{(i)}, P)_{L^2} \\
&= \sum_{i \in \mathbb{N}} \lambda^{(i)} |P^{(i)}|^2 (\Phi^{(i)}, \Phi^{(i)})_{L^2} - k^2 |P^{(i)}|^2 (\Phi^{(i)}, \Phi^{(i)})_{L^2} \\
&= \sum_{i \in \mathbb{N}} \left(\frac{\lambda^{(i)} - k^2}{1 - \lambda^{(i)}} \right) |P^{(i)}|^2
\end{aligned}
$$

T-coercivity of the Helmholtz problem

Suppose $\exists i_*$ s.t. $\lambda_1 < \cdots < \lambda_{i_*} < k^2 < \lambda_{i_*+1} < \ldots$

T-coercivity of the Helmholtz problem

Suppose $\exists i_*$ s.t. $\lambda_1 < \cdots < \lambda_{i_*} < k^2 < \lambda_{i_*+1} < \ldots$

We then construct $W := \operatorname{span}_{0 \le i \le i_*}\{\Phi^{(i)}\}$ and set $T := \mathrm{Id}_X - 2P_W$, i.e.

$$T\Phi^{(i)} = \begin{cases} -\Phi^{(i)} & \text{if } i \le i_*, \\ +\Phi^{(i)} & \text{if } i > i_*. \end{cases}$$

T-coercivity of the Helmholtz problem

Suppose $\exists i_*$ s.t. $\lambda_1 < \cdots < \lambda_{i_*} < k^2 < \lambda_{i_*+1} < \ldots$

We then construct $W := \operatorname{span}_{0 \leq i \leq i_*}\{\Phi^{(i)}\}$ and set $T := \operatorname{Id}_X - 2P_W$, i.e.

$$T\Phi^{(i)} = \begin{cases} -\Phi^{(i)} & \text{if } i \leq i_*, \\ +\Phi^{(i)} & \text{if } i > i_*. \end{cases}$$

- T is bijective, since it is self-inverse, i.e. $T^2 = \operatorname{Id}_X$.

We notice that $\mathcal{A}(\cdot,\cdot)$ is T-coercive, provided k^2 not an eigenvalue $\lambda^{(i)}$, since

$$\mathcal{A}(P,TP) = \sum_{i\leq i_*}\left(\frac{k^2-\lambda^{(i)}}{1+\lambda^{(i)}}\right)(\Phi^{(i)})^2 + \sum_{i>i_*}\left(\frac{\lambda^{(i)}-k^2}{1+\lambda^{(i)}}\right)(\Phi^{(i)})^2$$
$$\geq \alpha \sum_{i\in\mathbb{N}} \lambda^{(i)}(\Phi^{(i)})^2 = \alpha\|P\|^2_{H^1},$$

where $\alpha = \min_{i\geq 0}\left\{\left|\frac{\lambda^{(i)}-k^2}{1+\lambda^{(i)}}\right|\right\} > 0$.

Weak T-coercivity

Definition (weak T-coercivity)

A linear operator $A \in L(X)$ is called **weakly T-coercive** if there exists a bijective operator $T \in L(X)$ and $K \in L(X)$ compact s.t. $T^*A + K$ is coercive.

Weak T-coercivity

University of Oxford
Mathematical Institute

Definition (weak T-coercivity)

A linear operator $A \in L(X)$ is called **weakly T-coercive** if there exists a bijective operator $T \in L(X)$ and $K \in L(X)$ compact s.t. $T^*A + K$ is coercive.

- A is T-coercive if T^*A is bijective.

Weak T-coercivity

Definition (weak T-coercivity)

A linear operator $A \in L(X)$ is called **weakly T-coercive** if there exists a bijective operator $T \in L(X)$ and $K \in L(X)$ compact s.t. $T^*A + K$ is coercive.

- A is T-coercive if T^*A is bijective.
- A is weakly T-coercive if $T^*A + K$, B bijective and K compact.

Weak T-coercivity

University of Oxford
Mathematical Institute

Definition (weak T-coercivity)

A linear operator $A \in L(X)$ is called **weakly T-coercive** if there exists a bijective operator $T \in L(X)$ and $K \in L(X)$ compact s.t. $T^*A + K$ is coercive.

- A is T-coercive if T^*A is bijective.
- A is weakly T-coercive if $T^*A + K$, B bijective and K compact.

Lemma

If A is weakly T-coercive and injective, then A is bijective.

