The frequency decomposition multi-level method: a robust additive hierarchical basis preconditioner

Citation for published version (APA):

Document status and date:
Published: 01/01/1994

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
RANA 94-11
July 1994
The Frequency Decomposition
Multi-Level Method:
A robust additive hierarchical
basis preconditioner

by

R. Stevenson
The Frequency Decomposition Multi-Level Method: A robust additive hierarchical basis preconditioner

Rob Stevenson*

July, 1994

Abstract

Hackbusch's frequency decomposition multi-level method is characterized by the application of three additional coarse-grid corrections in parallel to the standard one. Each coarse-grid correction was designed to damp errors from a different part of the frequency spectrum. In this paper, we introduce a cheap variant of this method, partly based on semi-coarsening, which demands less recursive calls than the original version. Using the theory of the additive Schwarz methods, we will prove robustness of our method as a preconditioner applied to anisotropic equations.

Key words: Frequency decomposition, multi-level method, semi-coarsening, finite elements, hierarchical basis, additive Schwarz method, subspace decomposition, robustness.


1 Introduction

As is well-known, the rate of convergence of a multi-level method applied to a discretized elliptic boundary value problem is less than one uniformly in the toplevel. Yet, without a special choice of the components of the method, the rate of convergence tends to one as the problem becomes less elliptic (singularly perturbed problems), that is, the method is not robust. This paper concentrates on the question of robustness for so-called anisotropic problems. The classical way to obtain a robust multi-level method is to choose the smoother adapted to the problem. A disadvantage of this approach is that resulting smoothers are often expensive, not well parallelizable or, in three dimensions, hard to find.
An alternative approach is to add more coarse-grid corrections to the multi-level method. Representatives of this class of methods are Hackbusch’ Frequency Decomposition Multi-Level Method (FDMLM) ([2], [3], [5], [6]), that is subject of this paper, and the Multiple Semi-Coarsened Grids Method ([8], [9], [10]) introduced by Mulder.

In two dimensions, the FD Two-Level Method consists of four coarse-grid corrections, that can be performed in parallel, each of them designed to reduce errors in a (non-overlapping) part of the frequency spectrum. To speed up convergence, smoothers can be added to the algorithm but we shall not consider this option. In the (V-cycle) FDMLM, each of the four coarse-grid problems is solved by means of a recursive call, thus involving four coarse-grid corrections on the next coarser level. For a complete explanation of the ideas behind this method, we refer to the papers of Hackbusch. In [5], it has been proved that the FDTLM yields a robust preconditioner, that is, the condition number of the preconditioned system is bounded uniformly in the toplevel and the anisotropy. Up till now, robustness of the FDMLM is an open problem.

In this paper, we study a cheap variant of the FDMLM. As already was noted in [3], one of the coarse-grid problems generated by the FDTLM has a bounded condition number \( \lambda_{\text{max}} / \lambda_{\text{min}} \) uniformly with respect to the level and the anisotropy. Therefore, instead of applying a recursive call, this system can better be solved using a cheap iterative solver as e.g. Jacobi’s method. Apart from this, with our FDMLM, we solve two of the three remaining coarse-grid problems by means of only two instead of four coarse-grid corrections on the next coarser level by using semi-coarsening. It will appear then that also one of these two corrections yields a system with bounded condition number, which therefore can be solved cheaply. On the other system we apply the semi-coarsening idea recursively.

For any dimension \( d \), the complexity of the resulting algorithm is equivalent to the number of unknowns, even if one would apply more than one recursive calls on certain places in the algorithm. Considered as an additive Schwarz method or, in the terminology of [11], a Parallel Subspace Correction method, it consists of \( (\#\text{levels})^d \) subspace corrections compared to \( \sim (2^d)^\#\text{levels} \) subspace corrections for the FDMLM in its original form.

Using the theory of the additive Schwarz methods, we will prove robustness of our FDMLM as a preconditioner. To do that, we first reformulate the method in an abstract finite element context. This kind of formulation of a multi-level method was introduced in [1]. Then exploiting tensor products, the question of robustness will be reduced to the question of convergence of the method in one dimension applied to the identity and the Laplace operator.

In one dimension, the subspace decomposition that defines our method appears to be very similar to the decomposition of the finite element space into the differences of subsequent \( L^2 \)-orthogonal projections onto the finite element spaces corresponding to coarser grids. In particular, we will show that also our decomposition induces an \( L^2 \)-equivalent norm, which means convergence for the identity. The fact that the decomposition using \( L^2 \)-orthogonal projections yields an \( H^1 \)-equivalent norm plays a crucial role in the modern regularity free convergence proofs of standard multi-level methods (cf. [11], [12]). By adapting Xu’s proof of this result, we will prove the same for our decomposition and with that convergence for the Laplace operator.
Our FDMLM can be seen as block Jacobi's method after a basis transformation to a certain hierarchical basis. Our convergence result means that independent of the dimension, the stiffness matrix after this transformation has a bounded condition number uniformly in the level and the anisotropy.

In a forthcoming paper, we will report about numerical results obtained with the method.

Following [11], we shall use the notations $\lesssim, \gtrsim$ and $\asymp$. When we write

$$x_1 \lesssim y_1, x_2 \gtrsim y_2 \text{ and } x_3 \asymp y_3,$$

then there exists constants $C_1, C_2, C_3$ and $C_3$, that are independent of relevant parameters as the level or the anisotropy, such that

$$x_1 \leq C_1 y_1, x_2 \geq C_2 y_2 \text{ and } c_3 x_3 \leq y_3 \leq C_3 x_3.$$

## 2 Description of the method

### 2.1 Basic definitions

We start by giving some definitions for the one-dimensional case. Let $\Omega = (0,1)$, $h_k = 2^{-(k+1)}$ ($k \in \mathbb{N}_0 = \{0,1,2,\ldots\}$) and $\Omega^t_k = \Omega \cap h_k(Z + \frac{1}{2}i)$ ($t \in \{0,1\}$). Note that $\Omega^0_k = \Omega^0_{k-1} \cup \Omega^1_{k-1}$ (cf. figure 1). On the space of grid functions on $\Omega^t_k$, denoted by $\ell^2(\Omega^t_k)$, we define the scaled Euclidian scalar product by

$$\langle \mu, \nu \rangle_{\Omega^t_k} = h_k \sum_{x \in \Omega^t_k} \mu(x) \overline{\nu(x)}$$

and norm $\| \cdot \|_{\Omega^t_k} = \langle \cdot, \cdot \rangle_{\Omega^t_k}^{1/2}$.

The prolongations $p^t : \ell^2(\Omega^t_{k-1}) \rightarrow \ell^2(\Omega^t_k)$ are defined in difference stencil notation as $p^0 = \frac{1}{2} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$ (linear interpolation) and $p^1 = \frac{1}{2} \begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$. They satisfy

$$\text{range } p^0 \oplus \text{range } p^1 = \ell^2(\Omega^0_k). \quad (1)$$

The restrictions $r^t$ are defined as adjoints of the corresponding prolongations, that is, $r^0 = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$ and $r^1 = \frac{1}{4} \begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$. 

