Math 673

# Multigrid Methods: A Mostly Matrix-Based Approach

## Chapter 04: A Modern Energy-Norm Analysis of the Two-Grid Method

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

Fall 2024

# Chapter 04

# A Modern Energy-Norm Analysis of the Two-Grid Method

# Introduction

Like in the last chapter, we will consider the finite element approximation of the model problem in 1D:

$$\begin{cases} -u'' = f, & \text{in} \quad \Omega = (0,1), \\ u = 0, & \text{on} \quad \partial\Omega = \{0,1\}. \end{cases} \tag{1}$$

We will again use a uniform mesh and prove that the two-grid algorithm converges, borrowing many of the results from the previous chapter. But, the analysis style in this chapter will foreshadow the more general, and more modern, style used in the later chapters. In particular, the convergence theory will depend upon two properties, an *approximation property* and a *smoothing property*.

The reader will quickly realize how much simpler and more powerful is this modern energy approach to convergence. It is nothing short of remarkable.

# The Approximation Property

# Preliminaries

Recall we have defined the level-1 stiffness matrix, $\mathsf{A}_1 = [a_{1,i,j}] \in \mathbb{R}^{n_1 \times n_1}$, by the following equations:

$$a_{1,i,j} = \left(\psi_{1,j}', \psi_{1,i}'\right)_{L^2(0,1)} = \left(\psi_{1,i}', \psi_{1,j}'\right)_{L^2(0,1)}, \quad 1 \leq i,j \leq n_1.$$

This matrix is, recall, SPD, and is therefore orthogonally diagonalizable

$$\mathsf{A}_1 = \mathsf{V}_1 \mathsf{D}_1 \mathsf{V}_1^\top,$$

where $\mathsf{V}_1$ is the orthogonal matrix containing the normalized eigenvectors of $\mathsf{A}_1$,

$$\mathsf{V}_1 = \begin{bmatrix} | & | & & | \\ \tilde{\boldsymbol{v}}_1^{(1)} & \tilde{\boldsymbol{v}}_1^{(2)} & \cdots & \tilde{\boldsymbol{v}}_1^{(n_1)} \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}, \quad \tilde{\boldsymbol{v}}_1^{(k)} = \frac{\boldsymbol{v}_1^{(k)}}{\left\| \boldsymbol{v}_1^{(k)} \right\|_1}, \quad 1 \leq k \leq n_1,$$

and $\mathsf{D}_1$ is the diagonal matrix containing the eigenvalues of $\mathsf{A}_1$,

$$\mathsf{D}_1 = \operatorname{diag}\left(\lambda_1^{(1)}, \lambda_1^{(2)}, \cdots, \lambda_1^{(n_1)}\right) \in \mathbb{R}^{n_1 \times n_1}.$$

## Preliminaries

Recall the eigenvalues, $\lambda_1^{(k)}$, and eigenvectors, $\boldsymbol{v}_1^{(k)}$ of $\mathsf{A}_1$ are defined by

$$\lambda_1^{(k)} = \frac{4}{h_1}\sin^2\left(\frac{k\pi h_1}{2}\right) = \frac{2}{h_1}\left(1 - \cos(k\pi h_1)\right). \tag{2}$$

and

$$\left[\boldsymbol{v}_1^{(k)}\right]_i = v_{1,i}^{(k)} = \sin\left(k\pi x_{1,i}\right), \quad 1 \le i \le n_1, \tag{3}$$

respectively. Since $\mathsf{A}_1$ is SPD, we can define its square root as

$$\mathsf{A}_1^{1/2} = \mathsf{V}_1\mathsf{D}_1^{1/2}\mathsf{V}_1^\top,$$

where

$$\mathsf{D}_1^{1/2} = \mathrm{diag}\left(\sqrt{\lambda_1^{(1)}}, \sqrt{\lambda_1^{(2)}}, \cdots, \sqrt{\lambda_1^{(n_1)}}\right) \in \mathbb{R}^{n_1\times n_1}.$$

Of course, $\mathsf{A}_1^{1/2}$ is SPD, and $\mathsf{A}_1^{1/2}\mathsf{A}_1^{1/2} = \mathsf{A}_1$.