Robin boundary conditions

UNIVERSITY OF OXFORD
Mathematical Institute

Consider the Helmholtz problem with Robin boundary conditions, i.e. find $u \in H^1(\Omega)$ such that $\mathcal{A}(P, Q) = {}_{H^{-1}}\langle f, Q \rangle_{H^1}$ $\forall Q \in X$, where

$$\mathcal{A}(P, Q) := \underbrace{(\nabla P, \nabla Q)_{L^2} - k(P, Q)_{L^2}}_{\mathcal{A}_0(P,Q)} - ik\langle P, Q \rangle_{L^2(\partial\Omega)} = (f, Q)_{L^2}.$$

Robin boundary conditions

Mathematical Institute

Consider the Helmholtz problem with Robin boundary conditions, i.e. find $u \in H^1(\Omega)$ such that $\mathcal{A}(P,Q) = {}_{H^{-1}}\langle f, Q\rangle_{H^1}\ \forall Q \in X$, where

$$\mathcal{A}(P,Q) := \underbrace{(\nabla P, \nabla Q)_{L^2} - k(P,Q)_{L^2}}_{\mathcal{A}_0(P,Q)} - ik\langle P, Q\rangle_{L^2(\partial\Omega)} = (f, Q)_{L^2}.$$

Trace theorem

On bounded Lipschitz domains, the trace operator $\gamma_0 : H^1(\Omega) \to L^2(\partial\Omega)$ is compact.

Robin boundary conditions

Consider the Helmholtz problem with Robin boundary conditions, i.e. find $u \in H^1(\Omega)$ such that $\mathcal{A}(P, Q) = {}_{H^{-1}}\langle f, Q\rangle_{H^1}\ \forall Q \in X$, where

$$\mathcal{A}(P,Q) := \underbrace{(\nabla P, \nabla Q)_{L^2} - k(P,Q)_{L^2}}_{\mathcal{A}_0(P,Q)} - ik\langle P, Q\rangle_{L^2(\partial\Omega)} = (f, Q)_{L^2}.$$

Trace theorem

On bounded Lipschitz domains, the trace operator $\gamma_0 : H^1(\Omega) \to L^2(\partial\Omega)$ is compact.

- The operator $\langle Ku, v\rangle_{H^1} := -ik\langle \gamma_0 u, \gamma_0 v\rangle_{L^2(\partial\Omega)}$ is compact.

Robin boundary conditions

UNIVERSITY OF OXFORD
Mathematical Institute

Consider the Helmholtz problem with Robin boundary conditions, i.e. find $u \in H^1(\Omega)$ such that $\mathcal{A}(P, Q) = {}_{H^{-1}}\langle f, Q\rangle_{H^1}\ \forall Q \in X$, where

$$\mathcal{A}(P, Q) := \underbrace{(\nabla P, \nabla Q)_{L^2} - k(P, Q)_{L^2}}_{\mathcal{A}_0(P,Q)} - ik\langle P, Q\rangle_{L^2(\partial\Omega)} = (f, Q)_{L^2}.$$

Trace theorem

On bounded Lipschitz domains, the trace operator $\gamma_0 : H^1(\Omega) \to L^2(\partial\Omega)$ is compact.

- The operator $\langle Ku, v\rangle_{H^1} := -ik\langle \gamma_0 u, \gamma_0 v\rangle_{L^2(\partial\Omega)}$ is compact.
- A is weakly T-coercive (injectivity can also be shown)

Schatz argument

We begin observing that the sesquilinear form $\mathcal{A}_0$ satisfies the *Gårding inequality*, i.e.

$$\mathcal{G}\|P-P_h\|^2_{H^1_k} \leq Re\left\{\mathcal{A}_0(P-P_h, P-Q_h)\right\} + k^2\|P-P_h\|^2_{L^2},$$

where we have introduced the norm $\|P\|^2_{H^1_k} := \|\nabla P\|^2_{L^2} + k^2\|P\|^2_{L^2}$.

Schatz argument

We begin observing that the sesquilinear form $\mathcal{A}_0$ satisfies the *Gårding inequality*, i.e.

$$\mathcal{G}\|P-P_h\|^2_{H^1_k} \leq Re\left\{\mathcal{A}_0(P-P_h, P-Q_h)\right\} + k^2\|P-P_h\|^2_{L^2},$$

where we have introduced the norm $\|P\|^2_{H^1_k} := \|\nabla P\|^2_{L^2} + k^2\|P\|^2_{L^2}$.