---

Figure 1: "Standard" and "shifted" coarse-grids $\Omega^0_{k-1}$ and $\Omega^1_{k-1}$ respectively.
For the general $d$-dimensional case, we define the grids $\Omega_k^i = \Omega_k^{i_1} \times \cdots \times \Omega_k^{i_d}$ ($k \in N_0^d$, $i \in \{0,1\}^d$). We equip the space $\ell^2(\Omega_k^i)$ of grid functions on $\Omega_k^i$ with scalar product $\langle \mu, \nu \rangle_{\Omega_k^i} = h_k^d \sum_{x \in \Omega_k^i} \mu(x) \nu(x)$.

Since we will exploit tensor products quite often, we note here that $\ell^2(\Omega_k^i) = \bigotimes_{j=1}^d \ell^2(\Omega_k^{i_j})$ i.e. $\ell^2(\Omega_k^i) = \text{span} \{ \otimes_{j=1}^d u_j : u_j \in \ell^2(\Omega_k^{i_j}) \}$, where $(\otimes_{j=1}^d u_j)(x) := \prod_{j=1}^d u_j(x_j)$. Furthermore, we have $\langle \otimes_{j=1}^d u_j, \otimes_{j=1}^d v_j \rangle_{\Omega_k^i} = \prod_{j=1}^d \langle u_j, v_j \rangle_{\Omega_k^{i_j}}$.

### 2.2 Derivation of the (modified) FDMLM

First, we consider the two-dimensional case. In [2], the FD Two-Level Method to solve a system $A \mu = \beta$ on $\Omega_{10}^{00}$ was defined by

$$
\mu \leftarrow \mu - \sum_{i,j \in \{0,1\}} p_x^i \otimes p_y^j \left( r_x^i \otimes r_y^j A p_x^i \otimes p_y^j \right)^{-1} r_x^i \otimes r_y^j (A \mu - \beta),
$$

where thus $r_x^i \otimes r_y^j A p_x^i \otimes p_y^j$ acts on the space of grid functions on $\Omega_{j-1}^{i-1}$ ($= \Omega_{j-1}^{i-1} \times \Omega_{j-1}^{i-1}$). [It will be clear why we avoid the term two-grid method.] Using the abbreviations $p^{ij}$ and $r^{ij}$ for $p_x^i \otimes p_y^j$ and $r_x^i \otimes r_y^j$ respectively, we have

$$
p^{ij} = \frac{1}{4} \begin{bmatrix}
(-1)^{i+j} & -1 & -1
\end{bmatrix}^{'}
$$

and

$$
r^{ij} = (p^{ij})^{'} = \frac{1}{16} \begin{bmatrix}
(-1)^{i+j} & -1 & -1
\end{bmatrix}^{'}
$$

We consider only $A > 0$. Then because of the Galerkin approach, the error amplification operator of the method is given by

$$
\mu^{*} - \mu^{\text{new}} = \sum_{i,j \in \{0,1\}} \mathcal{P}^{ij}(\mu^{*} - \mu^{\text{old}}),
$$

where $\mu^{*}$ is the exact solution and $\mathcal{P}^{ij}$ is the projection from $\ell^2(\Omega_{10}^{00})$ onto range $p^{ij}$ orthogonal with respect to $\langle A \cdot, \cdot \rangle_{\Omega_{10}^{00}}$. The range of the standard prolongation $p^{00}$ contains the "smooth" functions. The one-dimensional prolongation $p^1$ was chosen such that the ranges of the $p^{ij}$ for $i,j \neq (0,0)$ contain the different types of oscillating functions, so that also errors of that kind are corrected.

We consider systems that arise from the application of the bilinear finite element method on

$$
\begin{cases}
-(a_1 \partial_1^2 + a_2 \partial_2^2)u = f & \text{on } \Omega^2 \\
u = 0 & \text{on } \partial \Omega^2,
\end{cases}
$$
that is,
\[
A = \frac{1}{6}a_1 h_{J-1}^{-2} \begin{bmatrix} 1 & 2 & -1 \\ 4 & 1 \\ \end{bmatrix} + \frac{1}{6}a_2 h_{J-1}^{-2} \begin{bmatrix} -1 & -2 & 1 \\ -1 \\ \end{bmatrix} : \ell^2(\Omega_{JJ}^{00}) \rightarrow \ell^2(\Omega_{JJ}^{00})
\]

where \(a_1, a_2 \geq 0\) and \(a_1 + a_2 > 0\). This kind of problems is called \textit{anisotropic} if \(a_1 \ll a_2\) or \(a_1 \gg a_2\). In [5], it was proved that the condition number \(\kappa(\Sigma_{i,j \in \{0,1\}} p^\prime((r^\prime A p^\prime)^{-1} r^\prime A)) \lesssim 1\) (uniformly in \(J\) and \(a_i\)), which means that the FDTLM yields a \textit{robust} preconditioner for these systems. Our aim is to prove the same for a multi-level version.

In its original form, the multi-level version consists of recursive calls for each of the four coarse-grid problems on the grids \(\Omega_{J-1J-1}^{00}, \Omega_{J-1J-1}^{01}, \Omega_{J-1J-1}^{10}\) and \(\Omega_{J-1J-1}^{11}\). Since \(r^0[-1 2 -1] p^0 = \frac{1}{4}[-1 2 -1]\), \(r^0[1 4 1] p^0 = [1 4 1]\), \(r^1[1 4 1] p^1 = 2I\) and
\[
r^1 [-1 2 -1] p^1 = \frac{1}{4} \begin{bmatrix} 7 & 3 \\ 3 & 10 \\ \end{bmatrix} =: \frac{1}{4} \begin{bmatrix} 3 & 10 \\ 3 & 10 \\ \end{bmatrix},
\]
the operators on the spaces of grid functions on these grids are
\[
r^{00} A_p^{00} = \frac{1}{6} a_1 h_{J-1}^{-2} \begin{bmatrix} 1 & 2 & -1 \\ 4 & 1 \\ \end{bmatrix} + \frac{1}{6} a_2 h_{J-1}^{-2} \begin{bmatrix} -1 & -2 & 1 \\ -1 \\ \end{bmatrix} : \ell^2(\Omega_{J-1J-1}^{00}),
\]
\[
r^{01} A_p^{01} = \frac{1}{3} a_1 h_{J-1}^{-2} \begin{bmatrix} 1 & 2 & -1 \\ 4 & 1 \\ \end{bmatrix} + \frac{1}{3} a_2 h_{J-1}^{-2} \begin{bmatrix} -1 & -3 & 3 \\ -1 \\ \end{bmatrix} : \ell^2(\Omega_{J-1J-1}^{01}),
\]
\[
r^{10} A_p^{10} = \frac{1}{6} a_1 h_{J-1}^{-2} \begin{bmatrix} 3 & 10 & 3 \\ 4 & 1 \\ \end{bmatrix} + \frac{1}{3} a_2 h_{J-1}^{-2} \begin{bmatrix} -1 & -3 & 3 \\ -1 \\ \end{bmatrix} : \ell^2(\Omega_{J-1J-1}^{10}),
\]
and
\[
r^{11} A_p^{11} = \frac{1}{3} a_1 h_{J-1}^{-2} \begin{bmatrix} 3 & 10 & 3 \\ 4 & 1 \\ \end{bmatrix} + \frac{1}{3} a_2 h_{J-1}^{-2} \begin{bmatrix} 3 & 10 & 3 \\ 4 & 1 \\ \end{bmatrix} : \ell^2(\Omega_{J-1J-1}^{11}).
\]