# Preliminaries

Now, recall, for any $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$,

$$\begin{aligned} \|\boldsymbol{v}_1\|_{A_1} &:= \sqrt{(\boldsymbol{v}_1, \boldsymbol{v}_1)_{A_1}} \\ &= \sqrt{(A_1\boldsymbol{v}_1, \boldsymbol{v}_1)_1} \\ &= \sqrt{\left(A_1^{1/2}\boldsymbol{v}_1, A_1^{1/2}\boldsymbol{v}_1\right)_1} \\ &= \left\|A_1^{1/2}\boldsymbol{v}_1\right\|_1 . \end{aligned}$$

In a similar way, we can define

$$\begin{aligned} \|\boldsymbol{v}_1\|_{A_1^2} &:= \sqrt{(A_1^2\boldsymbol{v}_1, \boldsymbol{v}_1)_1} \\ &= \sqrt{(A_1\boldsymbol{v}_1, A_1\boldsymbol{v}_1)_1} \\ &= \|A_1\boldsymbol{v}_1\|_1 . \end{aligned}$$

## Lemma

*Suppose that the two-level grid is uniform and nested and $\mathsf{A}_0$ satisfies the Galerkin condition. Then,*

$$\left\| (\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1 \right\|_1 = \sqrt{2\Lambda_1^{-1}} \left\| (\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1 \right\|_{\mathsf{A}_1}, \tag{4}$$

*for all $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$, where*

$$\Lambda_1 := \frac{4}{h_1}.$$

## Proof.

Expand $\boldsymbol{v}_1$ in the basis of eigenvectors of $\mathsf{A}_1$:

$$\boldsymbol{v}_1 = \sum_{k=1}^{n_1} v_{1,k} \boldsymbol{v}_1^{(k)}.$$

Recall that

$$(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1^{(k)} = \mathsf{S}_k \boldsymbol{v}_1^{(k)} + \mathsf{S}_k \boldsymbol{v}_1^{(k')}, \quad \text{(low frequency)}$$

for $1 \leq k \leq n_0 + 1$, and

$$(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1^{(k')} = \mathsf{C}_k \boldsymbol{v}_1^{(k)} + \mathsf{C}_k \boldsymbol{v}_1^{(k')}, \quad \text{(high frequency)}$$

for $1 \leq k \leq n_0 + 1$.

## Proof (Cont.)

Then

$$\begin{aligned}\left\|(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right\|_1^2 &= \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, (\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right)_1 \\ &= (n_1 + 1)\sum_{k=1}^{n_0+1} \delta_k \gamma_k^2,\end{aligned}$$

where

$$\delta_k = \begin{cases} \frac{1}{2}, & k = n_0 + 1, \\ 1, & \text{otherwise}, \end{cases}$$

and

$$\gamma_k := v_{1,k}\mathsf{S}_k + v_{1,k'}\mathsf{C}_k.$$

## Proof (Cont.)

Similarly,

$$\begin{aligned}\left\|(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right\|_{\mathsf{A}_1}^2 &= \left((\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1,(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right)_{\mathsf{A}_1}\\ &= \left(\mathsf{A}_1^{1/2}(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1,\mathsf{A}_1^{1/2}(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right)_1\\ &= (n_1+1)\sum_{k=1}^{n_0+1}\delta_k\gamma_k^2\left(\frac{\lambda_1^{(k)}+\lambda_1^{(k')}}{2}\right)\\ &= \frac{2}{h_1}(n_1+1)\sum_{k=1}^{n_0+1}\delta_k\gamma_k^2\\ &= \frac{1}{2}\Lambda_1\left\|(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right\|_1^2.\end{aligned}$$