Aubin-Nitsche duality trick

The following bound holds,

$$\|P-P_h\|_{L^2} \leq M^2\psi(X_h)\|P-P_h\|_{H^1_k},$$

where $\psi(X_h) := \sup\limits_{g\in L^2(\Omega)} \inf\limits_{Q_h\in X_h} \dfrac{\|g-Q_h\|_{H^1_k}}{\|g\|_{L^2}}$.

Schatz argument

Combining the Gårding inequality with the Aubin-Nitsche duality trick, we obtain

$$\mathcal{G}\|P - P_h\|^2_{H^1_k} \leq M\|P - Q_h\|_{H^1_k}\|P - P_h\|_{H^1_k} + M^2k^2\psi(X_h)^2\|P - P_h\|^2_{H^1_k}.$$

Schatz argument

Combining the Gårding inequality with the Aubin-Nitsche duality trick, we obtain

$$\mathcal{G}\|P-P_h\|^2_{H^1_k} \leq M\|P-Q_h\|_{H^1_k}\|P-P_h\|_{H^1_k}+M^2k^2\psi(X_h)^2\|P-P_h\|^2_{H^1_k}.$$

Imposing the following condition on $\psi(X_h)$,

$$\psi(X_h) \leq \left(\frac{\mathcal{G}}{2k^2M^2}\right)^{\frac{1}{2}},$$

we obtain the following error bound:

$$\mathcal{G}\|P-P_h\|_{H^1_k} \leq 2M \inf_{Q_h\in X_h}\|P-Q_h\|_{H^1_k}.$$

Schatz argument

Using *Bramble-Hilbert lemma*, we can bound $\psi(X_h)$ as

$$\psi(X_h) \leq C_{\mathcal{I}} h \|Z\|_{H^2(\Omega)} \leq \left(\frac{\mathcal{G}}{2k^2M^2}\right)^{\frac{1}{2}}.$$

Schatz argument

Using *Bramble-Hilbert lemma*, we can bound $\psi(X_h)$ as

$$\psi(X_h) \leq C_{\mathcal{I}} h \|Z\|_{H^2(\Omega)} \leq \left(\frac{\mathcal{G}}{2k^2 M^2}\right)^{\frac{1}{2}}.$$

Frequency regularity estimates

The following regularity estimate holds $\|P\|_{H^2} \leq (1+kC_\Omega)\|g\|_{L^2}$, for $P \in H^2(\Omega)$ such that $-\Delta P = g$.

Schatz argument

Using *Bramble-Hilbert lemma*, we can bound $\psi(X_h)$ as

$$\psi(X_h) \leq C_{\mathcal{I}} h \|Z\|_{H^2(\Omega)} \leq \left(\frac{\mathcal{G}}{2k^2M^2}\right)^{\frac{1}{2}}.$$

Frequency regularity estimates

The following regularity estimate holds $\|P\|_{H^2} \leq (1+kC_\Omega)\|g\|_{L^2}$, for $P \in H^2(\Omega)$ such that $-\Delta P = g$.

Combining the previous estimates we obtain,

$$h \lesssim C_{\mathcal{I}} \left(\frac{\mathcal{G}}{2k^2M^2}\right)^{\frac{1}{2}} (1+kC_\Omega)^{-1} \sim k^{-2}.$$

Discrete T-coercivity

Definition

We call a family of sesquilinear forms $(\mathcal{A}_h)_{h>0}$ on X_h **uniformly** T$_h$-coercive, if there exists bijective operators $T_h : X_h \to X_h$ and $\alpha > 0$ independent of h s.t.

$$\mathcal{A}_h(P_h, T_h P_h) \geq \alpha \|P_h\|_X^2 \qquad \forall P_h \in X_h.$$

Discrete T-coercivity

Definition

We call a family of sesquilinear forms $(\mathcal{A}_h)_{h>0}$ on X_h **uniformly** T_h-coercive, if there exists bijective operators $T_h : X_h \to X_h$ and $\alpha > 0$ independent of h s.t.

$$\mathcal{A}_h(P_h, T_h P_h) \geq \alpha \|P_h\|_X^2 \qquad \forall P_h \in X_h.$$

- If $\mathcal{A}_h$ is uniformly T_h-coercive, then the discrete problem is stable and quasi-optimal $\|P - P_h\|_X \lesssim \inf_{P_h \in X_h} \|P - Q_h\|_X$.

