Multigrid Lecture Notes [*,†]

Wenqiang Feng [‡], Steven Wise [*]

Abstract

Multigrid (MG) methods in numerical analysis are a group of algorithms for solving differential equations using a hierarchy of discretizations. MG method is based on the two facts: High frequency will be damped by smoother and Low frequency can be approximated well by coarse grid, since the low-frequency errors on a fine mesh can become high-frequency errors on a coarser mesh. MG methods can be used as solvers as well as preconditioners.

[†]Key words: UTK, PDE, Multigrid.

[‡]Department of Mathematics,University of Tennessee, Knoxville, TN, 37909, wfeng@math.utk..edu

[*]Department of Mathematics,University of Tennessee, Knoxville, TN, 37909, swise@math.utk..edu

Contents

List of Figures 3

List of Tables 4

0 Preliminaries 5
- 0.1 Vector norms . . . 5
- 0.2 Matrix norms . . . 5
- 0.3 Eigenvalues . . . 7
  - 0.3.1 Eigenvalues of Tridiagonal Matrices . . . 7
  - 0.3.2 Eigenvalues of Hermitians Matrices . . . 8
- 0.4 Spectral radius . . . 8
- 0.5 Hermitian and Symmetric . . . 10

1 Classical Iterative Methods 13
- 1.1 General Iterative Methods . . . 13
- 1.2 Jacobi Iterative Method . . . 18
- 1.3 Gauss-Seidel Method . . . 20
- 1.4 Richardson Method . . . 23
- 1.5 Symmetrization . . . 25
  - 1.5.1 symmetrized multiplication . . . 25
  - 1.5.2 Symmetric Gauss-Seidel Method . . . 26
  - 1.5.3 Regular Splittings . . . 26

2 The Two-Grid Algorithm 31

3 Rigorous (Quantitative) Fourier Analysis of the Two-Grid Method 40
- 3.1 The Damped Jacobi Smoother . . . 45
- 3.2 A Basic Mesh Assumption . . . 47
- 3.3 Prolongation and Restriction Operators . . . 50
- 3.4 Restriction Operator . . . 54
- 3.5 Quantitative Error Analysis . . . 55
- 3.6 Some Technical Results . . . 55
- 3.7 Qualitative Two-Grid Method Convergence . . . 60
- 3.8 Smoothing Revisited . . . 62

4 Multigrid 65
- 4.1 Richardson's Smoother . . . 74
- 4.2 Convergence of the Two-level Method Revisited . . . 74
- 4.3 Convergence of the V-cycle Algorithm . . . 78
- 4.4 Convergence of the V-Cycle Algorithm . . . 82

5 Conforming Finite Element Method 87
- 5.1 The Stiffness Matrices . . . 91
- 5.2 Strong Approximation Property . . . 94
- 5.3 Full Multigrid for Finite Element Methods . . . 97
- 5.4 Multigrid and Subspace Corrections . . . 99
- 5.5 Properties of the "Projections" . . . 102

6 Subspace Decompositions 109

6.1 Hierarchical Basis . . . 113
6.2 The BPX Preconditioner . . . 124

References **129**

List of Figures

1 The initial decomposition of matrix A. . . . 18
2 The curve of $\rho(T_{RC})$ as a function of ω . . . 24
3 The action of Prolongation and Restriction Operator. . . . 31
4 Restriction from fine grid to coarse grid on 1D uniform mesh. . . . 31
5 Prolongation from coarse grid to fine grid on 1D uniform mesh. . . . 32
6 Mesh on the fine and coarse grid of the two-grid method in 2D. . . . 32
7 The initial error and the absolute value of initial error in Fourier space. . . . 34
8 The error after 1_{st} coarse grid correction and post-smoothing with the absolute value of error in Fourier space. . . . 34
9 The error after 2_{nd} coarse grid correction and post-smoothing with the absolute value of error in Fourier space. . . . 35
10 The error after 3_{rd} coarse grid correction and post-smoothing with the absolute value of error in Fourier space. . . . 35
11 The error after 4_{th} coarse grid correction and post-smoothing with the absolute value of error in Fourier space. . . . 36
12 Piecewise linear basis function in 1D. . . . 41
13 Piecewise linear basis function ϕ_i in 1D. . . . 42
14 The value of $1 - \cos(\pi x)$. . . . 44
15 The upper and lower bound of $1 - \cos(\pi x)$. . . . 45
16 Fine and coarse grid for $n_1 = 3$. . . . 47
17 The eigenvalue of $\mu_1^{(k)}(\omega) = 1 - 2\omega \sin^2\left(\frac{k\pi h}{2}\right)$. . . . 48
18 Coarse grid $n_0 = 3$ in 1D. . . . 50
19 Fine grid $n_1 = 7$ in 1D. . . . 51
20 One full W- and V- Cycle for three levels. Left One full W-Cycle, Right: One full V- Cycle . . . 67
21 One full W- and V- Cycle for six levels. Left One full W-Cycle, Right: One full V- Cycle . . . 67
22 Mesh on the fine and coarse grid of the multigrid method in 2D. . . . 67
23 Bisecting the triangulation $\mathcal{T}_0$ in 1D. . . . 87
24 Quadrisecting the triangulation $\mathcal{T}_0$ in 2D. . . . 87
25 Recursively Mesh in 2D. . . . 88
26 Example of $N_{\ell,j}$ and $\psi_{\ell,i}$ in 2D. . . . 89
27 Mesh on the fine and coarse grid of the multigrid method in 2D. . . . 90
28 One full cycle with $r = 2$. . . . 98

List of Tables

0 Preliminaries

Let $\mathcal{M}_{m\times n} := \mathcal{M}_{m\times n}(\mathbb{C})$ be the set of all matrices with m rows and n columns and $\mathcal{M}_{m\times n}(\mathbb{R})$ be the subset of $\mathcal{M}_{m\times n}$ composed of matrices with only real entries. Denote by $\mathcal{M}_n := \mathcal{M}_{n\times n}$ the set of all square matrices of size $n \times n$, and by $\mathcal{M}_n(\mathbb{R})$ the subset of $\mathcal{M}_n$ composed of matrices with only real entries.

0.1 Vector norms

Definition 0.1. *(Vector Norms) A vector norm is a function $\|\cdot\| : \mathbb{R}^n\ \mathbb{R}$ satisfying the following conditions for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$*

1. *nonnegative* : $\|x\| \geq 0, \quad (\|x\| = 0 \Leftrightarrow x = 0)$,
2. *homegenity* : $\|\alpha x\| = |\alpha|\,\|x\|$,
3. *triangle inequality* : $\left\|x + y\right\| \leq \|x\| + \left\|y\right\|, \quad \forall x, y \in \mathbb{R}^n$,

0.2 Matrix norms

Definition 0.2. *(Matrix Norms) A matrix norm is a function $\|\cdot\| : \mathbb{R}^{m\times n}\ \mathbb{R}$ satisfying the following conditions for all $A, B \in \mathbb{R}^{m\times n}$ and $\alpha \in \mathbb{R}$*

1. *nonnegative* : $\|A\| \geq 0, \quad (\|A\| = 0 \Leftrightarrow x = 0), \forall A \in \mathcal{M}_n$,
2. *homegenity* : $\|\lambda x\| = |\lambda|\,\|x\|, \quad \forall \lambda \in \mathbb{C}$ *and* $\forall A \in \mathcal{M}_n$
3. *triangle inequality* : $\|A + B\| \leq \|A\| + \|B\|, \quad \forall A, B \in \mathcal{M}_n$,
4. *submultiplicativity* : $\|AB\| \leq \|A\|\|B\|, \ \forall A, B \in \mathcal{M}_n$.

Definition 0.3. *For $A \in \mathbb{R}^{m\times n}$, some of the most frequently matrix vector norms are*

1. *F-norm* : $\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{i=1}^{n} |a_{ij}|^2}$,
2. *1-norm* : $\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$,
3. *∞-norm* : $\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$,
4. *induced-norm* : $\|A\|_p = \sup_{x\in\mathbb{R}^n, x\neq 0} \dfrac{\|Ax\|_p}{\|x\|_p}$.

Lemma 0.4. *Let $\|\cdot\|$ is a matrix norm on $\mathcal{M}_n$, then*

$$\left\|A^k\right\| \leq \|A\|^k$$

Proof.

$$\left\|A^k\right\| = \sup_{0\neq x\in\mathbb{C}^n} \frac{\left\|A^k x\right\|}{\|x\|} \leq \sup_{0\neq x\in\mathbb{C}^n} \frac{\|A\|\left\|A^{k-1}x\right\|}{\|x\|} \leq \cdots \leq \|A\|^k.$$

□

Theorem 0.5. *(Neumann Series) Suppose that $A \in \mathbb{R}^{n\times n}$. If $\|A\| < 1$, then $(I-A)$ is nonsingular and*

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k \tag{0.1}$$

with

$$\frac{1}{1+\|A\|} \le \left\|(I-A)^{-1}\right\| \le \frac{1}{1-\|A\|}. \tag{0.2}$$

Moreover, if A is nonnegative, then $(I-A)^{-1} = \sum_{k=0}^{\infty} A^k$ is also nonnegative.

Proof.

1. (I-A) is nonsingular, i.e. $(I-A)^{-1}$ exits.

$$\begin{aligned}\left\|(I-A)x\right\| &\ge \|Ix\| - \|Ax\| \\ &\ge \|x\| - \|A\|\,\|x\| \\ &= (1-\|A\|)\,\|x\| \\ &= C\,\|x\|.\end{aligned}$$

   So, we get if $(I-A)x = 0$, then $x = 0$. Therefore, $ker(I-A) = 0$, then $(I-A)^{-1}$ exists.

2. Let $S_N = \sum_{k=0}^{N} A^k$, we want to show $(I-A)S_N \to I$, as $N \to \infty$.

$$(I-A)S_N = S_N - AS_N = \sum_{k=0}^{N} A^k - \sum_{k=1}^{N+1} A^k = A^0 - A^{N+1} = I - A^{N+1}.$$

   So by Lemma 0.4

$$\left\|(I-A)S_N - I\right\| = \left\|-A^{N+1}\right\| \le \|A\|^{N+1}.$$

   Since $\|A\| < 1$, then $\|A\|^{N+1} \to 0$. Therefore,

$$(I-A)\sum_{k=0}^{\infty} A^k = I.$$

   and

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k$$

3. bounded norm
   Since

$$1 = \|I\| = \left\|(I-A) * (I-A)^{-1}\right\|.$$

   So,

$$(1-\|A\|)\left\|(I-A)^{-1}\right\| \le 1 \le (1+\|A\|)\left\|(I-A)^{-1}\right\|.$$

   Therefore,

$$\frac{1}{1+\|A\|} \le \left\|(I-A)^{-1}\right\| \le \frac{1}{1-\|A\|}.$$

□

Lemma 0.6. *Suppose that $A \in \mathbb{R}^{n\times n}$. If $(I-A)$ is singular, then $\|A\| \geq 1$.*

Proof. Converse-negative proposition of If $\|A\| < 1$, then $(I-A)$ is nonsingular. □

Theorem 0.7. *Let A be a nonnegative matrix. then $\rho(A) < 1$ if only if $I-A$ is nonsingular and $(I-A)^{-1}$ is nonnegative.*

Proof.
1. $\Rightarrow$ By theorem (0.5).
2. $\Leftarrow$ since $I-A$ is nonsingular and $(I-A)^{-1}$ is nonnegative, by the Perron- Frobenius theorem, there is a nonnegative eigenvector u associated with $\rho(A)$, which is an eigenvalue, i.e.