$\square$

## Remark

*We will prove a similar result later in the general FEM framework, where, in particular, the meshes need not be uniform. In that setting, we will seek to prove that there is a constant $C_1 > 0$ such that*

$$\left\|(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right\|_1 \leq C_1\sqrt{\Lambda_1^{-1}}\left\|(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right\|_{\mathsf{A}_1}, \tag{5}$$

*for all $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$.*

## Theorem (The Approximation Property)

*Suppose that the two-level grid is uniform and nested and* $\mathsf{A}_0$ *satisfies the Galerkin condition. For any* $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$*,*

$$\left\| (\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1 \right\|_{\mathsf{A}_1} \leq C_1 \sqrt{\Lambda_1^{-1}} \left\| \boldsymbol{v}_1 \right\|_{\mathsf{A}_1^2}, \tag{6}$$

*where*

$$\Lambda_1 := \frac{4}{h_1} \quad \text{and} \quad C_1 = \sqrt{2}.$$

**Proof.**

Let $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$ be arbitrary. Then

$$\begin{aligned}
\left\|(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right\|_{\mathsf{A}_1}^2 &= \left(\mathsf{A}_1(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, (\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right)_1 \\
&= \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right)_1 \\
&= \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1\boldsymbol{v}_1\right)_1 - \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1\tilde{\Pi}_1\boldsymbol{v}_1\right)_1 .
\end{aligned}$$

## Proof (Cont.)

The second term on the RHS is zero, as we now show:

$$\begin{aligned}
\left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \tilde{\Pi}_1 \boldsymbol{v}_1\right)_{\mathsf{A}_1} &= \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1\tilde{\Pi}_1 \boldsymbol{v}_1\right)_1 \\
&= \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1(\mathsf{P}_0\mathsf{A}_0^{-1}\mathsf{R}_0 A_1)\boldsymbol{v}_1\right)_1 \\
&= \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, (\mathsf{A}_1\mathsf{P}_0\mathsf{A}_0^{-1}\mathsf{R}_0)\mathsf{A}_1 \boldsymbol{v}_1\right)_1 \\
&= \left((\mathsf{A}_1\mathsf{P}_0\mathsf{A}_0^{-1}\mathsf{R}_0)^\top(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1 \boldsymbol{v}_1\right)_1 \\
&= \left(\tilde{\Pi}_1(\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \boldsymbol{v}_1\right)_{\mathsf{A}_1} \\
&= 0,
\end{aligned}$$

since,

$$\tilde{\Pi}_1(\mathsf{I}_1 - \tilde{\Pi}_1) = \tilde{\Pi}_1 - \tilde{\Pi}_1^2 = \tilde{\Pi}_1 - \tilde{\Pi}_1 = 0.$$

## Proof (Cont.)

Therefore,

$$\begin{aligned}
\left\|(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right\|_{\mathsf{A}_1}^2 &= \left((\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1, \mathsf{A}_1\boldsymbol{v}_1\right)_1 \\
&\overset{\text{C.S.}}{\leq} \left\|(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right\|_1 \|\mathsf{A}_1\boldsymbol{v}_1\|_1 \\
&= \left\|(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right\|_1 \|\boldsymbol{v}_1\|_{\mathsf{A}_1^2} \\
&\overset{(4)}{\leq} \sqrt{2\Lambda_1^{-1}} \left\|(\mathsf{I}_1-\tilde{\Pi}_1)\boldsymbol{v}_1\right\|_{\mathsf{A}_1} \|\boldsymbol{v}_1\|_{\mathsf{A}_1^2} .
\end{aligned}$$

The result follows. □

## Corollary

*Suppose that the two-level grid is uniform and nested and* $\mathsf{A}_0$ *satisfies the Galerkin condition. Then*

$$\left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, (\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1\right)_{\mathsf{A}_1} = \left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \boldsymbol{v}_1\right)_{\mathsf{A}_1},$$

*or, equivalently ,*

$$\left((\mathsf{I}_1 - \tilde{\Pi}_1)\boldsymbol{v}_1, \tilde{\Pi}_1\boldsymbol{v}_1\right)_{\mathsf{A}_1} = 0.$$

## Remark

*Estimate* (6) *is called an* **approximation property** *with constant* $C_1 > 0$.