Discrete T-coercivity

Definition

We call a family of sesquilinear forms $(\mathcal{A}_h)_{h>0}$ on X_h **uniformly** T_h-coercive, if there exists bijective operators $T_h : X_h \to X_h$ and $\alpha > 0$ independent of h s.t.

$$\mathcal{A}_h(P_h, T_h P_h) \geq \alpha \|P_h\|_X^2 \qquad \forall P_h \in X_h.$$

- If $\mathcal{A}_h$ is uniformly T_h-coercive, then the discrete problem is stable and quasi-optimal $\|P - P_h\|_X \lesssim \inf_{P_h \in X_h} \|P - Q_h\|_X$. **Provided** it is usually enough to show that

$$\lim_{h\to 0} \|T - T_h\|_X = 0,$$

but this is an asymptotic result not explicit about h.

Discrete weak T-coercivity

UNIVERSITY OF OXFORD
Mathematical Institute

Theorem

Let $A = B + K$, where B is bijective and K compact and suppose that $\ker(A) = \{0\}$. If there exists a family of bijective operators $T_h \in \mathcal{L}(X_h)$ s.t. B is **uniformly $\mathbf{T}_h$-coercive** *on X_h, then there exists $h_0 > 0$ s.t. A is* **uniformly $\mathbf{T}_h$-coercive** *on X_h for $h < h_0$.*

- If A is weakly coercive and injective, and $(T^*)^{-1}B$ is uniformly T_h-coercive, then the discrete problem is stable and quasi-optimal

$$\|P - P_h\|_X \lesssim \inf_{Q_h \in X_h} \|P - Q_h\|_X.$$

Discrete Helmholtz

Find $P_h \in X_h$ such that for any $P_h \in X_h$ the following holds,

$$\mathcal{A}(P_h, Q_h) := \underbrace{(\nabla P_h, \nabla Q_h)_{L^2} - k(P_h, Q_h)_{L^2}}_{=:\mathcal{A}_0(P_h, Q_h)} - ik\langle P_h, Q_h\rangle_{L^2(\partial\Omega)} = (f, Q_h)_{L^2}.$$

Discrete Helmholtz

Find $P_h \in X_h$ such that for any $P_h \in X_h$ the following holds,

$$\mathcal{A}(P_h, Q_h) := \underbrace{(\nabla P_h, \nabla Q_h)_{L^2} - k(P_h, Q_h)_{L^2}}_{=:\mathcal{A}_0(P_h, Q_h)} - ik\langle P_h, Q_h\rangle_{L^2(\partial\Omega)} = (f, Q_h)_{L^2}.$$

- Only have to show that $\mathcal{A}_0$ is uniformly T_h-coercive.

Define $T_h : X_h \to X_h$ through

$$T_h e_h^{(i)} := \begin{cases} -e_h^{(i)} & \text{if } i \leq i_*, \\ +e_h^{(i)} & \text{if } i > i_*. \end{cases}$$

Discrete T-coercivity of $\mathcal{A}_0$

Following the same steps as before, we can expand P_h in terms of the eigenfunctions of $\mathcal{S} : X_h \to X_h$, i.e. $P_h = \sum_{i\in\mathbb{N}} P_h^{(i)}\Phi_h^{(i)}$.

$$\mathcal{A}_0(P_h, T_h P_h) := \sum_{i\leq i_*}\left(\frac{k-\lambda_h^{(i)}}{1+\lambda_h^{(i)}}\right)(P_h^{(i)})^2 + \sum_{i>i_*}\left(\frac{\lambda_h^{(i)}-k}{1+\lambda_h^{(i)}}\right)(P_h^{(i)})^2.$$

Discrete T-coercivity of $\mathcal{A}_0$

UNIVERSITY OF OXFORD
Mathematical Institute

Following the same steps as before, we can expand P_h in terms of the eigenfunctions of $\mathcal{S} : X_h \to X_h$, i.e. $P_h = \sum_{i \in \mathbb{N}} P_h^{(i)} \Phi_h^{(i)}$.

$$\mathcal{A}_0(P_h, T_h P_h) := \sum_{i \leq i_*} \left(\frac{k - \lambda_h^{(i)}}{1 + \lambda_h^{(i)}} \right) (P_h^{(i)})^2 + \sum_{i > i_*} \left(\frac{\lambda_h^{(i)} - k}{1 + \lambda_h^{(i)}} \right) (P_h^{(i)})^2.$$

- $\mathcal{A}_0$ is uniformly T_h-coercive if and only if $\lambda_h^{(i_*)} < k^2$.