As noted in [3], the condition number \(\kappa(r^{11} A_p^{11}) \lesssim 1\) (uniformly in \(h_{J-1} \) and \(a_i\)). So instead of applying a recursive call, the corresponding system can be solved using a cheap iterative solver. Furthermore, in [3] it was argued that in the cases \(a_1 \leq a_2\) or \(a_1 \geq a_2\) also one of the two operators \(r^{01} A_p^{01}\) and \(r^{10} A_p^{10}\) has a bounded condition number. Yet, this argument can not be applied to construct a method that is robust for the general variable coefficient case. Therefore, we will use another idea to further reduce the number of recursive calls.
Consider the following two operators that arise from \( A := r_0^1 A p_0^1 \) by means of semi-coarsening \( \Omega_{J-1,J-1}^{01} \) in the \( x \)-direction, that is, in the direction where we have not applied \( p^1 \) so far,

\[
r^0_x A p_x^0 = \frac{1}{3} a_1 h_{J-2}^2 \begin{bmatrix} -1 & 2 & -1 \\ \end{bmatrix} + \frac{1}{6} a_2 h_{J-1}^2 \begin{bmatrix} 3 \\ 10 \\ 3 \\ \end{bmatrix} \quad \text{on} \quad \ell^2(\Omega_{J-2,J-1}^{01})
\]

and

\[
r^1_x A p_x^1 = \frac{1}{3} a_1 h_{J-2}^2 \begin{bmatrix} 3 & 10 & 3 \\ \end{bmatrix} + \frac{1}{3} a_2 h_{J-1}^2 \begin{bmatrix} 3 \\ 10 \\ 3 \\ \end{bmatrix} \quad \text{on} \quad \ell^2(\Omega_{J-2,J-1}^{11}).
\]

Then the first operator is of the same type as \( A \) and so we can apply \((x\text{-})\)semi-coarsening recursively or, if \( J - 2 = 0 \), the operator has a bounded condition number and therefore the system can be solved using a cheap iterative solver. The second operator always has a bounded condition number. With a view to a recursive application, we note that this boundedness is not only uniform in \( h_{J-1} \) and \( a_i \), but also in the ratio \( h_{J-2}/h_{J-1} \). The argument is that for \( B_1, B_2 > 0 \), we have \( \kappa(B_1 + B_2) \leq \max\{\kappa(B_1), \kappa(B_2)\} \).

Analogously to the above procedure, we can solve the system on \( \Omega_{J-1,J-1}^{00} \) using semi-coarsening in the \( y \)-direction. Finally, as with the original version, the system on \( \Omega_{J-1,J-1}^{00} \) is solved with a recursive call of the entire method, with which this informal description of the modified FDMLM is completed (see figure 2).

Figure 2: Grids and prolongations defining the modified FDMLM in two-dimensions. [We dotted the lines at the bottom of the figure since the pictures of the grids correspond with \( J = 2 \), i.e. the three-level case.]
In view of the following, note that since e.g. on $\Omega_{J-1}^{0}$ no system is solved (unless $J-1 = 0$), but only coarse-grid corrections are invoked, this (intermediate) grid and the operators defined on it are only important for an efficient implementation. Because of the Galerkin approach, the mathematical properties of the resulting method are determined by the (sequence of) prolongations from the grids on which systems are (approximately) solved (leaves in the tree of figure 2) onto the finest grid $\Omega_{J}^{0}$. For example, for $J = 2$ these prolongations are $p_{x}^{0}p_{y}^{0} \otimes p_{y}^{0}p_{y}^{0}$, $p_{x}^{0}p_{x}^{0} \otimes p_{y}^{1}p_{y}^{0}$, $p_{x}^{1}p_{x}^{0} \otimes p_{y}^{1}p_{y}^{0}$, $p_{x}^{1}p_{x}^{0} \otimes p_{y}^{0}p_{y}^{1}$, $p_{x}^{0}p_{x}^{1} \otimes p_{y}^{1}p_{y}^{0}$, $p_{x}^{0}p_{x}^{1} \otimes p_{y}^{0}p_{y}^{1}$, $p_{x}^{1}p_{x}^{1} \otimes p_{y}^{1}p_{y}^{1}$, $p_{x}^{1} \otimes p_{y}^{1}$, $p_{x}^{1}$, $p_{y}^{1}$, i.e., tensor products of all possible combinations of $p_{x}^{0}$, $p_{x}^{1}$, $p_{y}^{0}$, $p_{y}^{1}$, and $p_{y}^{1}$.

We are now ready to give a formal description of the modified FDMLM.

**Algorithm 2.1** Let $A_{\mu} = \beta$ be a system on the $d$-dimensional grid $\Omega_{J}^{0}$. For $0 \leq k \leq J$, define the one-dimensional prolongation $p_{k} = p_{k}^{(J)}$ by

$$p_{k} = \left\{ \begin{array}{ll}
p_{0}^{0} \cdots p_{0}^{0} p_{1}^{1} : \ell^{2}(\Omega_{k-1}^{0}) \to \ell^{2}(\Omega_{J}^{0}) & k \geq 1 \\
p_{0}^{0} : \ell^{2}(\Omega_{k}^{0}) \to \ell^{2}(\Omega_{J}^{0}) & k = 0 \end{array} \right.$$

Note that because of (1), we have $\ell^{2}(\Omega_{J}^{0}) = \oplus_{k=1}^{J} \text{range} p_{k}$. For $k \in I := \{0, \ldots, J\}^{d}$, we define

$$p_{k} = \otimes_{j=1}^{d} p_{k_{j},x_{j}}, \quad r_{k} = p_{k}^{*}(= \otimes_{j=1}^{d} r_{k_{j},x_{j}}) \quad \text{and} \quad A_{k} = r_{k} A p_{k}.$$ 

Let now $B_{k}$ be such that $B_{k}^{-1}$ is a cheap approximation of $A_{k}^{-1}$. Then the (modified) FDMLM is defined by

$$\mu \leftarrow \mu - \sum_{k \in I} p_{k} B_{k}^{-1} r_{k} (A_{\mu} - \beta). \quad (2)$$

Our FDMLM is an example of an additive Schwarz method or Parallel Subspace Correction method with subspaces range $p_{k}$ satisfying $\ell^{2}(\Omega_{J}^{0}) = \oplus_{k \in I} \text{range} p_{k}$.

**Remark 2.2** Since e.g. $\Omega_{J-1}^{0}$ ($d = 2$) is coarsened only in the $x$-direction, the elementary one-dimensional prolongations $p_{0}$ and $p_{1}$, which are the building blocks of all prolongations in the algorithm, always map onto the space of grid functions on a non-shifted grid, that is, a grid $\Omega_{k}^{0}$ for some $k \in \mathbf{N}$. So in contrast to the original FDMLM, we do not have to construct boundary adaptations for $p_{0}$ and $p_{1}$ in order to maintain property (1).

We want to prove robustness of this method applied to anisotropic problems. As a consequence of the following lemma it is then sufficient to analyse the FDMLM with exact subspace corrections ($B_{k} = A_{k}$). The straightforward proof of this lemma is left to the reader.