# Richardson's Smoother

# Richardson's Smoother

In this section, we want to change our smoothing algorithm from damped Jacobi to Richardson's method. It turns out that this change is mostly cosmetic, as the underlying structure for the smoother will be essentially the same. Specifically, our Richardson smoother will have the following error transfer matrix:

$$\mathsf{K}_1 = \mathsf{I}_1 - \frac{1}{\Lambda_1}\mathsf{A}_1, \quad \Lambda_1 = \frac{4}{h_1}.$$

Recall

$$\lambda_1^{(k)} = \frac{2}{h_1}\left(1 - \cos(k\pi h_1)\right).$$

So,

$$0 < \lambda_1^{(1)} < \lambda_1^{(2)} < \cdots < \lambda_1^{(n_1)} = \rho(\mathsf{A}_1) < \frac{4}{h_1} = \Lambda_1.$$

$\Lambda_1$ is almost the spectral radius of $\mathsf{A}_1$; the last estimate is asymptotically sharp.

# Relation to Damped Jacobi

For damped Jacobi, recall that

$$
\begin{aligned}
\mathsf{K}_1 = \mathsf{K}_1(\omega) &= \mathsf{I}_1 - \omega \mathsf{D}^{-1}\mathsf{A}_1 \\
&= \mathsf{I}_1 - \omega \frac{h_1}{2}\mathsf{A}_1.
\end{aligned}
$$

If we take $\omega = \frac{1}{2}$ in damped Jacobi, we get our new Richardson smoother.

## Richardson's Smoother

The error transfer matrix, which, for Richardson's method is

$$\mathsf{K}_1 = \mathsf{I}_1 - \Lambda_1^{-1}\mathsf{A}_1 = \mathsf{I}_1 - \frac{h_1}{4}\mathsf{A}_1,$$

has the following eigenvalues:

$$\begin{aligned}\mu_1^{(k)}\left(\frac{1}{2}\right) &= \frac{1}{2}\cos(k\pi h_1) + 1 - \frac{1}{2} \\ &= 1 - \frac{\lambda_1^{(k)}}{\Lambda_1}.\end{aligned}$$

Consulting the figure on the next page, it follows easily that

$$\left|\mu_1^{(k)}\left(\frac{1}{2}\right)\right| = \mu_1^{(k)}\left(\frac{1}{2}\right) \leq \mu_1^{(n_0+1)}\left(\frac{1}{2}\right) = \frac{1}{2}, \quad n_0 + 1 \leq k \leq n_1.$$

In other words the smoother reduces all high frequency modes of the error by at least half.

# Richardson's Smoother

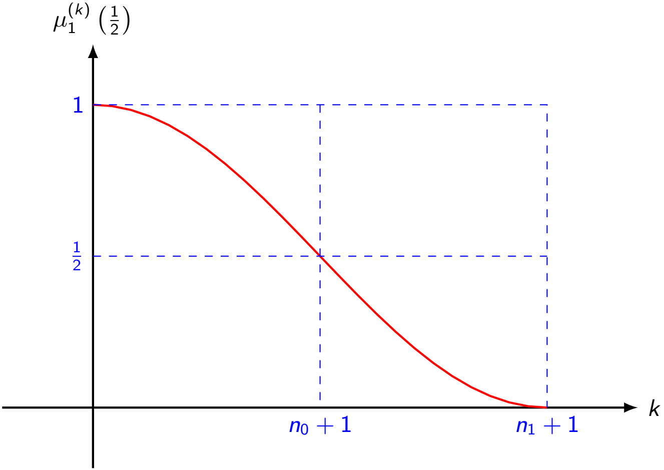


Figure: Eigenvalues of the error transfer matrix for the Richardson smoother.