$$Au = \rho(A)u$$

or

$$(I-A)^{-1}u = \frac{1}{1-\rho(A)}u.$$

since $I-A$ is nonsingular and $(I-A)^{-1}$ is nonnegative, this show that $1-\rho(A) > 0$, which implies

$$\rho(A) < 1.$$

□

0.3 Eigenvalues

0.3.1 Eigenvalues of Tridiagonal Matrices

Theorem 0.8. *(Eigenvalues of Tridiagonal Matrices) If $A = diag(b,a,b) \in \mathcal{M}_n$ is an tridiagonal matrix, then the eigenvaules of A are*

$$\lambda_k = a + 2b\cos(\theta_k),\ \ k = 1,2,3,\cdots,N$$

and its corresponding eigenvector are

$$\vec{\xi}_k = \sqrt{2}\,(\sin(1\theta_k),\sin(2\theta_k),\cdots,\sin(N\theta_k))$$

where

$$\theta_k = k\theta = k\pi h = \frac{k\pi}{N+1}.$$

Proof. It can be easily verified by the trigonometric identities

$$\sin(2\theta_k) = 2\sin(\theta_k)\cos(\theta_k),$$

and

$$2\sin(ki\theta_k)\cos(k\theta_k) = \sin(k(i-1)\theta_k) + \sin(k(i+1)\theta_k).$$

□

0.3.2 Eigenvalues of Hermitians Matrices

Definition 0.9. *(Hermitian Matrix) A matrix is Hermitian , if*

$$A^H = A.$$

Definition 0.10. *Let A be Hermitian , then the spectral of A, $\sigma(A)$, is real.*

Proof. Let $\lambda \in \sigma(A)$ with corresponding eigenvector v. Then

$$\begin{aligned} < Av, v > &= < \lambda v, v >= \lambda < v, v > \\ < Av, v > &= < v, A^H v >=< v, \bar{\lambda} v >= \bar{\lambda} < v, v > . \end{aligned}$$

Since $< v, v >\neq 0$,therefore $\lambda = \bar{\lambda}$. Hence λ is real. □

Definition 0.11. *Let A be Hermitian , then the different eigenvector are orthogonal i.e.*

$$< v_i, v_j >= 0, i \neq j. \tag{0.3}$$

Proof. Let λ_1, λ_2 be the arbitrary two different eigenvalues with corresponding eigenvector v_1, v_2. Then

$$\begin{aligned} < Av_1, v_2 > &= < \lambda_1 v_1, v_2 >= \lambda_1 < v_1, v_2 > \\ < Av_1, v_2 > &= < v_1, A^* v_2 >=< v_1, Av_2 >=< v, \lambda_2 v_2 >= \lambda_2 < v_1, v_2 > . \end{aligned}$$

Since $\lambda_1 \neq \lambda_2$,therefore $< v_1, v_2 >= 0$. □

Theorem 0.12. *(Spectral Theorem for Hermitian matrices) A is Hermitian , then A is unitary diagonalizable .*

$$A = UDU^{-1} = UDU^H, \tag{0.4}$$

where U is a unitary matrix , D is an diagonal matrix .

0.4 Spectral radius

Definition 0.13. *The spectral radius of a matrix $A \in \mathcal{M}_n$, defined as*

$$\rho(A) := \max\{|\lambda|, \ \lambda \text{ eigenvalue of } A\}.$$

Theorem 0.14. *If $\|\cdot\|$ is a matrix norm on $\mathcal{M}_n$, then for any $A \in \mathcal{M}_n$,*

$$\rho(A) \leq \|A\|.$$

Proof. Let λ be an eigenvalue of A, and let $x \neq 0$ be a corresponding eigenvector. Then $Ax = \lambda x$, we have

$$AX = \lambda X, \ \text{ where } \ X := [x|\cdots|x] \in \mathcal{M}_n \backslash \{0\}.$$

By the properties of Matrix norm, then we have

$$|\lambda| \|X\| = \|\lambda X\| = \|AX\| \leq \|A\| \|X\|.$$

Since, $\|X\| \neq 0$, then we have

$$|\lambda| \leq \|A\|.$$

Taking the maximum over all eigenvalues λ yields

$$\rho(A) \leq \|A\|.$$

□

Theorem 0.15. *If $A = A^*$, then $\rho(A) = \|A\|_2$.*

Proof. Since A is self-adjoint, there an orthonormal basis of eigenvector $x \in \mathbb{C}^n$, s.t.

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n.$$

Moreover, $Ae_i = \lambda_i e_i$, $\|e_i\| = 1$ and $(e_i, e_j) = 0$ when $i \not\equiv j, (e_j, e_j) = 1$. So,

$$\|x\|_{\ell^2}^2 = \sum_{i=1}^{n} |\alpha_i|^2,$$

since,

$$\begin{aligned}
(x,x) &= (\sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \alpha_j e_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \bar{\alpha}_j e_i e_j \\
&= \sum_{i=1}^{n} |\alpha_i|^2.
\end{aligned}$$

Since, $Ax = A(\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n) = \alpha_1 \lambda_1 e_1 + \alpha_2 \lambda_2 e_2 + \cdots + \alpha_n \lambda_n e_n$, then

$$\|Ax\|_{\ell^2}^2 = \sum_{i=1}^{n} |\lambda_i \alpha_i|^2 = \sum_{i=1}^{n} |\lambda_i|^2 |\alpha_i|^2 \leq \max\{|\lambda_i|\}^2 \sum_{i=1}^{n} |\alpha_i|^2.$$

Therefore,

$$\|Ax\|_{\ell^2} \leq \rho(A) \|x\|_{\ell^2},$$

i.e.

$$\|A\|_2 = \sup_{x \in \mathbb{C}^n} \frac{\|Ax\|_{\ell^2}}{\|x\|_{\ell^2}} \leq \rho(A).$$

Let k be the index, s.t: $|\lambda_n| = \rho(A)$ and $x = e_k, Ax = Ae_k = \lambda_n e_k$, so $\|Ax\|_{\ell^2} = |\lambda_n| = \rho(A)$ and

$$\|A\|_2 = \sup_{x \in \mathbb{C}^n} \frac{\|Ax\|_{\ell^2}}{\|x\|_{\ell^2}} \geq \frac{\|Ax\|_{\ell^2}}{\|x\|_{\ell^2}} = \rho(A).$$

□

0.5 Hermitian and Symmetric

Definition 0.16. *(Hermitian ans Symmetric) Let $\vec{u},\vec{v} \in \mathbb{C}^n$. Define*

$$(\vec{u},\vec{v}) := \vec{u}^{\,H}\vec{v} = \sum_{i=1}^{n} \overline{u}_i v_i.$$

1. *Let $A \in \mathbb{C}^{n\times n}$, we say that A is Hermitian iff*

$$A = A^H,$$

where $A^H \in \mathbb{C}^{n\times n}$ is the matrix satisfying

$$(\vec{u},A\vec{v}) = \left(A^H\vec{u},\vec{v}\right),$$

for all $\vec{u},\vec{v} \in \mathbb{C}^n$.

2. *$A \in \mathbb{R}^{n\times n}$ is called Symmetric iff*

$$A = A^T,$$

where $A^T \in \mathbb{R}^{n\times n}$ is the matrix satisfying

$$(\vec{u},A\vec{v}) = \left(A^T\vec{u},\vec{v}\right),$$

for all $\vec{u},\vec{v} \in \mathbb{R}^n$.

Definition 0.17. *(HPD and SPD) A is called Hermitian Positive Definite(HPD) iff $A = A^H$ and*

$$(\vec{u},A\vec{u}) > 0, \quad \forall \ \vec{u} \in \mathbb{C}^n_* := \mathbb{C}^n \backslash \left\{\vec{0}\right\}.$$

And A is called Symmetric Positive Definite(SPD) iff $A = A^T$ and

$$(\vec{u},A\vec{u}) > 0, \quad \forall \ \vec{u} \in \mathbb{R}^n_* := \mathbb{R}^n \backslash \left\{\vec{0}\right\}.$$

Remark 0.18. *Of course, if $A \in \mathbb{R}^{n\times n}$, then HPD and SPD mean the same things. We also have the usual simple formulas for the Hermitian and real transposes. For $A \in \mathbb{C}^{n\times n}$,*

$$\left[A^H\right]_{ij} = \overline{[A]_{ij}},$$

and if $A \in \mathbb{R}^{n\times n}$

$$A^H = A^T.$$

Definition 0.19. *(Orthogonal and Orthonormal) Two vectors $\vec{u},\vec{v} \in \mathbb{C}^n$ are called orthogonal (Orthogonal in canonical inner product) iff*

$$(\vec{u},\vec{v}) = 0.$$

The set $S = \{\vec{u}_1,\cdots,\vec{u}_\ell\} \in \mathbb{C}^n$ is called orthogonal iff

$$\left(\vec{u}_i,\vec{u}_j\right) = 0, \quad 1 \le i,j \le \ell, \quad i \ne j.$$

S is called orthonormal iff it is orthogonal and

$$(\vec{u}_i,\vec{u}_i) = 1, \quad 1 \le i \le \ell.$$

In this case,

$$\left(\vec{u}_i,\vec{u}_j\right) = \delta_{ij}, \quad 1 \le i,j \le \ell.$$

Theorem 0.20. *(Properties of Hermitian) Suppose that $A \in \mathbb{C}^{n\times n}$ is Hermitian, i.e. $A = A^H$,*

1. *All of the eigenvalues are real,*
2. *Eigenvector of A associated to distinct eigenvalues of A are orthogonal,*
3. *A has a full set of n eigenvectors that forms an orthonormal basis for $\mathbb{C}^n$.*

Proof.

1. Let $\lambda \in \sigma(A)$ with corresponding eigenvector $\vec{v}$. Then

$$\begin{aligned} <A\vec{v},\vec{v}> &= <\lambda\vec{v},\vec{v}> = \lambda <\vec{v},\vec{v}> \\ <Av,v> &= <v,A^H v> = <v,\bar{\lambda}v> = \bar{\lambda}<v,v>. \end{aligned}$$

Since $<\vec{v},\vec{v}> \ne 0$,therefore $\lambda = \bar{\lambda}$. Hence λ is real.

2. Let λ_1,λ_2 be the arbitrary two different eigenvalues with corresponding eigenvector $\vec{v}_1,\vec{v}_2$. Then

$$\begin{aligned} <A\vec{v}_1,\vec{v}_2> &= <\lambda_1\vec{v}_1,\vec{v}_2> = \lambda_1 <\vec{v}_1,\vec{v}_2> \\ <A\vec{v}_1,\vec{v}_2> &= <v_1,A^H\vec{v}_2> = <\vec{v}_1,A\vec{v}_2> = <\vec{v},\lambda_2\vec{v}_2> = \lambda_2 <\vec{v}_1,\vec{v}_2>. \end{aligned}$$

Since $\lambda_1 \ne \lambda_2$,therefore $<\vec{v}_1,\vec{v}_2> = 0$.

3. Gram-Schmidt Orthonormalization.

□

Theorem 0.21. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian,*

1. *If A is HPD then the eigenvalues of A are positive.*
2. *Conversely, if the eigenvalues are positive, then A is HPD.*

Proof.

1. Let $(\lambda,\vec{v})$ be arbitrary eigen-pair of A, then we have

$$\vec{v}^H A\vec{v} = \vec{v}^H\lambda\vec{v} = \lambda\vec{v}^H\vec{v} = \lambda(v_1^2+v_2^2+\cdots+v_n^2) > 0.$$

for all $\vec{v} \ne 0$. Hence, λ is positive.

2. Since $A \in \mathbb{C}^{n\times n}_{her}$, then the eigenvalue of A are real. Let λ be arbitrary eigenvalue of A, then

$$\begin{aligned} (A\vec{v},\vec{v}) &= (\lambda\vec{v},\vec{v}) = \lambda(\vec{v},\vec{v}), \\ (A\vec{v},\vec{v}) = (\vec{v},A^H\vec{v}) = (\vec{v},A\vec{v}) &= (\vec{v},\lambda\vec{v}) = \overline{\lambda}(\vec{v},\vec{v}), \end{aligned}$$

and then $\lambda = \overline{\lambda}$, so λ is real. Moreover, we have λ is positive, so

$$\vec{v}^H A\vec{v} = \vec{v}^H \lambda\vec{v} = \lambda\vec{v}^H\vec{v} = \lambda(v_1^2 + v_2^2 + \cdots + v_n^2) > 0.$$

for all $\vec{v} \neq 0$. Hence, A is Hermitian Positive Definite.