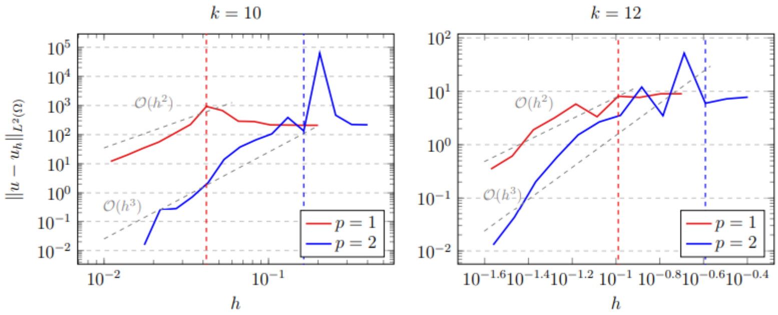
Discrete T-coercivity of $\mathcal{A}_0$

Following the same steps as before, we can expand P_h in terms of the eigenfunctions of $\mathcal{S} : X_h \to X_h$, i.e. $P_h = \sum_{i\in\mathbb{N}} P_h^{(i)} \Phi_h^{(i)}$.

$$\mathcal{A}_0(P_h, T_h P_h) := \sum_{i \leq i_*} \left(\frac{k - \lambda_h^{(i)}}{1 + \lambda_h^{(i)}} \right) (P_h^{(i)})^2 + \sum_{i > i_*} \left(\frac{\lambda_h^{(i)} - k}{1 + \lambda_h^{(i)}} \right) (P_h^{(i)})^2.$$

- $\mathcal{A}_0$ is uniformly T_h-coercive if and only if $\lambda_h^{(i_*)} < k^2$.
- This is equivalent to ensure that $\lambda_h^{(i_*)} - \lambda^{(i_*)} < k^2 - \lambda^{(i_*)}$

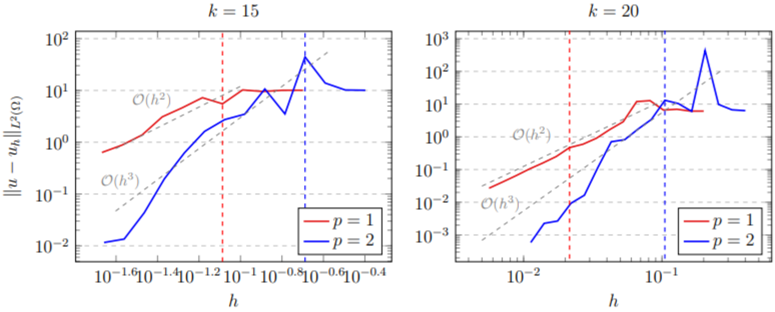
Discrete T-coercivity of $\mathcal{A}_0$



L^2-error of the approximation of the Helmholtz problem with Dirichlet boundary conditions against a computed reference solution.
The vertical lines indicate when $\lambda_h^{(i_*)} < k^2$.

Discrete T-coercivity of $\mathcal{A}_0$

UNIVERSITY OF OXFORD
Mathematical Institute



L^2-error of the approximation of the Helmholtz problem with Dirichlet boundary conditions against a computed reference solution.
The vertical lines indicate when $\lambda_h^{(i_*)} < k^2$.

Eigenvalue estimates

With classical eigenvalue estimates, we get

$$\lambda_h^{(i_*)} - \lambda^{(i_*)} \leq \lambda^{(i_*)} 4\sqrt{i_*} C_{\Omega,X} C_{\mathcal{I}} h^2,$$

where the constant $C_{\Omega,X}$ and $C_{\mathcal{I}}$ are defined as follows:

Eigenvalue estimates

With classical eigenvalue estimates, we get

$$\lambda_h^{(i_*)} - \lambda^{(i_*)} \leq \lambda^{(i_*)} 4\sqrt{i_*} C_{\Omega,X} C_{\mathcal{I}} h^2,$$

where the constant $C_{\Omega,X}$ and $C_{\mathcal{I}}$ are defined as follows:

- $C_{\Omega,X} = C_P/\kappa$ where C_P is the Poincaré constant and α is the coercivity constant of $-\Delta$.