# The Smoothing Property

Before we establish the so-called *smoothing property*, we need a technical lemma.

### Lemma

*For any* $m \in \mathbb{N}$,

$$\max_{0 \le x \le 1} x(1-x)^{2m} \le \frac{1}{2m}. \tag{7}$$

## Proof.

Define, for any $m \in \mathbb{N}$,

$$f(x) = x(1-x)^{2m}, \quad x \in [0,1].$$

Observe that

$$f(0) = f(1) = 0,$$

and, otherwise

$$f(x) > 0, \quad \forall\, x \in (0,1).$$

The derivative of $f$ is

$$f'(x) = (x-1)^{2m} + x(2m)(x-1)^{2m-1},$$

and there is a single zero for $f'(x)$ in $(0,1)$, which we label $x_0$. It is easy to show that

$$x_0 = \frac{1}{2m+1}.$$

## Proof (Cont.)

Then, the extreme value of $f$ satisfies

$$f(x_0) = \frac{1}{2m+1}\left(\frac{2m}{2m+1}\right)^{2m} \leq \frac{1}{2m+1} \leq \frac{1}{2m}.$$

The result follows. $\square$

## Theorem (The Smoothing Property)

*Suppose that the two-level grid is uniform and nested and $\mathsf{A}_0$ satisfies the Galerkin condition. Assume that smoothing is carried out by Richardson's method, or, equivalently, the damped Jacobi method with $\omega = \frac{1}{2}$. There is some constant $C_2 > 0$, such that*

$$\left\| \mathsf{K}_1^{m_1} \boldsymbol{v}_1 \right\|_{\mathsf{A}_1^2} \leq C_2 \sqrt{\frac{\Lambda_1}{m_1}} \left\| \boldsymbol{v}_1 \right\|_{\mathsf{A}_1}, \tag{8}$$

*for all $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$, for any $m_1 \in \mathbb{N}$. In particular, we can take*

$$C_2 = \sqrt{\frac{1}{2}}.$$

Proof.

First, observe that

$$\begin{aligned} \left\| \mathsf{K}_1^{m_1} \boldsymbol{v}_1 \right\|_{\mathsf{A}_1^2}^2 &= \left\| \mathsf{A}_1 \mathsf{K}_1^{m_1} \boldsymbol{v}_1 \right\|_1^2 \\ &= \left( \mathsf{A}_1 \mathsf{K}_1^{m_1} \boldsymbol{v}_1, \mathsf{A}_1 \mathsf{K}_1^{m_1} \boldsymbol{v}_1 \right)_1 . \end{aligned}$$

Using the eigenvector basis for $\mathsf{A}_1$, let us write

$$\boldsymbol{v}_1 = \sum_{k=1}^{n_1} v_{1,k} \boldsymbol{v}_1^{(k)}.$$

## Proof (Cont.)

Then

$$
\begin{aligned}
\|K_1^{m_1} \boldsymbol{v}_1\|_{A_1^2}^2 &= \frac{n_1+1}{2} \sum_{k=1}^{n_1} \left(\lambda_1^{(k)} v_{1,k}\right)^2 \left(\mu_1^{(k)}\left(\frac{1}{2}\right)\right)^{2m_1} \\
&= \Lambda_1 \left(\frac{n_1+1}{2}\right) \sum_{k=1}^{n_1} \left(\frac{\lambda_1^{(k)}}{\Lambda_1}\right)\left(1-\frac{\lambda_1^{(k)}}{\Lambda_1}\right)^{2m_1} \lambda_1^{(k)} v_{1,k}^2 \\
&\leq \Lambda_1 G(m_1) \left(\frac{n_1+1}{2}\right) \sum_{k=1}^{n_1} \lambda_1^{(k)} v_{1,k}^2 \\
&= \Lambda_1 G(m_1) \|\boldsymbol{v}_1\|_{A_1}^2 ,
\end{aligned}
$$

where

$$
G(m) := \max_{1 \leq k \leq n_1} \left(\frac{\lambda_1^{(k)}}{\Lambda_1}\right)\left(1-\frac{\lambda_1^{(k)}}{\Lambda_1}\right)^{2m} .
$$