□

Definition 0.22. *(A-inner product) Let $A \in \mathbb{C}^{n\times n}$ be HPD. Define for all $\vec{u},\vec{v} \in \mathbb{C}^n$*

$$(\vec{u},\vec{v}) := (A\vec{u},\vec{v}).$$

Theorem 0.23.
1. *Let $A \in \mathbb{C}^{n\times n}$ be HPD. Then $(\cdot,\cdot)_A$ defines a bona-fide inner product on $\mathbb{C}^n$.*
2. *Conversely, suppose $(\cdot,\cdot)_*$ is an inner product on $\mathbb{C}^n$. Then there exists an HPD matrix (w.r.t. the canonical inner product) $C \in \mathbb{C}^{n\times n}$ such that*

$$(\vec{u},\vec{v})_C = (\vec{u},\vec{v})_*, \quad \forall\, \vec{u},\vec{v} \in \mathbb{C}^n.$$

Theorem 0.24. *Suppose that A and B are $n \times n$ HPD matrices with respect to $(\cdot,\cdot)$. Then BA is HPD with respect to $(\cdot,\cdot)_A$ and $(\cdot,\cdot)_{B^{-1}}$.*

Proof.
1. Hermitian-ness:

$$\begin{aligned} (BA\vec{u},\vec{v})_A &= (BA\vec{u},A\vec{v}) \\ &= (A\vec{u},BA\vec{v}) \\ &= (\vec{u},BA\vec{v})_A, \end{aligned}$$

for all $\vec{u},\vec{v} \in \mathbb{C}^n$.

2. Positive Definite-ness: Suppose $\vec{u} \in \mathbb{C}^n_*$, $\vec{z} := A\vec{u}$. Then $\vec{z}\mathbb{C}^n_*$. Hence

$$\begin{aligned} (BA\vec{u},\vec{u})_A &= (BA\vec{u},A\vec{u}) \\ &= (A\vec{z},\vec{z}) \\ &> 0. \quad (\vec{z} \neq \vec{0}) \end{aligned}$$

The proof for $(\cdot,\cdot)_{B^{-1}}$ is similar.

□

1 Classical Iterative Methods

In this section, we shall discuss the classical linear iterative methods on solving the linear operator equation

$$A\vec{u} = \vec{f}, \tag{1.1}$$

which arises from the Finite Difference Method (FDM) or Finite Element Method (FEM). We will assume that A is invertible, at least, often will assume that A is HPD.

1.1 General Iterative Methods

Definition 1.1. *(General Iterative Scheme) A general linear two layer iterative scheme reads*

$$B_k^{-1}\left(\frac{\vec{u}^{\,k+1} - \vec{u}^{\,k}}{\alpha_k}\right) + A\vec{u}^{\,k} = \vec{f}.$$

1. *$\alpha_k \in R, B_k^{-1} \in \mathbb{C}^{n\times n}$–iterative parameters*
2. *If $\alpha_k = \alpha, B_k^{-1} = B^{-1}$, then the method is stationary.*
3. *If $B_k^{-1} = I$, then the method is explicit.*

From now on, we consider the stationary scheme ($\alpha > 0$), i.e

$$B^{-1}\left(\frac{\vec{u}^{\,k+1} - \vec{u}^{\,k}}{\alpha}\right) + A\vec{u}^{\,k} = \vec{f}. \tag{1.2}$$

Then we get

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + \alpha B(\vec{f} - A\vec{u}^{\,k}). \tag{1.3}$$

Definition 1.2. *(General Linear Iterative Scheme (GLIS)) Suppose that $B \in \mathbb{C}^{n\times n}$ consider the iteration: Given $\vec{u}^{\,0}$, find $\vec{u}^{\,1}, \vec{u}^{\,2}, \cdots$, such that*

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}). \tag{1.4}$$

This is known as a General Linear Iterative Scheme (GLIS) .

Remark 1.3. *The application of B should be simple and cheap (computationally inexpensive). In some sense B should approximate A^{-1}.*

Remark 1.4. *In the General Linear Iteration Scheme (GLIS) (1.4), the matrix B which should approximate $\in A$ is often called a preconditioner of A, especially, when A and B are both HPD. Suppose $\vec{u}^{\,0} = \vec{0}$. Then*

$$\vec{u}^{\,1} = B\vec{f}.$$

Thus the action of the preconditioner B on a vector f is equivalentt to 1 iteration of the GLIS with $\vec{u}^{\,0} = \vec{0} \in \mathbb{C}^n$. A GLIS (where B is HPD) may used (viewed as) a precondition for a Krylov method. A "good" preconditioned has the property that

$$\kappa(BA) = \frac{\lambda_n(BA)}{\lambda_1(BA)} = \mathcal{O}(1).$$

where

$$0 < \lambda_1(BA) \le \lambda_2(BA) \le \cdots \le \lambda_n(BA),$$

are the eigenvalues of the HPD matrix BA. (This is HPD w.r.t. $(\cdot,\cdot)_A$ *and* $(\cdot,\cdot)_{B^{-1}}$*.)*

Definition 1.5. *(Error Transfer Operator of the GLIS) Let $\vec{u}$ be the exact solution and $\vec{u}^{\,k}$ be the approximate solution at k step. Then we define error, residual and error Transfer Operator as follows:*

$$\text{Error}: \quad \vec{e}^{\,k} := \vec{u} - \vec{u}^{\,k}. \tag{1.5}$$

$$\text{Residual}: \quad \vec{r}^{\,k} := \vec{f} - A\vec{u}^{\,k}. \tag{1.6}$$

$$\text{Error transfer operator}: \quad \vec{e}^{\,k+1} := (I - BA)\vec{e}^{\,k} := T\vec{e}^{\,k}. \tag{1.7}$$

Remark 1.6. *Since from the general iterative scheme we can get*

$$\begin{aligned} \vec{u} &= \vec{u} + B(\vec{f} - A\vec{u}), \\ \vec{u}^{\,k+1} &= \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}). \end{aligned}$$

Then, subtracting the above two equations yields

$$\vec{e}^{\,k+1} = \vec{e}^{\,k} - BA\vec{e}^{\,k} = (I - BA)\vec{e}^{\,k} := T\vec{e}^{\,k}.$$

$T = I - BA$ is the error transfer operator.

Definition 1.7. *(Energy norm w.r.t A) The Energy norm associated with A is*

$$\|\vec{u}\|_A = (A\vec{u}, \vec{u});$$

Theorem 1.8. *(convergence in energy norm) Suppose A is HPD and B is invertible. If $Q := B^{-1} - \frac{\alpha}{2}A > 0$, then GLIS (1.3) converges with A norm, i.e.*

$$\|\vec{e}^{\,k}\|_A \to 0.$$

Proof. Let $\vec{u}$ be the exact solution and $\vec{u}^{\,k}$ be the approximate solution at k step, then

$$\begin{aligned} \vec{u} &= \vec{u} + \alpha B(\vec{f} - A\vec{u}), \\ \vec{u}^{\,k+1} &= \vec{u}^{\,k} + \alpha B(\vec{f} - A\vec{u}^{\,k}). \end{aligned}$$

From the definite of Error (1.5), we have

$$\begin{aligned} \vec{e}^{\,k+1} &= \vec{e}^{\,k} + \alpha B\left(\vec{f} - A\vec{u} - \vec{f} + A\vec{u}^{\,k}\right) \\ &= \vec{e}^{\,k} - \alpha BA\vec{e}^{\,k}. \end{aligned}$$

Let $\vec{v}^{\,k+1} = \vec{e}^{\,k+1} - \vec{e}^{\,k}$, then

$$\frac{1}{\alpha}B^{-1}\vec{v}^{\,k+1} + A\vec{e}^{\,k} = 0.$$

Taking inner product with $\vec{v}^{\,k+1}$ gives

$$\frac{1}{\alpha}(B^{-1}\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) + (A\vec{e}^{\,k}, \vec{v}^{\,k+1}) = 0.$$

Since

$$\vec{e}^{\,k} = \frac{1}{2}(\vec{e}^{\,k+1} + \vec{e}^{\,k}) - \frac{1}{2}(\vec{e}^{\,k+1} - \vec{e}^{\,k}) = \frac{1}{2}(\vec{e}^{\,k+1} + \vec{e}^{\,k}) - \frac{1}{2}\vec{v}^{\,k+1},$$

then

$$\begin{aligned}
0 &= \frac{1}{\alpha}(B^{-1}\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) + (A\vec{e}^{\,k}, \vec{v}^{\,k+1}) \\
&= \frac{1}{\alpha}(B^{-1}\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) + \frac{1}{2}(A(\vec{e}^{\,k+1} + \vec{e}^{\,k}), \vec{v}^{\,k+1}) - \frac{1}{2}(A\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) \\
&= \frac{1}{\alpha}((B^{-1} - \frac{\alpha}{2}A)\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) + \frac{1}{2}(A(\vec{e}^{\,k+1} + \vec{e}^{\,k}), \vec{v}^{\,k+1}) \\
&= \frac{1}{\alpha}((B^{-1} - \frac{\alpha}{2}A)\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) + \frac{1}{2}(A(\vec{e}^{\,k+1} + \vec{e}^{\,k}), \vec{e}^{\,k+1} - \vec{e}^{\,k}) \\
&= \frac{1}{\alpha}((B^{-1} - \frac{\alpha}{2}A)\vec{v}^{\,k+1}, \vec{v}^{\,k+1}) + \frac{1}{2}(\left\|\vec{e}^{\,k+1}\right\|_A^2 - \left\|\vec{e}^{\,k}\right\|_A^2)
\end{aligned}$$

By assumption, $Q := B^{-1} - \frac{\alpha}{2}A > 0$, i.e. there exists a $m > 0$, s.t.

$$(Q\vec{u}, \vec{u}) \geq m\left\|\vec{u}\right\|_2^2.$$

Therefore,

$$\frac{m}{\alpha}\left\|\vec{v}^{\,k+1}\right\|_2^2 + \frac{1}{2}(\left\|\vec{e}^{\,k+1}\right\|_A^2 - \left\|\vec{e}^{\,k}\right\|_A^2) \leq 0.$$

i.e.

$$\frac{2m}{\alpha}\left\|\vec{v}^{\,k+1}\right\|_2^2 + \left\|\vec{e}^{\,k+1}\right\|_A^2 \leq \left\|\vec{e}^{\,k}\right\|_A^2.$$

Hence

$$\left\|\vec{e}^{\,k+1}\right\|_A^2 \leq \left\|\vec{e}^{\,k}\right\|_A^2.$$

and

$$\left\|\vec{e}^{\,k+1}\right\|_A^2 \to 0.$$

□

Definition 1.9. *(Convergence) Let $C \in \mathbb{C}^{n\times n}$. C is called convergent iff for every $\epsilon > 0$, there exists a $K \in \mathbb{N}$ such that if $k > K$ then for all $1 \leq i, j \leq n$*

$$\left|\left[C^k\right]_{i,j}\right| \leq \epsilon.$$

We have the following equivalences:

Theorem 1.10. *The following statements are equivalent*

1. $C \in \mathbb{C}^{n\times n}$ *is convergent,*
2. $$\lim_{k\to\infty} \left\|C^k\right\| = 0,$$
   for some induced matrix norm (operator norm),
3. $$\lim_{k\to\infty} \left\|C^k\right\| = 0,$$
   for all induced matrix norms.
4. $\rho(C) < 1$, *where*
   $$\rho(C) = \max_{1\le i\le n} |\lambda_i|,$$
   and $\lambda_i \in \mathbb{C}$, $1 \le i \le n$, *are the eigenvalues of* C.
5. $$\lim_{k\to\infty} C^k\vec{x} = \vec{0},$$

for every $\vec{x} \in \mathbb{C}^n$.