**Lemma 2.3** Let $A > 0$ and $B_{k} > 0$ ($k \in I$). Define $\Lambda = \max_{k \in I} \lambda_{\text{max}}(B_{k}^{-1} A_{k})$ and $\lambda = \min_{k \in I} \lambda_{\text{min}}(B_{k}^{-1} A_{k})$. Then

$$\kappa \left( \sum_{k \in I} p_{k} B_{k}^{-1} r_{k} A \right) \leq \frac{\Lambda}{\lambda} \kappa \left( \sum_{k \in I} p_{k} A_{k}^{-1} r_{k} A \right).$$
Analogously to the two-dimensional case, for $d$-dimensional anisotropic problems we have that $\kappa(A_k) \leq 1$ ($k \in I$). So already the simple Richardson iteration, that is $B_k = \rho(A_k)$, gives $\frac{1}{\kappa} = \max_{k \in I} \kappa(A_k) \leq 1$.

### 2.3 Computational complexity

For ease of presentation we consider the two-dimensional case. The general case can be handled using induction.

Assume that the application of $B_k^{-1}$ ($k \in I$) costs a number of operations that is equivalent to the number of points of the grid in question. For $k_1 \leq k_2$, let $W_{k_1k_2}^0$ ($W_{k_2k_1}^0$) be the number of arithmetic operations necessary to treat a system on $\Omega_{k_1k_2}^0$ ($\Omega_{k_2k_1}^0$) using the recursive application of semi-coarsening in the $x$- ($y$-)direction. Then we have $W_{k_1k_2}^0 \approx \#\Omega_{k_1k_2}^0 + W_{k_1k_2}^0$, which gives $W_{k_1k_2}^0 \approx \#\Omega_{k_1k_2}^0$ and analogously $W_{k_2k_1}^0 \approx \#\Omega_{k_2k_1}^0$. Finally, let $W_{kk}^0$ be the number of arithmetic operations necessary for an entire FDMLM call on $\Omega_{kk}^0$. We conclude that

\[
W_{kk}^0 \approx W_{k-1k-1}^0 + W_{k-1k-1}^0 + \#\Omega_{kk}^0
\]

\[
\approx W_{k-1k-1}^0 + \#\Omega_{kk}^0,
\]

which implies $W_{kk}^0 \approx \#\Omega_{kk}^0$. Note that since $\#\Omega_{k-1k-1}^0 / \#\Omega_{kk}^0 = \frac{1}{4}$, more than one recursive calls on $\Omega_{k-1k-1}^0$ can be applied. The number of recursive calls involving semi-coarsening should be restricted to one.

### 3 Proof of robustness of the FDMLM

#### 3.1 Coordinate free finite element formulation

To facilitate the analysis, we reformulate the algorithm in a more abstract context. We start by giving some definitions for the one-dimensional case.

For $i \in \{0,1\}$, $k \in \mathbb{N}_0$, define $\delta_{k,x} \in \ell^2(\Omega_k^i)$ by $\delta_{k,x}(y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$.

Define $P_k : \ell^2(\Omega_k^i) \to H_0^1(\Omega) \subset L^2(\Omega)$ as the linear interpolation operator using zero boundary values. Put $M_k = \text{range} P_k$, that is, $M_k$ is the linear finite element space corresponding to the grid $\Omega_k^0$. The basis $\{\phi_{k,x}^0 := P_k \delta_{k,x}^0 : x \in \Omega_k^0\}$ is the standard (nodal) basis of $M_k$. We equip $M_k$ with scalar product

$\langle \cdot, \cdot \rangle_{M_k} = \langle P_k^{-1} \cdot, P_k^{-1} \cdot \rangle_{\Omega_k^0}$

and norm $\| \cdot \|_{M_k} = \langle \cdot, \cdot \rangle_{M_k}^{1/2}$. It is well-known that $\| \cdot \|_{M_k} \approx \| \cdot \|_{L^2}$ (uniformly in $k$).

For $k > 1$, we define $V_k = \text{range}(P_k p_1 : \ell^2(\Omega_{k-1}^0) \to H_0^1(\Omega))$. We will call the basis $\{\phi_{k,x}^1 := P_k p_1 \delta_{k-1,x}^1 : x \in \Omega_{k-1}^0\}$ the standard basis of $V_k$. Using $P_k p_1 = P_{k-1}$, we find that (1) is equivalent to

$M_k = M_{k-1} + M_k \ V_k$.  

(4)
So, with the definition \( V_0 = \mathcal{M}_0 \), the union of the bases of \( V_0, \ldots, V_k \) forms a basis of \( \mathcal{M}_k \) which is therefore called a \textit{hierarchical basis} (cf. figure 3). Note that (4) does not imply that \( V_0, \ldots, V_k \) are mutually orthogonal with respect to some scalar product.

For \( k \in \mathbb{N}_0 \), let \( I_k : V_k \to L^2(\Omega) \) be the inclusion operator. Since

\[
P_J p_k = \begin{cases} 
 P_J p_1^0 \cdots p_{J-k}^0 p_{J-k+1}^1 = P_k p_1^1 & 1 \leq k \leq J \\
 P_J p_0^0 \cdots p_{J-k}^0 = P_0 & k = 0
\end{cases}
\]

we find that \( p_k \) is the representation of \( I_k \) with respect to the standard bases on \( V_k \) and \( \mathcal{M}_J \).

As usual, for some basis \( \{\psi_i : i \in I_W\} \) of a subspace \( W \subset L^2(\Omega) \), we define the dual basis \( \{\tilde{\psi}_i : i \in I_W\} \) of \( W \) by \( (\tilde{\psi}_i, \psi_j)_{L^2} = \begin{cases} 
 1 & i = j \\
 0 & i \neq j
\end{cases} \). Let \( V_k : L^2(\Omega) \to V_k \) be the adjoint of \( I_k \) with respect to the \( L^2 \)-scalar product on both spaces, that is, \( V_k \) is the \( L^2 \)-orthogonal projection onto \( V_k \). Then we find that the representation of \( V_k|_{\mathcal{M}_J} \) with respect to the dual bases on \( \mathcal{M}_J \) and \( V_k \) is equal to the matrix adjoint of \( p_k \), that is, \( 2r_k \).

In the multi-dimensional case everything is defined using tensor products, that is,

\[
\mathcal{M}_k = \otimes_{j=1}^d \mathcal{M}_{kj}, \quad V_k = \otimes_{j=1}^d V_{kj}, \quad I_k = \otimes_{j=1}^d I_{kj} \quad \text{and} \quad V_k(= I_k^*) = \otimes_{j=1}^d V_{kj}.
\]

So, \( I_k : V_k \to L^2(\Omega^d) \) is the inclusion operator and \( V_k : L^2(\Omega^d) \to V_k \) is the \( L^2 \)-orthogonal projection onto \( V_k \). We equip \( \mathcal{M}_k \) and \( V_k \) with standard bases that are obtained by making tensor products of the standard basis functions of its factors, that is, the standard bases consist of functions of the form \( \phi_{kj,x} = \otimes_{j=1}^d \phi_{kj,x}^j \). Concerning dual basis functions, note that \( \tilde{\phi}_{kj,x} = \otimes_{j=1}^d \tilde{\phi}_{kj,x}^j \). Using the abbreviation \( \mathbf{m} \) for multi-indices \( (m, \ldots, m) \in \mathbb{N}_0^d \), we
conclude that for \( k \in I \), \( p_k \) is the representation of \( I_k \) with respect to the standard bases on \( V_k \) and \( M_J \) and \( 2^d r_k \) is the representation of \( V_k | M_J \) with respect to the dual bases.