Eigenvalue estimates

With classical eigenvalue estimates, we get

$$\lambda_h^{(i_*)} - \lambda^{(i_*)} \leq \lambda^{(i_*)} 4\sqrt{i_*} C_{\Omega,X} C_{\mathcal{I}} h^2,$$

where the constant $C_{\Omega,X}$ and $C_{\mathcal{I}}$ are defined as follows:

- $C_{\Omega,X} = C_P/\kappa$ where C_P is the Poincaré constant and α is the coercivity constant of $-\Delta$.
- $C_{\mathcal{I}}$ is the interpolation constant.

Eigenvalue estimates

With classical eigenvalue estimates, we get

$$\lambda_h^{(i_*)} - \lambda^{(i_*)} \leq \lambda^{(i_*)} 4\sqrt{i_*} C_{\Omega,X} C_{\mathcal{I}} h^2,$$

where the constant $C_{\Omega,X}$ and $C_{\mathcal{I}}$ are defined as follows:

- $C_{\Omega,X} = C_P/\kappa$ where C_P is the Poincaré constant and α is the coercivity constant of $-\Delta$.
- $C_{\mathcal{I}}$ is the interpolation constant.

So to have uniform T_h-coercivity, we want to ensure that

$$h^2 < \frac{k^2 - \lambda^{(i_*)}}{4\sqrt{i_*}\lambda^{(i_*)} C_{\Omega,X} C_{\mathcal{I}}}.$$

Quasi-optimality of $\mathcal{A}_0$

Theorem

The bilinear form $\mathcal{A}_0$ is uniformly T_h-coercive on X_h, if h is chosen such that

$$h^2 < \frac{k^2 - \lambda^{(i_*)}}{4\sqrt{i_*}\lambda^{(i_*)} C_{\Omega,X} C_{\mathcal{I}}}.$$

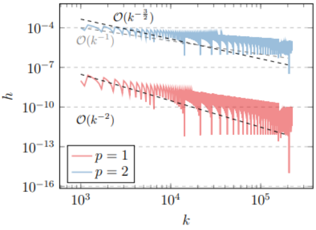
Quasi-optimality of $\mathcal{A}_0$

UNIVERSITY OF OXFORD
Mathematical Institute

Theorem

The bilinear form $\mathcal{A}_0$ is uniformly T_h-coercive on X_h, if h is chosen such that

$$h^2 < \frac{k^2 - \lambda^{(i_*)}}{4\sqrt{i_*}\lambda^{(i_*)} C_{\Omega,X} C_{\mathcal{I}}}.$$



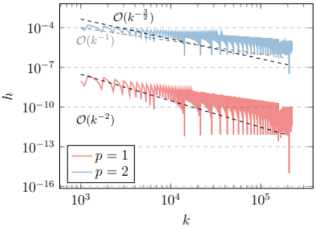
Quasi-optimality of $\mathcal{A}_0$

University of Oxford Mathematical Institute

Theorem

The bilinear form $\mathcal{A}_0$ is uniformly T_h-coercive on X_h, if h is chosen such that

$$h^2 < \frac{k^2 - \lambda^{(i_*)}}{4\sqrt{i_*}\lambda^{(i_*)} C_{\Omega,X} C_{\mathcal{I}}}.$$



- This guarantees that the discrete problem is quasi-optimal provided h is small enough.

Adaptive scheme

Construct the mesh, with the minimal number of elements, that guarantees the quasi-optimality of the Helmholtz problem:

Adaptive scheme

Construct the mesh, with the minimal number of elements, that guarantees the quasi-optimality of the Helmholtz problem:

- **Determine i_*:** either we know the eigenvalues, or we have to approximate them well enough (but we can choose any method we like to do this).

Adaptive scheme

Construct the mesh, with the minimal number of elements, that guarantees the quasi-optimality of the Helmholtz problem:

- **Determine i_*:** either we know the eigenvalues, or we have to approximate them well enough (but we can choose any method we like to do this).
- **Solving the Laplace eigenvalue problem adaptively:** Solve the Laplace eigenvalue problem on a sequence of refined meshes and check whether $k^2 - \lambda_h^{(i_*)} < 0$. If yes, we can stop because h is small enough s.t. we have uniform T_h-coercivity. (needs to use the same discretization as for Helmholtz)