Thus we have the following convergence results.

Theorem 1.11. *(Sufficient & necessary condition for convergence) The sufficient & necessary condition for convergence of (1.4) is*

$$\rho(T) < 1,$$

for any starting vector $\vec{u}^{\,0} \in \mathbb{C}^n$*.Where* $\rho(T)$ *is the spectral radius of* $T = I - BA$.

Proof. Let $\vec{u}$ be the exact solution and $\vec{u}^k$ be the approximate solution at k step. Then, from Error transfer operator, we have $\vec{e}^{k+1} = T\vec{e}^k$. Hence

$$\vec{e}^{\,k} = T^k\vec{e}^{\,0}.$$

- $\Rightarrow$ Suppose (1.4) converges (it must converge to $\vec{u} \in \mathbb{C}^n$, incidentally). Then

  $$\lim_{k\to\infty} \vec{e}^{\,k} = \vec{0},$$

  for any $\vec{u}^{\,0} \in \mathbb{C}^n$. Let $(\lambda, \vec{w}) \in \mathbb{C}\times\mathbb{C}^n$ be the eigen-pair of T and set $\vec{e}^{\,0} = \vec{w}$. Then

  $$\vec{e}^{\,k} = \lambda^k\vec{e}^{\,0}.$$

  This implies that $|\lambda| < 1$. Since λ was an arbitrary eigenvalue,

  $$\rho(T) < 1$$

  .

- $\Leftarrow$ If $\rho(T) < 1$, then T by theorem 1.10,

$$\lim_{k\to\infty} T^k \vec{x} = \vec{0},$$

for any $\vec{x} \in \mathbb{C}^n$. Hence

$$\lim_{k\to\infty} \vec{e}^{\,k} = \lim_{k\to\infty} T^k \vec{e}^{\,0} = \vec{0},$$

$\square$

Theorem 1.12. *(Sufficient condition for convergence) The sufficient condition for convergence of (1.4) is*

$$\|T\| < 1.$$

Where $\|\cdot\|$ is any induced matrix norm.

Proof. By theorem 0.14, we have

$$\rho(T) \le \|T\| < 1.$$

Using theorem 1.11 gives the result. $\square$

Theorem 1.13. *(Error estimates) If (1.3) is convergent, then we have*

1. $\left\|\vec{u} - \vec{u}^{\,k}\right\| \le \|T\|^k \left\|\vec{u} - \vec{u}^{\,0}\right\|$
2. $\left\|\vec{u} - \vec{u}^{\,k}\right\| \le \frac{\|T\|^k}{1-\|T\|} \left\|\vec{u}^{\,1} - \vec{u}^{\,0}\right\|$

Proof. 1. By Lemma 0.4

$$\begin{aligned} \left\|\vec{e}^{\,k}\right\| &= \left\|T^k \vec{e}^{\,0}\right\| \\ &\le \left\|T^k\right\| \left\|\vec{e}^{\,0}\right\| \\ &\le \|T\|^k \left\|\vec{e}^{\,0}\right\|, \end{aligned}$$

i.e.

$$\left\|\vec{u} - \vec{u}^{\,k}\right\| \le \|T\|^k \left\|\vec{u} - \vec{u}^{\,0}\right\|.$$

2. Now, write

$$\begin{aligned} \left\|\vec{e}^{\,0}\right\| &= \left\|\vec{u} - \vec{u}^{\,0}\right\| \\ &= \left\|\vec{u} - \vec{u}^{\,1} + \vec{u}^{\,1} - \vec{u}^{\,0}\right\| \\ &\le \left\|\vec{u} - \vec{u}^{\,1}\right\| + \left\|\vec{u}^{\,1} - \vec{u}^{\,0}\right\| \end{aligned}$$

From part 1, we have

$$\left\|\vec{u} - \vec{u}^{\,1}\right\| \le \|T\| \left\|\vec{u} - \vec{u}^{\,0}\right\| = \|T\| \left\|\vec{e}^{\,0}\right\|.$$

Therefore,

$$\left\|\vec{e}^{\,0}\right\| \le \|T\| \left\|\vec{e}^{\,0}\right\| + \left\|\vec{u}^{\,1} - \vec{u}^{\,0}\right\|.$$

Hence

$$\left\|\vec{e}^{\,0}\right\| \le \frac{1}{1-\|T\|}\left\|\vec{u}^{\,1}-\vec{u}^{\,0}\right\|.$$

Substituting into first estimate yields

$$\left\|\vec{e}^{\,k}\right\| \le \|T\|^k \left\|\vec{u}-\vec{u}^{\,0}\right\| = \|T\|^k \left\|\vec{e}^{\,0}\right\| \le \frac{\|T\|^k}{1-\|T\|}\left\|\vec{u}^{\,1}-\vec{u}^{\,0}\right\|.$$

□

Let $A \in \mathbb{C}^{n\times n}$, we start with the following matrix decomposition

$$A = D - L - U, \tag{1.8}$$

where D is the diagonal of A, $-L$ and $-U$ are the strict lower part and the strict upper part, respectively, as illustrated in (Figure.1).

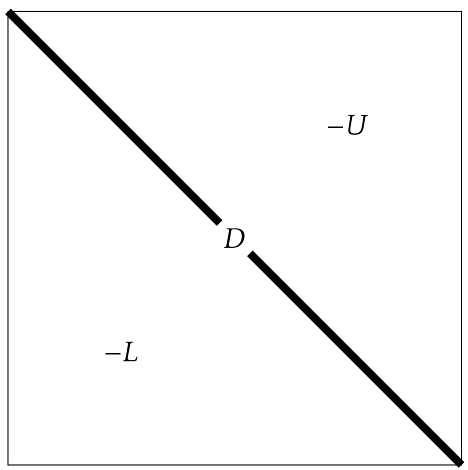


Figure 1: The initial decomposition of matrix A.

If A is HPD, then D is real and $U = L^H$.

1.2 Jacobi Iterative Method

Assume D is invertible, then the linear system (1.1) can be rewritten as

$$D\vec{u} = (U+L)\vec{u} + \vec{f}.$$

Definition 1.14. *(Jacobi Method) Define*

$$D\vec{u}^{\,k+1} = (U+L)\vec{u}^{\,k} + \vec{f}.$$

for a given starting value $\vec{u}^{\,0} \in \mathbb{C}^n$, then Jacobi Method scheme reads

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + D^{-1}\left(\vec{f} - A\vec{u}^{\,k}\right). \tag{1.9}$$

Remark 1.15. *Since the matrix decomposition (1.8) and*

$$D\vec{u}^{\,k+1} = (U+L)\vec{u}^{\,k} + \vec{f},$$

we have

$$\begin{aligned}\vec{u}^{\,k+1} &= D^{-1}(U+L)\vec{u}^{\,k} + D^{-1}\vec{f}\\ &= D^{-1}(D-A)\vec{u}^{\,k} + D^{-1}\vec{f}\\ &= \vec{u}^{\,k} + D^{-1}\left(\vec{f} - A\vec{u}^{\,k}\right).\end{aligned}$$

Hence Jacobi Method is a GLIS and $B_J = D^{-1}$. Moreover, from the definition of (1.14), we have

$$D\left(\vec{u}^{\,k+1} - \vec{u}^{\,k}\right) + (L+D+U)\vec{u}^{\,k} = L\vec{u}^{\,k} + D\vec{u}^{\,k+1} + U\vec{u}^{\,k} = \vec{f}.$$

So, the Jacobi iterative method can be written as

$$\sum_{j<i} a_{ij}u_j^k + a_{ii}u_i^{k+1} + \sum_{j>i} a_{ij}u_j^k = f_i,$$

or

$$u_i^{k+1} = \frac{1}{a_{ii}}\left(f_i - \sum_{j\neq i} a_{ij}u_j^k\right).$$

Definition 1.16. *(Preconditioner and Error Transfer Operator for Jacobi Method) The preconditioner of Jacobi Method is*

$$B_J = D^{-1}.$$

And the error transfer operator for Jacobi Method is as follows

$$T_J = I - D^{-1}A.$$

Remark 1.17. *Since $B_J = D^{-1}$ and the definition of the error transform operator (1.5), we have*

$$T_J = I - D^{-1}A.$$

Theorem 1.18. *(convergence of the Jacobi Method) If A is diagonal dominant , then the Jacobi Method convergences.*

Proof. We want to show If A is diagonal dominant , then $\|T_J\| < 1$, then Jacobi Method convergences. From the definition of T, we know that T for Jacobi Method is as follows

$$T_J = I - D^{-1}A.$$

In the matrix form is

$$T = \begin{pmatrix} 1 & & 0\\ & \ddots & \\ 0 & & 1\end{pmatrix} - \begin{pmatrix} \frac{1}{a_{11}} & & 0\\ & \ddots & \\ 0 & & \frac{1}{a_{nn}}\end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix} = \left[t_{ij}\right] = \begin{cases} t_{ij} = 0,\ i=j,\\ t_{ij} = -\frac{a_{ij}}{a_{ii}},\ i\neq j.\end{cases}$$

So,

$$\|T\|_\infty = \max_i \sum_j |t_{ij}| = \max_i \sum_{i\neq j} |\frac{a_{ij}}{a_{ii}}|.$$

Since A is diagonal dominant, so

$$|a_{ii}| \geq \sum_{j\neq i} |a_{ij}| + \delta.$$

Therefore,

$$1 \geq \sum_{j\neq i} \frac{|a_{ij}|}{|a_{ii}|} + \frac{\delta}{|a_{ii}|}.$$

Hence, $\|T\|_\infty < 1$ □

Definition 1.19. *(Weighted Jacobi Method) Assume D in (1.8) is invertible, define for $0 < \omega \leq 1$,*

$$\begin{aligned} \vec{z} &= \vec{u}^{\,k} + D^{-1}(\vec{f} - A\vec{u}^{\,k}) \\ \vec{u}^{\,k+1} &= \omega\vec{z} + (1-\omega)\vec{u}^{\,k}. \end{aligned}$$

for a given starting value $\vec{u}^{\,0} \in \mathbb{C}^n$, then Weighted Jacobi Method scheme reads

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + \omega D^{-1}\left(\vec{f} - A\vec{u}^{\,k}\right). \tag{1.10}$$

Remark 1.20. *Weighted Jacobi Method is a GLIS with*

$$B_{WJ} = \omega D^{-1}.$$

1.3 Gauss-Seidel Method

Assume $D-L$ is invertible, define

$$\begin{aligned} \vec{u}^{\,k+1} &= (D-L)^{-1}U\vec{u}^{\,k} + (D-L)^{-1}\vec{f} \\ &= \vec{u}^{\,k} + (D-L)^{-1}(\vec{f} - A\vec{u}^{\,k}). \end{aligned}$$

Definition 1.21. *(Forward Gauss-Seidel Method) Suppose $D-L$ is invertible, Forward Gauss-Seidel Method scheme reads*

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + B_{GS}(\vec{f} - A\vec{u}^{\,k}),$$

where

$$B_{GS} = (D-L)^{-1}.$$

Definition 1.22. *(Error Transfer Operator of Forward Gauss-Seidel Method) The error transfer operator for Forward Gauss-Seidel Method is as follows*

$$T_{GS} = (D-L)^{-1}U.$$

Remark 1.23.

$$\begin{aligned} T_{GS} &= I-(D-L)^{-1}A \\ &= I-(D-L)^{-1}(D-L-U) \\ &= (D-L)^{-1}U. \end{aligned}$$

Theorem 1.24. *(convergence of the Forward Gauss-Seidel Method) If A is diagonal dominant , then the Forward Gauss-Seidel Method convergences.*