Finally, for a given system \( A \mu = \beta \) on \( \Omega_0^J \), let us define \( A : M_J \rightarrow M_J \) by
\[
(A \phi_{J,y}^0, \phi_{J,x}^0)_{L^2} = A_{xy} \quad (x,y \in \Omega_0^J).
\]
(5)

Then \( A \) is the representation of \( A \) with respect to the standard basis on its domain and dual basis on its image. With the definition \( A_k = V_k AI_k \), we arrive at the conclusion that (2), with \( B_k = A_k \), is a matrix formulation of the iteration
\[
 u \leftarrow u - \sum_{k \in I} I_k A_k^{-1} V_k (Au - f),
\]
(6)

where \( u = \sum_{x \in \Omega_0^J} \mu(x) \phi_{J,x}^0, f = \sum_{x \in \Omega_0^J} \beta(x) \phi_{J,x}^0 \) (that is, \( \beta(x) = (f, \phi_{J,x}^0)_{L^2} \)). Because the condition number \( \kappa \) was defined as the quotient of the largest and smallest eigenvalue, clearly we have \( \kappa(\sum_{k \in I} p_k A_k^{-1} V_k A) = \kappa(\sum_{k \in I} I_k A_k^{-1} V_k A) \).

**Remark 3.1** We defined the operator \( A \) using the matrix \( A \). Of course, the usual procedure is the other way round. If \( a \) is a bilinear form on \( H_0^1(\Omega^d) \) and \( A : M_J \rightarrow M_J \) is defined by \((Au,v)_{L^2} = a(u,v) \) \((u,v \in M_J)\), then \( A \) defined by (5) is called the **stiffness matrix** with respect to the (multi-linear) basis \( \{ \phi_{J,x}^0 : x \in \Omega_0^J \} \).

**Remark 3.2** Consider the hierarchical basis of \( M_J = \oplus_{k \in I} V_k \) that is obtained by taking the union of the standard bases of the \( V_k \). The iteration (6) with respect to this basis, that is, the hierarchical basis for the solution and its dual for the right-hand side, is just block Jacobi's method with a partitioning into blocks corresponding to the spaces \( V_k \). As we have seen, for anisotropic problems, the diagonal blocks have bounded condition number and so robustness of (6) implies that, properly scaled, the stiffness matrix with respect to this hierarchical basis has a bounded condition number uniformly in the level and the anisotropy.

### 3.2 Main theorem; reduction to one dimensional cases

Since \( M_J = \oplus_{k \in I} V_k \), there exist unique projections \( Z_k = Z_k^{(J)} : M_J \rightarrow V_k \) such that \( \sum_{k \in I} Z_k = I \) on \( M_J \). Note that \( Z_k I_{k'} = 0 \) if \( k \neq k' \) and that \( Z_k I_k \) is the identity on \( V_k \).

**Lemma 3.3** Define \( W = \sum_{k \in I} Z_k^* A_k Z_k : M_J \rightarrow M_J \). Then \( W^{-1} \) exists and is equal to \( \sum_{k \in I} I_k A_k^{-1} V_k \).

**Proof** \( \sum_{k \in I} Z_k^* A_k Z_k \sum_{k' \in I} I_{k'} A_{k'}^{-1} V_{k'} = \sum_{k \in I} Z_k V_k = (\sum_{k \in I} I_k Z_k)^* = I. \)

A consequence of this lemma is that for \( A > 0 \), the condition number \( \kappa(\sum_{k \in I} I_k A_k^{-1} V_k A) \) is the quotient \( \Gamma \gamma \) of the optimum constants in the inequalities \( \gamma W \leq I \leq \Gamma W \) or, by \( A_k = V_k AI_k \),
\[
\gamma \sum_{k \in I} (AZ_k u, Z_k u)_{L^2} \leq (Au, u)_{L^2} \leq \Gamma \sum_{k \in I} (AZ_k u, Z_k u)_{L^2} \quad (u \in M_J).
\]
(7)

We are now ready to formulate our main theorem.
Theorem 3.4 For non-negative constants $a_j$ and $b$ with $\sum_{j=1}^{d} a_j + b > 0$, let
\[
a(u, v) = \sum_{j=1}^{d} \int_{\Omega^d} a_j \partial_j u \partial_j v + b \int_{\Omega^d} uv \quad (u, v \in H_0^1(\Omega^d))
\]
and let $A : \mathcal{M}_J \to \mathcal{M}_J$ be defined by $(Au, v)_{L^2} = a(u, v)$ $(u, v \in \mathcal{M}_J)$. Then we have
\[
\kappa(\sum_{k \in I} I_k A^{-1} V_k A) \lesssim 1 \quad \text{(uniformly in } J, a_j \text{ and } b).
\]
Remark 3.5 From (7), we immediately see that the theorem can be extended to all operators $A$ for which there exist $c, C > 0$ such that
\[
c(Au, u)_{L^2} \leq (Au, u)_{L^2} \leq C(Au, u)_{L^2}
\]
for some $A$ as described in the theorem. Examples are linear finite element discretizations or discretizations of elliptic boundary value problems with non-constant coefficients. With a view to the non-constant coefficient case, we note that clearly the theorem does not yield boundedness of the condition number that is uniform in $C/c$.

To prove theorem 3.4, we first note that when (7) is satisfied by $A^{(1)}$ and $A^{(2)}$, then it also satisfied by $c_1 A^{(1)} + c_2 A^{(2)}$ for any $c_1, c_2 \geq 0$. As a consequence, we only have to consider the quadratic forms $(Au, u)_{L^2} = \|u\|_{L^2}^2$ and $(Au, u)_{L^2} = \|\partial_i u\|_{L^2}^2$. By definition of a tensor product space, it is sufficient to check (7) for $u = \otimes_{j=1}^{d} u_j$, where $u_j \in \mathcal{M}_J$.

As a special case of the general definition, the one-dimensional $Z^{(J)}_k : \mathcal{M}_J \to \mathcal{V}_k$ was defined by $\sum_{k=0}^{J} Z^{(J)}_k = I$. Clearly, for $k \in I$, we have $Z_k = Z^{(J)}_k$. From \(\sum_{k \in I} \otimes_j u_j \|Z^{(J)}_k\|_{L^2(\Omega)^d} = \prod_j \|u_j\|_{L^2(\Omega)}\), \(\|\partial_j \otimes_j u_j \|_{L^2(\Omega)^d} = \|u_j\|_{L^2(\Omega)} \prod_{j \neq i} \|u_j\|_{L^2(\Omega)}\),
\[
\sum_{k \in I} \|Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)^d}^2 = \sum_{k \in I} \|Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)^d}^2 = \sum_{k \in I} \prod_j \|Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)}^2 = \prod_j \|Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)}^2
\]
and analogously
\[
\sum_{k \in I} \|\partial_j Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)^d}^2 = \sum_{k \in I} \|\partial_j Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)^d}^2 = \sum_{j \neq i} \|Z^{(J)}_k \otimes_j u_j \|_{L^2(\Omega)}^2
\]
we conclude that it suffices to prove the following norm equivalences in one dimension:
\[
\sum_{k=0}^{J} \|Z^{(J)}_k u\|_{L^2}^2 \simeq \|u\|_{L^2}^2 \quad (8)
\]
and
\[
\sum_{k=0}^{J} \|Z^{(J)}_k u\|_{H^1}^2 \simeq \|u\|_{H^1}^2 \quad (9)
\]
($u \in \mathcal{M}_J$), where \(\|u\|_{H^1} := (\cdot, \cdot)^{\frac{1}{2}}_{H^1}\) and $(u, v)_{H^1} := (u', v')_{L^2}$.