## Proof (Cont.)

Upon rescaling and using the lemma we just proved, we have

$$G(m) \leq \max_{0 \leq x \leq 1} x(1-x)^{2m} \overset{(7)}{\leq} \frac{1}{2m}.$$

Therefore

$$\|\mathsf{K}_1^{m_1} \boldsymbol{v}_1\|_{\mathsf{A}_1^2} \leq \sqrt{\frac{1}{2}} \sqrt{\frac{\Lambda_1}{m_1}} \|\boldsymbol{v}_1\|_{\mathsf{A}_1}, \qquad (9)$$

and the result follows with $C_2 = \sqrt{\frac{1}{2}}$. □

### Remark

*Estimate* (8) *is called a* **smoothing property** *with constant* $C_2 > 0$.

# Convergence in the Energy Norm

Now that we have the powerful smoothing and approximation properties, the energy-norm convergence of the two-grid method is simple matter.

**Theorem (Convergence of the One-Sided Two-Grid Method)**

*Suppose that the two-level grid is uniform and nested and $\mathsf{A}_0$ satisfies the Galerkin condition. Suppose $m_2 = 0$, and with $\omega = \frac{1}{2}$ (Richardson). Then the two grid method converges provided $m_1$ is sufficiently large. Moreover, we have the error estimate*

$$\left\| \boldsymbol{e}_1^{\ell+1} \right\|_{\mathsf{A}_1} \leq C_1 C_2 m_1^{-1/2} \left\| \boldsymbol{e}_1^{\ell} \right\|_{\mathsf{A}_1},$$

*where $C_1, C_2 > 0$ are as given in the previous results. Since we can assume $C_1 C_2 = 1$, it follows that the method converges if $m_1 \geq 2$.*

## Proof.

Recall, the error transfer matrix in this case is

$$\mathsf{E}_1 = (\mathsf{I}_1 - \tilde{\Pi}_1)\mathsf{K}_1^{m_1}.$$

So

$$\begin{aligned}
\left\| \boldsymbol{e}_1^{\ell+1} \right\|_{\mathsf{A}_1} &= \left\| (\mathsf{I}_1 - \tilde{\Pi}_1)\mathsf{K}_1^{m_1} \boldsymbol{e}_1^{\ell} \right\|_{\mathsf{A}_1} \\
&\overset{(6)}{\leq} C_1 \sqrt{\Lambda_1^{-1}} \left\| \mathsf{K}_1^{m_1} \boldsymbol{e}_1^{\ell} \right\|_{\mathsf{A}_1^2} \\
&\overset{(8)}{\leq} C_1 \sqrt{\Lambda_1^{-1}} C_2 \sqrt{\Lambda_1} m_1^{-\frac{1}{2}} \left\| \boldsymbol{e}_1^{\ell} \right\|_{\mathsf{A}_1} \\
&= C_1 C_2 m_1^{-\frac{1}{2}} \left\| \boldsymbol{e}_1^{\ell} \right\|_{\mathsf{A}_1}.
\end{aligned}$$

□

## Remark

*In this simplified setting, we know that* $C_1 C_2 = 1$*. For the more general FEM case that we will explore in the future – for which the* $h_1$*-independent constants* $C_1 > 0$ *and* $C_2 > 0$*, may not be known explicitly – we will still be guaranteed that the method converges at a uniform,* $h_1$*-independent rate provide* $m_1 \geq 1$ *is large enough so that*

$$0 < C_1 C_2 m_1^{-\frac{1}{2}} < 1.$$