Proof. We want to show If A is diagonal dominant , then $\|T_{GS}\| < 1$, then the Forward Gauss-Seidel Method convergences. From the definition of T, we know that T for Forward Gauss-Seidel Method is as follows

$$T_{GS} = (D-L)^{-1}U.$$

Next, we will show $\|T_{GS}\| < 1$. Since A is diagonal dominant, so

$$|a_{ii}| \geq \sum_{j\neq i}|a_{ij}| + \delta = \sum_{j>i}|a_{ij}| + \sum_{j<i}|a_{ij}| + \delta.$$

So,

$$|a_{ii}| - \sum_{j<i}|a_{ij}| \geq \sum_{j>i}|a_{ij}| + \delta,$$

which implies

$$\gamma = max_i\left\{\frac{\sum_{j>i}|a_{ij}|}{|a_{ii}|-\sum_{j<i}|a_{ij}}|\right\} \leq 1.$$

Now, we will show $\|T_{GS}\| < \gamma$. Let $x \in \mathbb{C}^n$ and $y = Tx$, i.e.

$$y = T_{GS}x = (D-L)^{-1}Ux.$$

Let i_0 be the index such that $\left\|y\right\|_\infty = |y_{i_0}|$, then we have

$$|((D-L)y)_{i_0}| = |(Ux)_{i_0}| = |\sum_{j>i_0} a_{i_0j}x_j| \leq \sum_{j>i_0}|a_{i_0j}||x_j| \leq \sum_{j>i_0}|a_{i_0j}|\|x\|_\infty.$$

Moreover

$$|((D-L)y)_{i_0}| = |\sum_{j<i_0} a_{i_0j}y_j + a_{i_0i_0}y_j| \geq |a_{i_0i_0}y_j| - |\sum_{j<i_0} a_{i_0j}y_j| = |a_{i_0i_0}|\left\|y\right\|_\infty - |\sum_{j<i_0} a_{i_0j}y_j| \geq |a_{i_0i_0}|\left\|y\right\|_\infty - \sum_{j<i_0}|a_{i_0j}|\left\|y\right\|_\infty.$$

Therefore, we have

$$|a_{i_0i_0}|\left\|y\right\|_\infty - \sum_{j<i_0}|a_{i_0j}|\left\|y\right\|_\infty \leq \sum_{j>i_0}|a_{i_0j}|\|x\|_\infty,$$

which implies

$$\left\|y\right\|_\infty \leq \frac{\sum_{j>i_0}|a_{i_0j}|}{|a_{i_0i_0}|-\sum_{j<i_0}|a_{i_0j}|}\|x\|_\infty.$$

So,

$$\|T_{GS}x\|_\infty \le \gamma \|x\|_\infty ,$$

which implies

$$\|T_{GS}\|_\infty \le \gamma < 1.$$

□

Definition 1.25. *(Backward Gauss-Seidel Method) Suppose $D-U$ is invertible, Backward Gauss-Seidel Method scheme reads*

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + B_{BGS}(\vec{f} - A\vec{u}^{\,k}),$$

where

$$B_{BGS} = (D-U)^{-1}.$$

Definition 1.26. *(Error Transfer Operator of Backward Gauss-Seidel Method) The error transfer operator for Backward Gauss-Seidel Method is as follows*

$$T_{BGS} = (D-U)^{-1}L.$$

Remark 1.27.

$$\begin{aligned} T_{BGS} &= I-(D-U)^{-1}A \\ &= I-(D-U)^{-1}(D-L-U) \\ &= (D-U)^{-1}L. \end{aligned}$$

Theorem 1.28. *(convergence of the Backward Gauss-Seidel Method) If A is diagonal dominant , then the Backward Gauss-Seidel Method convergences.*

Proof. We want to show If A is diagonal dominant , then $\|T_{BGS}\| < 1$, then the Forward Gauss-Seidel Method convergences. From the definition of T, we know that T for Forward Gauss-Seidel Method is as follows

$$T_{BGS} = (D-U)^{-1}L.$$

Next, we will show $\|T_{BGS}\| < 1$. Since A is diagonal dominant, so

$$|a_{ii}| \ge \sum_{j\neq i} |a_{ij}| + \delta = \sum_{j>i} |a_{ij}| + \sum_{j<i} |a_{ij}| + \delta.$$

So,

$$|a_{ii}| - \sum_{j>i} |a_{ij}| \ge \sum_{j<i} |a_{ij}| + \delta,$$

which implies

$$\gamma = max_i \left\{ \frac{\sum_{j<i} |a_{ij}|}{|a_{ii}| - \sum_{j>i} |a_{ij}|} \right\} \le 1.$$

Now, we will show $\|T_{BGS}\| < \gamma$. Let $x \in \mathbb{C}^n$ and $y = Tx$, i.e.

$$y = T_{BGS}x = (D-U)^{-1}Lx.$$

Let i_0 be the index such that $\|y\|_\infty = |y_{i_0}|$, then we have

$$|((D-U)y)_{i_0}| = |(Lx)_{i_0}| = |\sum_{j<i_0} a_{i_0 j}x_j| \le \sum_{j<i_0} |a_{i_0 j}||x_j| \le \sum_{j<i_0} |a_{i_0 j}|\,\|x\|_\infty .$$

Moreover

$$|((D-U)y)_{i_0}| = |\sum_{j>i_0} a_{i_0 j}y_j + a_{i_0 i_0}y_j| \ge |a_{i_0 i_0}y_j| - |\sum_{j>i_0} a_{i_0 j}y_j| = |a_{i_0 i_0}|\,\|y\|_\infty - |\sum_{j>i_0} a_{i_0 j}y_j| \ge |a_{i_0 i_0}|\,\|y\|_\infty - \sum_{j>i_0} |a_{i_0 j}|\,\|y\|_\infty .$$

Therefore, we have

$$|a_{i_0 i_0}|\,\|y\|_\infty - \sum_{j>i_0} |a_{i_0 j}|\,\|y\|_\infty \le \sum_{j<i_0} |a_{i_0 j}|\,\|x\|_\infty ,$$

which implies

$$\|y\|_\infty \le \frac{\sum_{j<i_0} |a_{i_0 j}|}{|a_{i_0 i_0}| - \sum_{j>i_0} |a_{i_0 j}|}\,\|x\|_\infty .$$

So,

$$\|T_{BGS}x\|_\infty \le \gamma\,\|x\|_\infty ,$$

which implies

$$\|T_{BGS}\|_\infty \le \gamma < 1.$$

□

1.4 Richardson Method

Suppose $\omega \in \mathbb{C}_* := \mathbb{C}\setminus\{0\}$, consider the splitting

$$A = \omega I + A - \omega I,$$

suppose

$$A\vec{u} = \vec{f}.$$

Then

$$A\vec{u} = (\omega I - A)\vec{u} + \vec{f}.$$

Definition 1.29. *(Richardson Method) Define Richardson Method scheme as*

$$A\vec{u}^{\,k+1} = (\omega I - A)\vec{u}^{\,k} + \vec{f}.$$

So that

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + \omega^{-1}(\vec{f} - A\vec{u}^{\,k}). \tag{1.11}$$

Definition 1.30. *(Preconditioner and Error Transfer Operator for Gauss-Seidel Method) The preconditioner B_{RC} and error transfer operator T_{RC} of Gauss-Seidel Method are as follows*

$$\begin{aligned} B_{RC} &= \omega^{-1} I, \\ T_{RC} &= I - \omega A. \end{aligned}$$

Theorem 1.31. *(convergence of the Richardson Method) Let $A = A^H > 0$ (HPD). If $0 < \omega < \frac{2}{\lambda_{\max}}$, then the Richardson Method convergences. Moreover, the best acceleration parameter is given by*

$$\omega_{opt} = \frac{2}{\lambda_{\min} + \lambda_{\max}},$$

in which, similarly, $\lambda_{\min}$ is the smallest eigenvalue of $A^T A$.

Proof.

1. From the above lemma, we know that the error transform operator is as follows

$$T_{RC} = I - \omega (B)^{-1} A = I - \omega A.$$

Let $\lambda \in \sigma(A)$, then $\nu := 1 - \omega\lambda \in \sigma(T)$. From the sufficient and & necessary condition for convergence, we know if $\sigma(T) < 1$, then Richardson Method convergences, i.e.

$$|1 - \omega\lambda| < 1,$$

which implies

$$-1 < 1 - \omega\lambda_{\max} \le 1 - \omega\lambda_{\min} < 1.$$

So, we get $-1 < 1 - \omega\lambda_{\max}$, i.e.

$$\omega < \frac{2}{\lambda_{\max}}.$$

2. The minimum is attachment at $|1 - \omega\lambda_{\max}| = |1 - \omega\lambda_{\min}|$(Figure.2), i.e.

$$\omega\lambda_{\max} - 1 = 1 - \omega\lambda_{\min}.$$

Therefore, we get

$$\omega_{opt} = \frac{2}{\lambda_{\min} + \lambda_{\max}}.$$

□

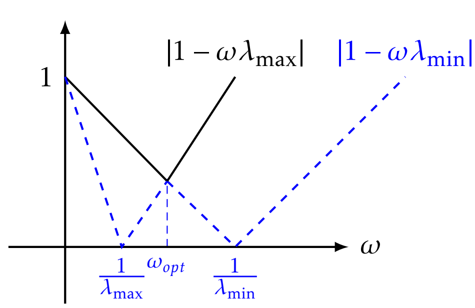


Figure 2: The curve of $\rho(T_{RC})$ as a function of ω

1.5 Symmetrization

1.5.1 symmetrized multiplication

Definition 1.32. *(symmetrized multiplication GLIS) Suppose $A, B \in \mathbb{C}^{n\times n}, \vec{f} \in \mathbb{C}^n$, the Symmetrized Multiplication GLIS (SMGLIS) is defined as follows: given $\vec{u}^{\,0}$, find $\vec{u}^{\,1}, \vec{u}^{\,2}, \cdots$ via*

$$\vec{u}^{\,k+1/2} = \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}), \tag{1.12}$$

$$\vec{u}^{\,k+1} = \vec{u}^{\,k+1/2} + B^H(\vec{f} - A\vec{u}^{\,k+1/2}). \tag{1.13}$$

Remark 1.33. *If the blue part in (1.13) is $\vec{u}^{\,k}$, then the symmetrization is called symmetrized additive method.*

Lemma 1.34. *(symmetrized multiplication GLIS) The SMGLIS can be written as*

$$\vec{u}^{\,k+1} = \vec{u}^{\,k} + B_{SM}(\vec{f} - A\vec{u}^{\,k}), \tag{1.14}$$

where

$$B_{SM} = B + B^H - B^H AB. \tag{1.15}$$

If B is invertible, then

$$B_{SM} = B^H(B^{-H} + B^{-1} - A)B.$$

If A is Hermitian, then B_{SM} is as well.

Proof. 1. Plugging (1.12) into (1.13) yields

$$\begin{aligned}
\vec{u}^{\,k+1} &= \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}) + B^H\left(\vec{f} - A\left[\vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k})\right]\right) \\
&= \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}) + B^H\left(\vec{f} - A\vec{u}^{\,k} - AB(\vec{f} - A\vec{u}^{\,k})\right) \\
&= \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}) + B^H\left(\vec{f} - A\vec{u}^{\,k}\right) - B^H AB(\vec{f} - A\vec{u}^{\,k}) \\
&= \vec{u}^{\,k} + \left(B + B^H - B^H AB\right)(\vec{f} - A\vec{u}^{\,k}).
\end{aligned}$$

2. If B is invertible, then

$$\begin{aligned}
B_{SM} &= B + B^H - B^H AB \\
&= \left(I + B^H B^{-1} - B^H A\right)B \\
&= \left(B^H B^{-H} + B^H B^{-1} - B^H A\right)B \\
&= B^H(B^{-H} + B^{-1} - A)B.
\end{aligned}$$

3. If A is Hermitian, then B_{SM} is Hermitian, since

$$B_{SM} = B + B^H - B^H AB,$$

then

$$\begin{aligned}
B_{SM}^H &= B^H + B - B^H A^H B \\
&= B^H + B - B^H AB \\
&= B_{SM}.
\end{aligned}$$

□

Remark 1.35. *We could also consider a symmetrized additive method?*

$$\vec{u}^{\,k+1/2} = \vec{u}^{\,k} + B(\vec{f} - A\vec{u}^{\,k}), \tag{1.16}$$
$$\vec{u}^{\,k+1} = \vec{u}^{\,k+1/2} + B^H(\vec{f} - A\vec{u}^{\,k}). \tag{1.17}$$

Then

$$B_{SA} = B + B^H.$$

But this is not as useful to us.