We start with constructing an explicit formula for the one-dimensional projections $Z^{(J)}_k$:
Lemma 3.6 Let $Y_k : M_{k+1} \to M_k$ be the projection on $M_k$ orthogonal with respect to $<,>_M$. Then

$$Z_k^{(J)} = \begin{cases} 
Y_0 \cdots Y_{J-1} & k = 0 \\
(I - Y_{k-1})Y_k \cdots Y_{J-1} & 1 \leq k \leq J
\end{cases} \quad (10)$$

Proof From $M_{J-1} \oplus ^{+}_{k=1} V_k = M_J$ and $M_{J-1} = \oplus ^{J-1}_{k=1} V_k$, it follows that $Z_k^{(J)} = (I - Y_{J-1})$ and $Z_k^{(J)} = Z_k^{(J-1)}Y_{J-1}$ ($0 \leq k \leq J - 1$).

Remark 3.7 As noted before, $\| \cdot \|_{M_k} \cong \| \cdot \|_{L^2}$ on $M_k$ (uniformly in $k$). If $\| \cdot \|_{M_k}$ would be equal to $\| \cdot \|_{L^2}$ (which is not the case), then $Z_k^{(J)}$ defined by (10) would be equal to $(Q_k - Q_{k-1})_{M_J}$, where $Q_k : L^2(\Omega) \to M_k$ is the $L^2$-orthogonal projection onto $M_k$ and $Q_{-1} := 0$. For the decomposition $u = \sum_{k=0}^{J} (Q_k - Q_{k-1})u$, (8) is trivially true. The corresponding relation (9) is famous and it is the key to the modern regularity free convergence proofs of standard multi-level methods (cf. [11, 12]). Unlike the decomposition $u = \sum_{k=0}^{J} (Q_k - Q_{k-1})u$, our decomposition was not introduced as a clever trick for the analysis of an overlapping subspace correction method, but it was yielded by the method itself.

In the next two subsections 3.3 and 3.4, we will prove the norm equivalences (8) and (9) respectively. We will prove (8) by estimating the angles between the spaces $V_k$ with respect to the $L^2$-scalar product for our one-dimensional regular grid case. For the same case, an alternative proof exploiting the standard bases of the $V_k$ will appear in [7]. We will prove (9) in an abstract framework using (8), $\sum_{k=1}^{J} \| (Q_k - Q_{k-1})u \|_{L^2}^2 \cong \| u \|_{L^2}^2$ ($u \in M_J$) and $\| \cdot \|_{M_k} \cong \| \cdot \|_{L^2}$ on $M_k$.

3.3 An $L^2$-equivalent norm on $M_J$

Since $Z_k^{(J)}$ is the projection from $M_J$ onto $V_k$ satisfying $\sum_{k=0}^{J} Z_k^{(J)} = I$, the validity of (8) depends on the angle between the spaces $V_k$.

Lemma 3.8 Let $\theta_{kl}$ be the smallest constant satisfying

$$| (u, v)_{L^2} | \leq \theta_{kl} \| u \|_{L^2} \| v \|_{L^2} \quad \text{ for all } u \in V_k, v \in V_l$$

(strengthened Cauchy-Schwarz inequality). Put $\Theta = (\theta_{kl})_{k,l}$ and assume $\rho(\Theta - I) < 1$. Then

$$1 - \rho(\Theta - I) \leq \| u \|_{L^2}^2 / \sum_{k=0}^{J} \| Z_k^{(J)}u \|_{L^2}^2 \leq 1 + \rho(\Theta - I) \quad (u \in M_J).$$

Proof Use

$$\| u \|_{L^2}^2 = \sum_{k=0}^{J} Z_k^{(J)}u, \sum_{l=0}^{J} Z_l^{(J)}u_{L^2} = \sum_{k=0}^{J} \| Z_k^{(J)}u \|_{L^2}^2 + \sum_{0 \leq k \neq l \leq J} (Z_k^{(J)}u, Z_l^{(J)}u)_{L^2}$$
and

\[ | \sum_{0 \leq k \neq l < J} (Z_k^{(j)} u, Z_l^{(j)} u)_{L^2} | \leq \sum_{0 \leq k \neq l < J} \theta_{kl} \| Z_k^{(j)} u \|_{L^2} \| Z_l^{(j)} u \|_{L^2} \leq \rho(\Theta - I) \sum_{k=0}^{J} \| Z_k^{(j)} u \|_{L^2}^2. \]

Using the fact that for \( k > l \), \( \mathcal{V}_k \) is orthogonal to \( \mathcal{M}_l \supset \mathcal{V}_l \) with respect to \( \langle \cdot, \cdot \rangle_{\mathcal{M}_k} \) (use (4)), we obtain the following estimate for \( \theta_{kl} \):

**Proposition 3.9** For \( k > l \),

\[ \theta_{kl} \leq \| M_k^\frac{k}{2} p^1 (r^1 M_k p^1)^{-1} r^1 (M_k - I) p_0 \cdots p_0 M_l^{-\frac{1}{2}} \|_{\Omega_k^0 - \Omega_l^0}, \]

where \( M_m := R_m P_m \) (mass-matrix) and \( R_m \) is the adjoint of \( P_m : \ell^2(\Omega_m^0) \rightarrow L^2(\Omega) \).

**Proof** Since \( P_k \) and \( p^1 \) are injective, \( M_k \) and \( r^1 M_k p^1 \) are invertible. The operators \( P_k p^1 (r^1 M_k p^1)^{-1} r^1 R_k \) and \( P_k M_k^{-1} R_k \) are the \( L^2 \)-orthogonal projections on \( \mathcal{V}_k (= \text{range} P_k p^1) \) and \( \mathcal{M}_k (= \text{range} P_k) \), in particular, these operators are the identity on \( \mathcal{V}_k \) and \( \mathcal{M}_k \) respectively. So for \( u \in \mathcal{V}_k \subset \mathcal{M}_k \) and \( v \in \mathcal{V}_l \), we have

\[ 0 = \langle u, v \rangle_{\mathcal{M}_k} = \langle P_k M_k^{-1} R_k u, v \rangle_{\mathcal{M}_k} = \langle M_k^{-1} R_k u, P_k^{-1} v \rangle_{\Omega_k^0} = (u, P_k M_k^{-1} P_k^{-1} v)_{L^2} \]

and thus

\[ (u, v)_{L^2} = (u, (I - P_k M_k^{-1} P_k^{-1}) v)_{L^2} = (P_k p^1 (r^1 M_k p^1)^{-1} r^1 R_k u, P_k (I - M_k^{-1}) P_k^{-1} v)_{L^2} \]

\[ = (u, P_k p^1 (r^1 M_k p^1)^{-1} r^1 (M_k - I) P_k^{-1} v)_{L^2}. \]

Application of the Cauchy-Schwarz inequality shows that

\[ \theta_{kl} \leq \| P_k p^1 (r^1 M_k p^1)^{-1} r^1 (M_k - I) P_k^{-1} \|_{\mathcal{M}_l \supset \mathcal{V}_l} \| L^2 \rightarrow L^2. \]

For \( u = P_m x \in \mathcal{M}_m \), we have \( \| u \|_{L^2}^2 = (P_m x, P_m x)_{L^2} = \| M_m^{\frac{1}{2}} x \|_{\Omega_m^0}^2 \) or \( \| u \|_{L^2} = \| M_m^{\frac{1}{2}} P_m^{-1} u \|_{\Omega_m^0}. \)

Using this we get

\[ \theta_{kl} \leq \| M_k^{\frac{1}{2}} p^1 (r^1 M_k p^1)^{-1} r^1 (M_k - I) P_k^{-1} P_l M_l^{-\frac{1}{2}} \|_{\Omega_k^0 - \Omega_l^0}. \]

Since \( P_k^{-1} P_l = \prod_{j=0}^{(k-l)} \), the proof is complete.