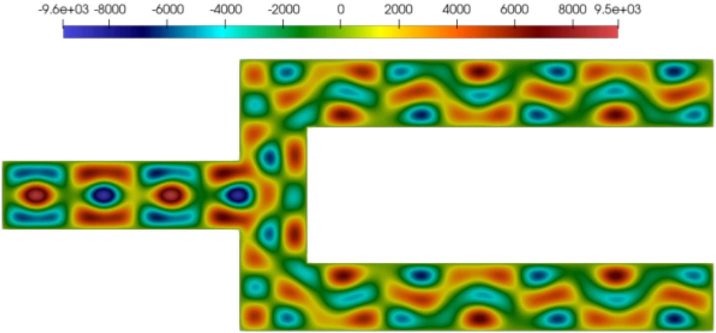
Adaptive scheme

UNIVERSITY OF OXFORD
Mathematical Institute

Construct the mesh, with the minimal number of elements, that guarantees the quasi-optimality of the Helmholtz problem:

- **Determine i_*:** either we know the eigenvalues, or we have to approximate them well enough (but we can choose any method we like to do this).
- **Solving the Laplace eigenvalue problem adaptively:** Solve the Laplace eigenvalue problem on a sequence of refined meshes and check whether $k^2 - \lambda_h^{(i_*)} < 0$. If yes, we can stop because h is small enough s.t. we have uniform T_h-coercivity. (needs to use the same discretization as for Helmholtz)
- **Solve the Helmholtz problem.**