1.5.2 Symmetric Gauss-Seidel Method

Definition 1.36. *(Symmetric Gauss-Seidel Method) Assume $A = A^H$, then*

$$U^H = L, \;\; D \in \mathbb{R}^{n\times n},$$

in the splitting (1.8). And

$$B_{GS} = (D-L)^{-1}.$$

So that

$$\begin{aligned} B_{GS}^H &= (D-L)^{-H} \\ &= (D-U)^{-1} \\ &= B_{BGS}. \end{aligned}$$

We have

$$\begin{aligned} B_{SM} := &= B_{SGS} \\ &= B_{GS}^H (B_{GS}^{-H} + {B_{GS}}^{-1} - A) B_{GS} \\ &= (D-U)^{-1}(D-U+D-L-A)(D-L)^{-1}, \end{aligned}$$

with cancellation, we get

$$B_{SGS} = (D-U)^{-1} D (D-L)^{-1}. \tag{1.18}$$

Of course,

$$B_{SGS}^H = B_{SGS}.$$

as desired.

1.5.3 Regular Splittings

In this subsection, we consider the following matrix splitting

$$A = M - N \in \mathbb{C}^{n\times n}.$$

where A is associate with the linear system (1.1).

Theorem 1.37. *Consider the iterative scheme*

$$M\vec{u}^{\,k+1} = N\vec{u}^{\,k} + \vec{f}, \tag{1.19}$$

where

$$A = M - N \in \mathbb{C}^{n\times n}.$$

If both A and $M + M^H - A$ are HPD, then (1.19) converges.

There are two ways to prove this theorem, one way is based on computing the spectral radius, the other way is energy method.

1. *Proof.* We can rewrite (1.19) as a GLIS, i.e.

$$\begin{aligned}
\vec{u}^{\,k+1} &= M^{-1}N\vec{u}^{\,k} + M^{-1}\vec{f} \\
&= \vec{u}^{\,k} - M^{-1}M\vec{u}^{\,k} + M^{-1}N\vec{u}^{\,k} + M^{-1}\vec{f} \\
&= \vec{u}^{\,k} + M^{-1}\left(-M\vec{u}^{\,k} + N\vec{u}^{\,k} + \vec{f}\right) \\
&= \vec{u}^{\,k} + M^{-1}\left(\vec{f} - A\vec{u}^{\,k}\right).
\end{aligned}$$

Observe that M is nonsingular, otherwise $M + M^H - A$ is not HPD. Then by the definition of Error Transfer operator (1.7),

$$T = I - BA = I - M^{-1}A. \tag{1.20}$$

If we can prove that $\rho(T) < 1$, then this method converges. Let $(\lambda, \vec{\omega})$ be an eigenpair of T in (1.20). Then we have

$$T\vec{\omega} = \lambda\vec{\omega} \Rrightarrow (I - M^{-1}A)\vec{\omega} = \lambda\vec{\omega} \Rrightarrow (M - A)x = \lambda M x \Rrightarrow (1-\lambda)M\vec{\omega} = A\vec{\omega}.$$

(a) $\lambda \neq 1$. Otherwise $\lambda = 1$, then $A\vec{\omega} = 0$, contradicting A is HPD.

(b) $\lambda \leq 1$. Since, $(1-\lambda)M\vec{\omega} = A\vec{\omega}$, multiplying $\vec{\omega}^H$ yields

$$(1-\lambda)\vec{\omega}^H M\vec{\omega} = \vec{\omega}^H A\vec{\omega}.$$

Since $\lambda \neq 1$, we have

$$\vec{\omega}^H M\vec{\omega} = \frac{1}{1-\lambda}\vec{\omega}^H A\vec{\omega}.$$

taking conjugate transpose of which gives

$$\vec{\omega}^H M^H\vec{\omega} = \frac{1}{\overline{1-\lambda}}\vec{\omega}^H A^H\vec{\omega} = \frac{1}{\overline{1-\lambda}}\vec{\omega}^H A\vec{\omega}.$$

Adding the above two equations and subtracting $\vec{\omega}^H A\vec{\omega}$, we have

$$\begin{aligned}
\vec{\omega}^H(M + M^H - A)\vec{\omega} &= \left(\frac{1}{1-\lambda} + \frac{1}{\overline{1-\lambda}} - 1\right)\vec{\omega}^H A\vec{\omega} \\
&= \left(\frac{\lambda}{1-\lambda} + \frac{1}{\overline{1-\lambda}}\right)\vec{\omega}^H A\vec{\omega} \\
&= \frac{1-\lambda^2}{|1-\lambda|^2}\vec{\omega}^H A\vec{\omega}.
\end{aligned}$$

Since $M+M^H-A$ and A are SPD, then $\vec{\omega}^H(M+M^H-A)\vec{\omega}>0$ and $\vec{\omega}^H A\vec{\omega}>0$. It must be that

$$1-\lambda^2>0.$$

Hence

$$|\lambda|<1.$$

□

2. *Proof.* Let $\vec{u}$ be the exact solution and $\vec{u}^{\,k}$ be the approximate solution at k step, then

$$\begin{aligned}\vec{u} &= \vec{u}+M^{-1}\left(\vec{f}-A\vec{u}\right),\\ \vec{u}^{\,k+1} &= \vec{u}^{\,k}+M^{-1}\left(\vec{f}-A\vec{u}^{\,k}\right).\end{aligned}$$

From the definite of Error (1.5), we have

$$\begin{aligned}\vec{e}^{\,k+1} &= \vec{e}^{\,k}+M^{-1}\left(\vec{f}-A\vec{u}-\vec{f}+A\vec{u}^{\,k}\right)\\ &= \vec{e}^{\,k}-M^{-1}A\vec{e}^{\,k}.\end{aligned}$$

Let $\vec{v}^{\,k+1}=\vec{e}^{\,k+1}-\vec{e}^{\,k}$, then

$$M\vec{v}^{\,k+1}+A\vec{e}^{\,k}=0.$$

Taking the conjugate transport of the above equation, then we get

$$M^H\vec{v}^{\,k+1}+A^H\vec{e}^{\,k}=M^H\vec{v}^{\,k+1}+A\vec{e}^{\,k}=0.$$

Adding the last two equations yields

$$\frac{M+M^H}{2}v^{k+1}+Ae^k=0.$$

Let $B_s=\frac{M+M^H}{2}$ and taking the inner product of both sides with $\vec{v}^{\,k+1}$ gives

$$(B_s\vec{v}^{\,k+1},\vec{v}^{\,k+1})+(A\vec{e}^{\,k},\vec{v}^{\,k+1})=0.$$

Since

$$\vec{e}^{\,k}=\frac{1}{2}(\vec{e}^{\,k+1}+\vec{e}^{\,k})-\frac{1}{2}(\vec{e}^{\,k+1}-\vec{e}^{\,k})=\frac{1}{2}(\vec{e}^{\,k+1}+\vec{e}^{\,k})-\frac{1}{2}\vec{v}^{\,k+1},$$

then

$$\begin{aligned}0 &= (B_s\vec{v}^{\,k+1},\vec{v}^{\,k+1})+(A\vec{e}^{\,k},\vec{v}^{\,k+1})\\ &= (B_s\vec{v}^{\,k+1},\vec{v}^{\,k+1})+\frac{1}{2}(A(\vec{e}^{\,k+1}+\vec{e}^{\,k}),\vec{v}^{\,k+1})-\frac{1}{2}(A\vec{v}^{\,k+1},\vec{v}^{\,k+1})\\ &= ((B_s-\frac{1}{2}A)\vec{v}^{\,k+1},\vec{v}^{\,k+1})+\frac{1}{2}(A(\vec{e}^{\,k+1}+\vec{e}^{\,k}),\vec{v}^{\,k+1})\\ &= ((B_s-\frac{1}{2}A)\vec{v}^{\,k+1},\vec{v}^{\,k+1})+\frac{1}{2}(A(\vec{e}^{\,k+1}+\vec{e}^{\,k}),\vec{e}^{\,k+1}-\vec{e}^{\,k})\\ &= ((B_s-\frac{1}{2}A)\vec{v}^{\,k+1},\vec{v}^{\,k+1})+\frac{1}{2}(\left\|\vec{e}^{\,k+1}\right\|_A^2-\left\|\vec{e}^{\,k}\right\|_A^2)\end{aligned}$$

By assumption, $Q:=B_s-\frac{1}{2}A=\frac{M+M^T-A}{2}>0$, i.e. there exists $m>0$, s.t.

$$(Q\vec{u},\vec{u})\geq m\left\|\vec{u}\right\|_2^2.$$

Therefore,

$$m\left\|\vec{v}^{k+1}\right\|_2^2 + \frac{1}{2}(\left\|\vec{e}^{k+1}\right\|_A^2 - \left\|\vec{e}^{k}\right\|_A^2) \leq 0.$$

i.e.

$$2m\left\|\vec{v}^{k+1}\right\|_2^2 + \left\|\vec{e}^{k+1}\right\|_A^2 \leq \left\|\vec{e}^{k}\right\|_A^2.$$

Hence

$$\left\|\vec{e}^{k+1}\right\|_A^2 \leq \left\|\vec{e}^{k}\right\|_A^2.$$

and

$$\left\|\vec{e}^{k+1}\right\|_A^2 \to 0.$$

□

Theorem 1.38. *Suppose that A is HPD, then the Forward, Backward and Symmetric Gauss-Seidel Method converge.*

Proof. In the language in Theorem.1.37.

1. For Forward Gauss-Seidel Method,

$$M = B_{GS} = D - L.$$

So, we have

$$M + M^H - A = D - L + (D - L)^H - A = D - L + D^H - U - A = D.$$

Since A is HPD, so D is HPD. Now we can apply Theorem.1.37.

2. For Backward Gauss-Seidel Method,

$$M = B_{BGS} = D - U.$$

So, we have

$$M + M^H - A = D - U + (D - U)^H - A = D - U + D^H - L - A = D.$$

Since A is HPD, so D is HPD. Now we can apply Theorem.1.37.

3. For Symmetric Gauss-Seidel Method, we have

$$M = B_{SGS} = (D - U)^{-1} D (D - L)^{-1}.$$

So, we have

$$M + M^H - A = B_{SGS} + B_{SGS}^H - A.$$

Since A and B_{SGS} are HPD, so $M + M^H - A$ is HPD. Now we can apply Theorem.1.37.

□

Definition 1.39. *(Regular Splittings)[3] Let A, M, N be three given matrices satisfying*

$$A = M - N \in \mathbb{C}^{n \times n}. \tag{1.21}$$

The pair of matrices M, N is a regular splitting of A, if M is nonsingular and M^{-1} and N are nonnegative .

Theorem 1.40. *(The spectral radius estimation of Regular Splittings[3]) Let M, N be a regular splitting of A. Then*

$$\rho(M^{-1}N) < 1$$

if only if A is nonsingular and A^{-1} is nonnegative.

Proof.