**Remark 3.10** Let \( k > l \). From the above proof it appears that the upper bound for \( \theta_{kl} \) from proposition 3.9 is equal to the smallest constant \( \theta'_{kl} \) satisfying

\[ \langle (u, v)_{L^2} \rangle \leq \theta'_{kl} \| u \|_{L^2} \| v \|_{L^2} \quad \text{for all } u \in \mathcal{V}_k, v \in \mathcal{M}_l. \]

Similarly, one could check that

\[ \theta_{kl} = \| M_k^{\frac{1}{2}} p^1 (r^1 M_k p^1)^{-1} r^1 (M_k - I) p_0 \cdots p_0 P_l (r^1 M_l p^1)^{-\frac{1}{2}} \|_{\Omega_k^0 - \Omega_{k-1}^0}. \]
We will now estimate the upper bound for $\theta_{kl}$ from proposition 3.9 in our one-dimensional regular grid case. Define $\overline{p} : \ell^2(\Omega_{m-1}) \to \ell^2(\Omega_m)$ by $(\overline{p}u)(x) = \begin{cases} 2u(x) & x \in \Omega_{m-1}^0 \\ 0 & x \in \Omega_m^0 \setminus \Omega_{m-1}^0 \end{cases}$ and $\overline{r} : \ell^2(\Omega_m^0) \to \ell^2(\Omega_{m-1}^0)$ by $(\overline{r}u)(x) = u(x)$. Then $\overline{r} = (\overline{p})^*$. It holds that

$$p^0 = \frac{1}{4}[1 \ 2 \ 1]p^0, \quad p^1 = \frac{1}{4}[-1 \ 2 \ -1]p^1, \quad r^0 = \overline{r} \frac{1}{4}[1 \ 2 \ 1] \quad \text{and} \quad r^1 = \overline{r} \frac{1}{4}[-1 \ 2 \ -1].$$

The mass-matrix $M_m$ is given by the difference stencil $\frac{1}{6}[1 \ 4 \ 1]$. It satisfies the relation

$$(M_m - I)p^0 = \frac{1}{4}p^0(M_m - I). \quad (11)$$

The set $\{\psi_{m}^{(i)} : x \mapsto \sqrt{2} \sin(\pi i x)\}_{i \in \{1, \ldots, n_m\}}$, where $n_m := h_m^{-1} - 1$, forms an orthonormal basis of $\ell^2(\Omega_m^0)$ consisting of eigenvectors $\psi_{m}^{(i)}$ of $M_m$, $\frac{1}{4}[1 \ 2 \ 1]$ and $\frac{1}{4}[-1 \ 2 \ -1]$ with eigenvalues $\frac{1}{3}(2 + \cos(\pi h_m))$, $\frac{1}{2}(1 + \cos(\pi h_m))$ and $\frac{1}{2}(1 - \cos(\pi h_m))$ respectively. It holds that

$$\overline{r}^1 \psi_{m}^{(i)} = \overline{r}^1 \psi_{m}^{(n_m+1-i)} = \psi_{m}^{(i)}|_{\Omega_{m-1}^0}, \quad \overline{p}^1(\psi_{m}^{(i)}|_{\Omega_{m-1}^0}) = \psi_{m}^{(i)} + \psi_{m}^{(n_m+1-i)},$$
and

$$\overline{p}^0 \psi_{m-1}^{(i)} = \psi_{m}^{(i)} - \psi_{m}^{(n_m+1-i)}$$

($\{i \in \{1, \ldots, n_m-1\}\}$).

After a basis transformation to the orthonormal bases, straightforward computations using the relations above now show that

$$\theta_{kk-1} \leq \|M_m^{\frac{1}{2}}p^1(r^1M_kp^1)^{-1}r^1p^0(M_{k-1} - I)M_k^{\frac{1}{2}}\|_{\alpha^0_{k-1} - \alpha^0_{k-1}}$$

$$= \max\left\{\left\|\frac{1}{4}\begin{bmatrix} (2 - x)^{\frac{1}{2}}(1 + x) \\ (2 + x)^{\frac{1}{2}}(1 - x) \end{bmatrix} \cdot x(x^2 - 1) \right\| : x = \cos(\pi h_k), i \in \{1, \ldots, n_{k-1}\}\right\}$$

$$\leq \max_{x \in [0, 1]} \frac{x(1 - x^2)}{2(1 + 2x^2)^{\frac{1}{2}}} =: \eta \approx 0.153.$$

and, by a repeated application of (11), that for $l < k - 1$

$$\theta_{kl} \leq \|M_k^{\frac{1}{2}}p^1(r^1M_kp^1)^{-1}r^1p^0\|_{\alpha^0_{k-1} - \alpha^0_{k-1}} \|M_l^{\frac{1}{2}} - I\| \|M_l^{\frac{1}{2}} - I\|_{\alpha^0_{k-1} - \alpha^0_{k-1}}$$

$$= \max\left\{\left\|\frac{1}{4}\sqrt{3}\begin{bmatrix} (2 - x)^{\frac{1}{2}}(1 + x) \\ (2 + x)^{\frac{1}{2}}(1 - x) \end{bmatrix} \right\| : x = \cos(\pi h_k), i \in \{1, \ldots, n_{k-1}\}\right\}$$

$$\leq \frac{1}{4}\sqrt{3} \cdot \frac{1}{4}\sqrt{3} \cdot \frac{1}{4}\sqrt{3} = \sqrt{2}(\frac{1}{4}\sqrt{2})^{k-l}. \quad (12)$$

By using lemma 3.8 and $\rho(\Theta - I) \leq \|\Theta - I\|_\infty \leq 2(\eta + \sqrt{2}\sum_{n \geq 2}(\frac{1}{4}\sqrt{2})^n) \approx 0.85 < 1$, we conclude that (8) is valid.
3.4 An $H^1$-equivalent norm on $\mathcal{M}_J$

A consequence of (8) is that

$$\|Z_l^{(k)}\|_{L^2-L^2} \lesssim 1 \quad (\text{uniformly in } k \geq l).$$  \hspace{1cm} (12)

Let $Q_k : L^2(\Omega) \to \mathcal{M}_k$ be the $L^2$-orthogonal projection on $\mathcal{M}_k$. Then it is well-known that

$$\|I - Q_k\|_{L^2-H^1} \lesssim h_k.$$  \hspace{1cm} (13)

From the inverse estimate

$$\|\cdot\|_{H^1} \lesssim h_k^{-1}\|\cdot\|_{L^2} \quad \text{on } \mathcal{M}_k,$$  \hspace{1cm} (14)

we also have

$$\|Q_k\|_{H^1-H^1} \gtrsim 1.$$  \hspace{1cm} (15)

In [11, Appendix], a compact proof is given of

$$\sum_{k=1}^{J} \|(Q_k - Q_{k-1})u\|_{H^1}^2 \approx \|u\|_{H^1}^2 \quad (u \in \mathcal{M}_J)$$  \hspace{1cm} (16)

(where $Q_{-1} := 0$). As we will see this proof with $Z_k^{(J)}$ playing the role of $Q_k - Q_{k-1}$ will also yield (9). A crucial property of $Z_k^{(J)}$ is stated in the following lemma.