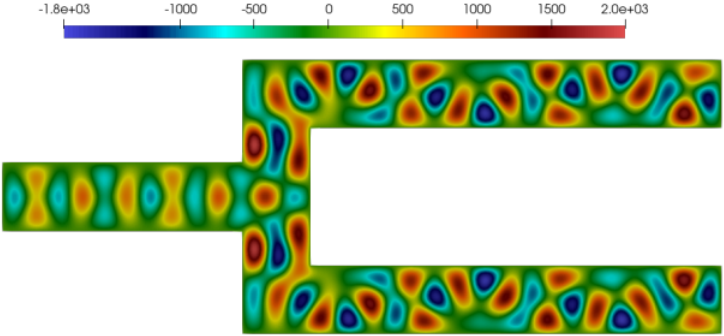
Adpative scheme: numerical examples



(a) $Number of DoFs: 21521.$

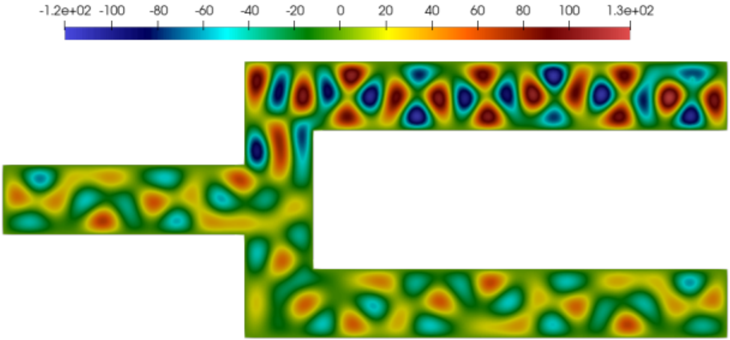
Adpative scheme: numerical examples

UNIVERSITY OF OXFORD
Mathematical Institute



(b) $Number of DoFs : 84769.$

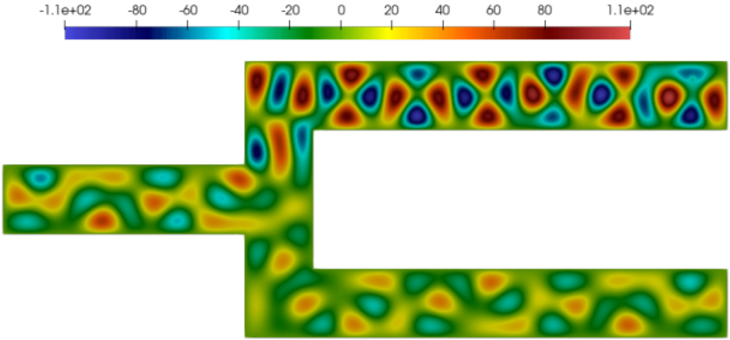
Adpative scheme: numerical examples



(c) $Number of DoFs : 336449.$

Adpative scheme: numerical examples

UNIVERSITY OF OXFORD
Mathematical Institute



(d) $Number of DoFs : 1340545.$

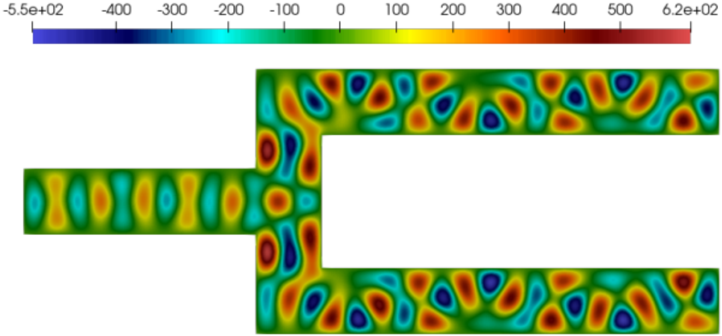
Adaptive^2 scheme

- In the adaptive scheme, we use a **Babuška–Rheinboldt** estimator to adaptively refine the mesh.
- The main idea is to use the **Babuška–Rheinboldt** estimator not one the desired solution or on a specific eigenfunction, but rather on the first $i_* + \ell$ eigenfunctions, i.e.

$$\eta_K = i_*^{-1} \sum_{i=1}^{i_*+\ell} h_K^2 \| \Delta e_h^{(i)} + \lambda_h^{(i)} e_h^{(i)} \|_{L^2(K)}^2 + \frac{h_K}{2} \| \nabla e_h^{(i)} \cdot n \|_{L^2(\partial K \setminus \partial \Omega)}^2 .$$

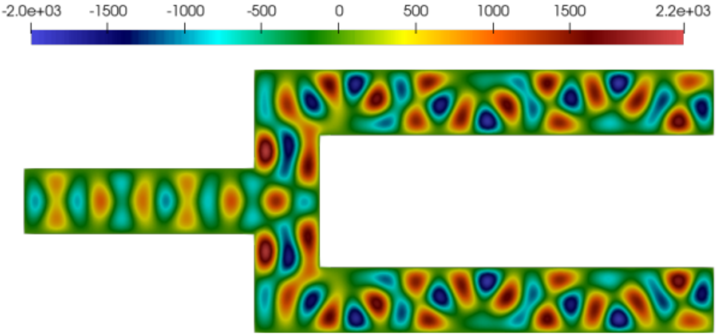
- We then refine the mesh in the elements where the indicator is larger. We can also adapt a **Dörfler marking** strategy.

Adpative^2 scheme: numerical examples



(a) $Number of DoFs : 77615.$

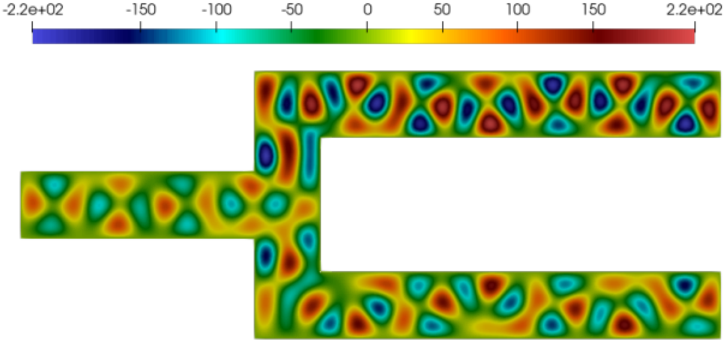
Adpative^2 scheme: numerical examples



(b) $Number of DoFs : 86733.$

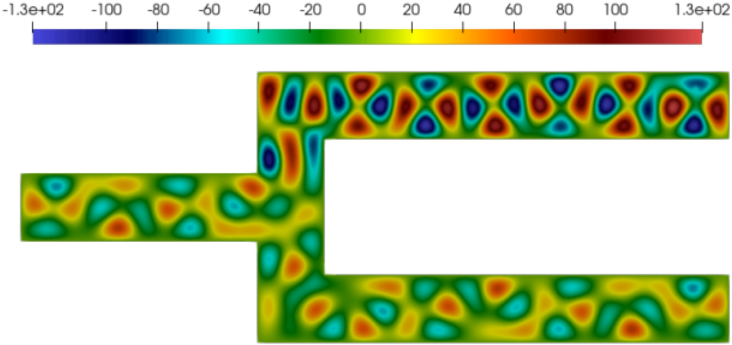
Adpative^2 scheme: numerical examples

UNIVERSITY OF OXFORD
Mathematical Institute



(c) $Number of DoFs$: 161102.

Adpative^2 scheme: numerical examples



(d) $Number of DoFs: 279034.$