1. Define $G = M^{-1}N$, since $\rho(G) < 1$, then $I - G$ is nonsingular. And then $A = M(I - G)$, so A is nonsingular. So, by Theorem.0.7 satisfied, since $G = M^{-1}N$ is nonsingular and $\rho(G) < 1$, then we have $(I - G)^{-1}$ is nonnegative as is $A^{-1} = (I - G)^{-1}M^{-1}$.
2. $\Leftarrow$: since A,M are nonsingular and A^{-1} is nonnegative, then $A = M(I - G)$ is nonsingular. Moreover

$$\begin{aligned} A^{-1}N &= \left(M(I - M^{-1}N)\right)^{-1} N \\ &= (I - M^{-1}N)^{-1}M^{-1}N \\ &= (I - G)^{-1}G. \end{aligned}$$

Clearly, $G = M^{-1}N$ is nonnegative by the assumptions, and as a result of the Perron-Frobenius theorem, there is a nonnegative eigenvector x associated with $\rho(G)$ which is an eigenvalue, such that

$$Gx = \rho(G)x.$$

Therefore

$$A^{-1}Nx = \frac{\rho(G)}{1 - \rho(G)}x.$$

Since x and $A^{-1}N$ are nonnegative, this shows that

$$\frac{\rho(G)}{1 - \rho(G)} \geq 0.$$

and this can be true only when $0 \leq \rho(G) \leq 1$. Since $I - G$ is nonsingular, then $\rho(G) \neq 1$, which implies that $\rho(G) < 1$.

□

2 The Two-Grid Algorithm

Definition 2.1. *(Two-Grid Algorithm Components)* *Suppose* $n_0, n_1 \in \mathbb{Z}$ *and* $n_1 \geq n_0 > 0$. *Suppose* $A_1 \in \mathbb{R}^{n_1 \times n_1}$ *is SPD and* $\vec{f} \in \mathbb{R}^{n_1}$, A_1 *is called the fine grid operator (matrix). Define some*

$$R_0 \in \mathbb{R}^{n_0 \times n_1}, \quad R_0 \vec{v}_1 \in \mathbb{R}^{n_0}, \quad \forall \vec{v}_1 \in \mathbb{R}^{n_1}.$$

R_0 *is called a restriction operator (matrix). We assume that* R_0 *has full rank* $rank(R_0) = n_0$ *and*

$$A_0 = \underbrace{R_0}_{n_0 \times n_1} \underbrace{A_1}_{n_1 \times n_1} \underbrace{R_0^T}_{n_1 \times n_0} \in \mathbb{R}^{n_0 \times n_0}.$$

A_0 *is called the coarse grid operator (matrix). Define*

$$P_0 = R_0^T \in \mathbb{R}^{n_1 \times n_0}.$$

Then, as Figure.3 described

$$A_0 = R_0 A_1 P_0.$$

P_0 *is called the Prolongation Operator (matrix).*

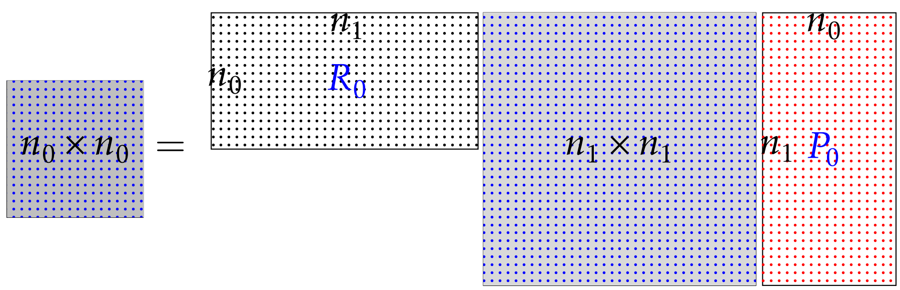


Figure 3: The action of Prolongation and Restriction Operator.

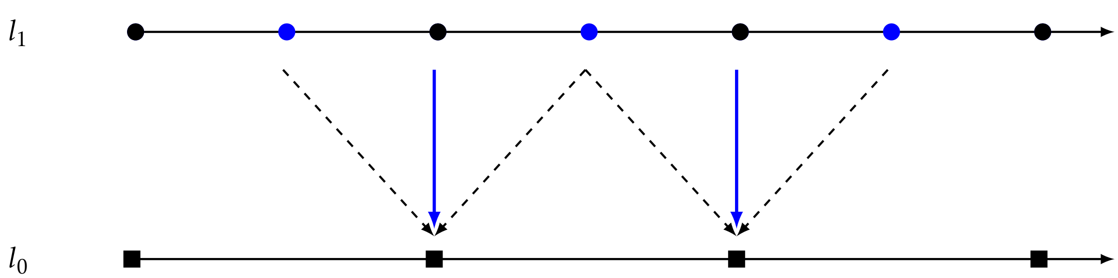


Figure 4: Restriction from fine grid to coarse grid on 1D uniform mesh.

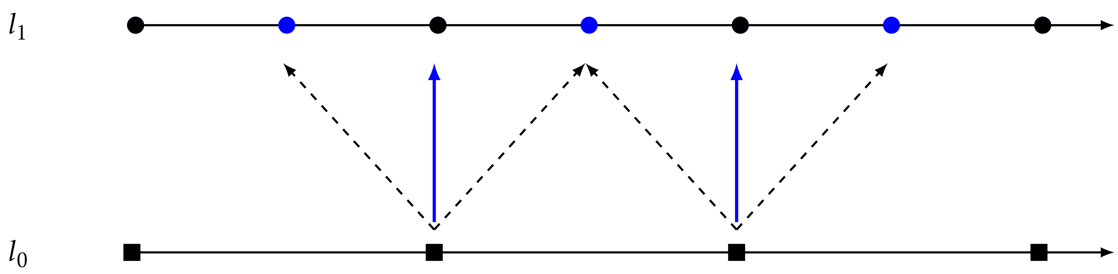


Figure 5: Prolongation from coarse grid to fine grid on 1D uniform mesh.

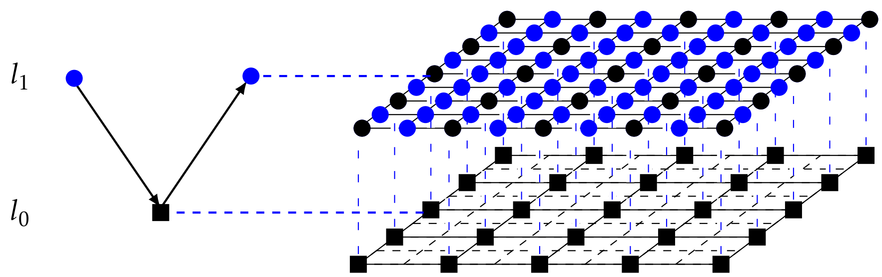


Figure 6: Mesh on the fine and coarse grid of the two-grid method in 2D.

Proposition 2.2. *$A_0 \in \mathbb{R}^{n_0 \times n_0}$ is SPD.*

Proof. 1. Symmetry-ness: Since $A_0 = R_0 A_1 R_0^T$ and A_1 is SPD, then

$$A_0^T = R_0 A_1^T R_0^T = R_0 A_1 R_0^T = A_0.$$

2. Positive Definite-ness: Let $\vec{\omega} \in \mathbb{R}^{n_0}$ an arbitrary vector and $\vec{\omega} \neq 0$. Since R_0 has full rank, then $R_0^T \vec{\omega} \neq 0$. Moreover, since A_1 is SPD,

$$\vec{\omega}^T A_0 \vec{\omega} = \vec{\omega}^T R_0 A_1 R_0^T \vec{\omega} = \left(R_0^T \vec{\omega}\right)^T A_1 \left(R_0^T \vec{\omega}\right) > 0.$$

□

We wish to solve the following: find $\vec{u}^{\,1} \in \mathbb{R}^{n_1}$ such that

$$A_1 \vec{u}_1 = \vec{f}_1.$$

Algorithm 2.3. *Given $\vec{u}_1^{\,k} \in \mathbb{R}^n$, compute*

$$\vec{u}_1^{\,k+1} := TG\left(\vec{f}_1, \vec{u}_1^{\,k}, m_1, m_2\right)$$

1. *Pre-smoothing (Figure.7-Figure.11):*
   - $\vec{u}_1^{\,(1,0)} := \vec{u}_1^{\,k}$
   - $\vec{u}_1^{\,(1,\sigma+1)} := \vec{u}_1^{\,(1,\sigma)} + S_1\left(\vec{f}_1 - A_1\vec{u}_1^{\,(1,\sigma)}\right),\ 1 \le \sigma \le m_1 - 1$
   - $\vec{u}_1^{\,(1)} := \vec{u}_1^{\,(1,m_1)}$
2. *Coarse-Grid correction (Figure.7-Figure.11):*
   - $\vec{r}_1^{\,(1)} := \vec{f}_1 - A_1\vec{u}_1^{\,(1)}$
   - $\vec{r}_0^{\,(1)} := R_0\vec{r}_1^{\,(1)} = R_0\left(\vec{f}_1 - A_1\vec{u}_1^{\,(1)}\right)$
   - $\vec{q}_0^{\,(1)} := A_0^{-1}\vec{r}_0^{\,(1)}$
   - $\vec{u}_1^{\,(2)} := \vec{u}_1^{\,(1)} + R_0^T\vec{q}_0^{\,(1)}$
3. *Post-smoothing (Figure.7-Figure.11):*
   - $\vec{u}_1^{\,(3,0)} := \vec{u}_1^{\,(2)}$
   - $\vec{u}_1^{\,(3,\sigma+1)} := \vec{u}_1^{\,(3,\sigma)} + S_1^T(\vec{f}_1 - A_1\vec{u}_1^{\,(3,\sigma)}),\ 0 \le \sigma \le m_2 - 1$
   - $\vec{u}_1^{\,(3)} := \vec{u}_1^{\,(3,m_2)}$
   - $\vec{u}_1^{\,k+1} := \vec{u}_1^{\,(3)}$

Remark 2.4. *Let us examine the Coarse-Grid Correction in a very special suppose $n_0 = n_1$ and*

$$R_0 = I_1, \quad (n_1 \times n_1 \text{ Identity}).$$

Then

$$A_0 = R_0 A_1 P_0 = R_0 A_1 R_0^T.$$

Coarse-Grid Correction:

- $\vec{r}_1^{\,(1)} := \vec{f}_1 - A_1\vec{u}_1^{\,(1)}$
- $\vec{r}_0^{\,(1)} := R_0\vec{r}_1^{\,(1)} = \vec{r}_1^{\,(1)}$
- $\vec{q}_0^{\,(1)} := A_0^{-1}\vec{r}_0^{\,(1)} = A_0^{-1}\vec{r}_1^{\,(1)} = \vec{e}_1^{\,(1)}$
- $$\begin{aligned}\vec{u}_1^{\,(2)} &:= \vec{u}_1^{\,(1)} + R_0^T\vec{q}_0^{\,(1)} \\ &= \vec{u}_1^{\,(1)} + R_0^T\vec{e}_1^{\,(1)} \\ &= \vec{u}_1^{\,(1)} + \vec{e}_1^{\,(1)} \\ &= \vec{u}_1^{\,(1)} + \vec{u}_1 - \vec{u}_1^{\,(1)} = \vec{u}_1\end{aligned}$$

Where

$$A_1\vec{u}_1 = \vec{f}_1.$$

Of course, the Algorithm.2.3 should terminated at this stage, because we have the exact solution.

Now, in general

$$\vec{q}_0^{\,(1)} = \vec{e}_1^{\,(1)}.$$

But

$$R_0^T \vec{q}_0^{\,(1)} \approx \vec{e}_1^{\,(1)}.$$

can be a quite good approximation in some sense. We will revisit this issue later.