**Lemma 3.11** For $1 \leq k \leq J$, $\|Z_k^{(J)}\|_{L^2-H^1} \lesssim h_{k-1}$.

**Proof** From lemma 3.6 we know that for $1 \leq k \leq J$, $Z_k^{(J)} = (I - Y_{k-1})Y_k \cdots Y_{J-1}$, where $Y_i : \mathcal{M}_{i+1} \to \mathcal{M}_i$ is the projection on $\mathcal{M}_i$ orthogonal with respect to $\langle \cdot, \cdot \rangle_{\mathcal{M}_{i+1}}$. From (13) it follows that for $u \in \mathcal{M}_k$,

$$\|(I - Y_{k-1})u\|_{L^2} \approx \|(I - Y_{k-1})u\|_{M_k} \leq \|(I - Q_{k-1})u\|_{M_k} \approx \|(I - Q_{k-1})u\|_{L^2} \lesssim h_{k-1}\|u\|_{H^1},$$

which gives the proof for the case $k = J$.

For $1 \leq k < J$, we write

$$Z_k^{(J)} = Z_k^{(J)}Q_J = (Z_k^{(J-1)}Q_{J-1} + Z_k^{(J-1)}(Y_{J-1} - Q_{J-1})Q_J$$

$$= Z_k^{(J-1)}Q_{J-1} + Z_k^{(J-1)}(Y_{J-1} - Q_{J-1})Q_J$$

$$\vdots$$

$$= Z_k^{(k)}Q_k + \sum_{l=k+1}^{J} Z_k^{(l-1)}(Y_{l-1} - Q_{l-1})Q_l.$$
From (17), (15), (12) and (13), we now get
\[
\| Z_k^{(J)} \|_{L^2 - H^1} \leq \| Z_k^{(k)} \|_{L^2 - H^1} \| Q_k \|_{H^1 - H^1} + \sum_{l=k+1}^J \| Z_k^{(l-1)} \|_{L^2 - L^2} \| Y_{l-1} - Q_{l-1} \|_{L^2 - H^1} \| Q_l \|_{H^1 - H^1} \\ \lesssim h_{k-1} + \sum_{l=k+1}^J h_{l-1} \approx h_{k-1}.
\]

**Remark 3.12** An obvious approach for estimating \( \| Z_k^{(J)} \|_{L^2 - H^1} \) for \( 1 \leq k \leq J \) would be to use
\[
\| Z_k^{(J)} \|_{L^2 - H^1} \leq \| I - Y_{k-1} \|_{L^2 - H^1} \| Y_k \cdots Y_{J-1} \|_{H^1 - H^1} \lesssim h_{k-1} \| Y_k \cdots Y_{J-1} \|_{H^1 - H^1}.
\]
Yet, the \( H^1 \)-stability \( \| Y_k \cdots Y_{J-1} \|_{H^1 - H^1} \lesssim 1 \) (uniformly in \( k < J \)) is not known a priori.

[Assuming lemma 3.11, it can be proved using \( Y_k \cdots Y_{J-1} = (I - Q_k - \sum_{l=k+1}^J Z_l^{(J)}) + Q_k. \)]

**Theorem 3.13** \( \sum_{k=0}^J \| Z_k^{(J)} u \|_{L^2}^2 \approx \| u \|_{H^1}^2 \) (\( u \in M_J \)).

**Proof** [Based on [11, Appendix].] Let \( u \in M_J \) and \( u_l = (Q_l - Q_{l-1})u \). Then \( u_l \in M_l \) and so for \( l < k \leq J \), \( Z_k^{(J)} u_l = 0 \). For \( k \leq l \), it follows from (14), (12) and (13) that
\[
\| Z_k^{(J)} u_l \|_{H^1} \lesssim h_k^{-1} \| Z_k^{(J)} u_l \|_{L^2} \lesssim h_k^{-1} \| u_l \|_{L^2} \lesssim h_k^{-1} h_l \| u_l \|_{H^1}.
\]

Let \( l \wedge m = \min\{l, m\} \). Writing \( u = \sum_{l=0}^J u_l \), we get
\[
\sum_{k=0}^J \| Z_k^{(J)} u \|_{H^1}^2 = \sum_{k=0}^J \sum_{l,m=0}^J (Z_k^{(J)} u_l, Z_k^{(J)} u_m)_{H^1} = \sum_{l,m=0}^J (Z_k^{(J)} u_l, Z_k^{(J)} u_m)_{H^1} \\ \leq \sum_{l,m=0}^J \sum_{k=0}^{l \wedge m} h_k^{-2} h_l h_m \| u_l \|_{H^1} \| u_m \|_{H^1} \lesssim \sum_{l,m=0}^J h_k^{l \wedge m} h_l h_m \| u_l \|_{H^1} \| u_m \|_{H^1} \\ \lesssim \sum_{l=0}^J \| u_l \|_{H^1}^2 \approx \| u \|_{H^1}^2
\]
by (16).

In [11, lemma 6.1], it was proved that for \( k \leq l \), \( u \in M_k \) and \( v \in M_l \),
\[
\| (u, v)_{H^1} \| \lesssim (h_k h_l)^{-\frac{1}{2}} \| u \|_{H^1} \| v \|_{L^2}.
\]
Using this and \( \| Z_m^{(J)} u \|_{L^2} = \| (Z_m^{(J)})^2 u \|_{L^2} \lesssim h_{m-1} \| Z_m^{(J)} u \|_{H^1} \) by lemma 3.11, we obtain for \( u \in M_J \),
\[
\| u \|_{H^1}^2 = \sum_{k,l=0}^J (Z_k^{(J)} u, Z_l^{(J)} u)_{H^1} \lesssim \sum_{k,l=0}^J (h_k h_l)^{-\frac{1}{2}} \min\{h_{k-1}, h_{l-1}\} \| Z_k^{(J)} u \|_{H^1} \| Z_l^{(J)} v \|_{H^1} \\ \lesssim \sum_{k=0}^J \| Z_k^{(J)} u \|_{H^1}^2,
\]
which completes the proof.
Acknowledgements

The author wishes to acknowledge the hospitality of the Institut für Informatik und Praktische Mathematik, Universität Kiel, where much of this work was performed. Furthermore, he would like to thank Professors W. Hackbusch (Kiel) and M. Dryja (Warsaw) and Dr. A. Reusken (Eindhoven) for fruitful discussions.

References