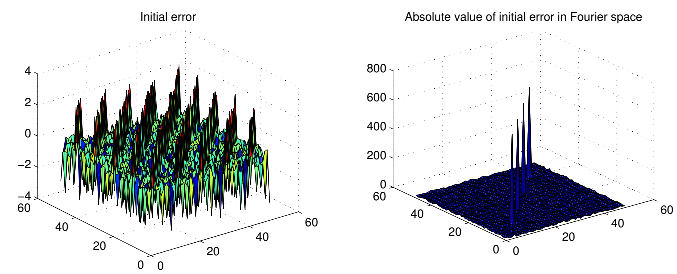


Figure 7: The initial error and the absolute value of initial error in Fourier space.

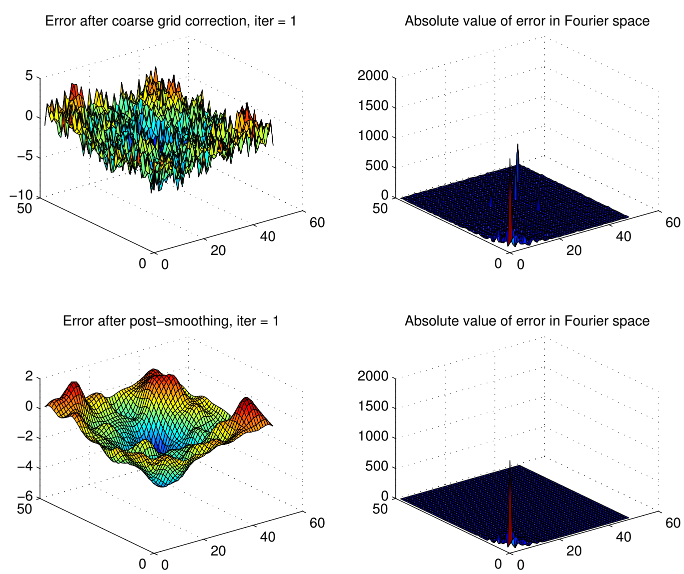


Figure 8: The error after 1_{st} coarse grid correction and post-smoothing with the absolute value of error in Fourier space.

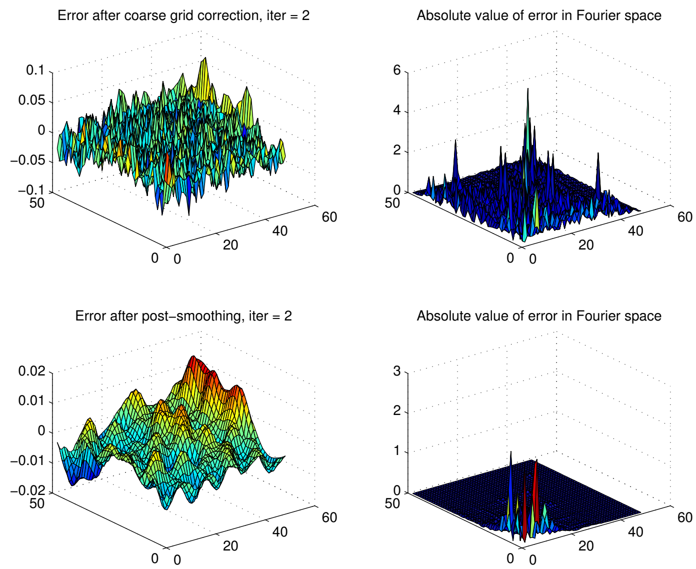


Figure 9: The error after 2_{nd} coarse grid correction and post-smoothing with the absolute value of error in Fourier space.

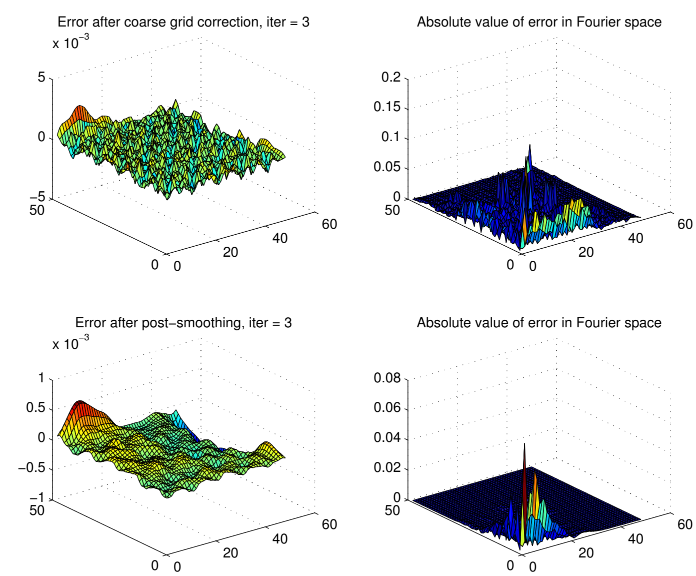


Figure 10: The error after 3_{rd} coarse grid correction and post-smoothing with the absolute value of error in Fourier space.

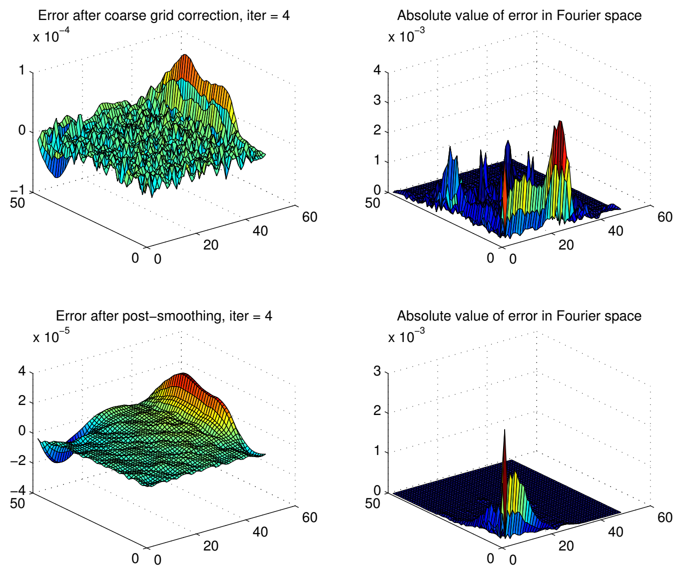


Figure 11: The error after 4_{th} coarse grid correction and post-smoothing with the absolute value of error in Fourier space.

From now, let us show that the two-grid method is a GLIS and also calculate the error propagation matrix, we need the following definitions first.

Definition 2.5. *(Coarse-Grid Ritz Projection matrix and Galerkin Condition) Referring to the definitions in Definition. (2.1), we define the matrix*

$$\tilde{\Pi}_1 = R_0^T \Pi_0 \in \mathbb{R}^{n_1 \times n_1},$$

where

$$\Pi_0 = \underbrace{A_0^{-1}}_{n_0 \times n_0} \underbrace{R_0}_{n_0 \times n_1} \underbrace{A_1}_{n_1 \times n_1} \in \mathbb{R}^{n_0 \times n_1}.$$

$\tilde{\Pi}_1$ is called the Coarse-Grid Ritz Projection matrix . We say that the coarse-grid matri, A_0, satisfies the Galerkin Condition iff

$$A_0 = R_0 A_1 R_0^T.$$

Proposition 2.6. *$\tilde{\Pi}_1$ as defined in Defintion.2.6 us a "bona fides" projection matrix, i.e.*

$$\left(\tilde{\Pi}_1\right)^2 = \tilde{\Pi}_1,$$

provided A_0 satisfies the Galerkin Condition.

Proof.

$$\left(\tilde{\Pi}_1\right)^2 = R_0^T \Pi_0 R_0^T \Pi_0$$

$$
\begin{aligned}
&= R_0^T A_0^{-1} \underbrace{R_0 A_1 R_0^T}_{A_0} A_0^{-1} R_0 A_1 \\
&= R_0^T A_0^{-1} R_0 A_1 \\
&= R_0^T \Pi_0 = \tilde{\Pi}_1.
\end{aligned}
$$

□

Corollary 2.7. *If the Galerkin Condition holds $I - \tilde{\Pi}_1$ is a projection matrix, i.e.*

$$\left(I_1 - \tilde{\Pi}_1\right)^2 = I_1 - \tilde{\Pi}_1,$$

Definition 2.8. *(Smoothing error matrix) Define the smoothing error matrix as*

$$K_1 = I_1 - S_1 A_1.$$

Definition 2.9. *(Adjoint) Let $M \in \mathbb{R}^{n_1 \times n_1}$ and $\vec{u}_1, \vec{v}_1 \in \mathbb{R}^{n_1}$. Define the adjoint with respect to the inner product*

$$(\vec{u}_1, \vec{v}_1)_1 := \vec{u}_1^T \vec{v}_1,$$

via

$$(M\vec{u}_1, \vec{v}_1)_1 := \left(\vec{u}_1, M^T \vec{v}_1\right)_1.$$

Define the adjoint with respect to the inner product

$$(\vec{u}_1, \vec{v}_1)_{A_1} := (\vec{u}_1, A_1 \vec{v}_1)_1,$$

via

$$(M\vec{u}_1, \vec{v}_1)_{A_1} := (\vec{u}_1, M^* \vec{v}_1)_{A_1}.$$

It follows by a simple calculation that

$$K_1{}^* = I_1 - S_1^T A_1.$$

In other words, for any $\vec{u}_1, \vec{v}_1 \in \mathbb{R}^{n_1}$,

$$
\begin{aligned}
((I_1 - S_1 A_1)\vec{u}_1, \vec{v}_1)_{A_1} &= ((I_1 - S_1 A_1)\vec{u}_1, A_1 \vec{v}_1)_1 \\
&= \left(\vec{u}_1, (I_1 - S_1 A_1)^T A_1 \vec{v}_1\right)_1 \\
&= \left(\vec{u}_1, \left(I_1 - A_1^T S_1^T\right) A_1 \vec{v}_1\right)_1 \\
&= \left(\vec{u}_1, \left(I_1 - A_1 S_1^T\right) A_1 \vec{v}_1\right)_1 \\
&= \left(\vec{u}_1, \left(A_1 - A_1 S_1^T A_1\right) \vec{v}_1\right)_1 \\
&= \left(\vec{u}_1, A_1 \left(I_1 - S_1^T A_1\right) \vec{v}_1\right)_1 \\
&= \left(\vec{u}_1, \left(I_1 - S_1^T A_1\right) \vec{v}_1\right)_{A_1}.
\end{aligned}
$$

Theorem 2.10. *For the two-grid Algorithm.2.3, we have*

$$\vec{e}_1^{\,k+1} = E_1 \vec{e}_1^{\,k},$$

where

$$E_1 = (K_1^*)^{m_2}\left(I_1 - \tilde{\Pi}_1\right)(K_1)^{m_2},$$

and

$$\vec{e}_1^{\,k} := \vec{u}_1^{\,E} - \vec{u}_1^{\,k}$$

where $\vec{u}_1^{\,E}$ is the exact solution. Furthermore, the two-grid method is a GLIS, i.e., there is some $B_1 \in \mathbb{R}^{n_1 \times n_1}$ such that

$$E_1 = I_1 - B_1 A_1.$$

Proof. Set

$$\vec{e}_1^{\,(i)} := \vec{u}_1^{\,E} - \vec{u}_1^{\,(i)}, \quad i = 1,2,3.$$

Then

$$\vec{e}_1^{\,(1)} := (I_1 - S_1 A_1)^{m_1} \vec{e}_1^{\,k} = K_1^{m_1} \vec{e}_1^{\,k}.$$

after the pre-smoothing step. Next we have the coarse-grid correction step.

$$\begin{aligned}
\vec{e}_1^{\,(2)} &= \vec{u}_1^{\,E} - \vec{u}_1^{\,(2)} \\
&= \vec{u}_1^{\,E} - \vec{u}_1^{\,(1)} - R_0^T \vec{q}_0^{\,(1)} \\
&= \vec{e}_1^{\,(1)} - R_0^T A_0^{-1} \vec{r}_0^{\,(1)} \\
&= \vec{e}_1^{\,(1)} - R_0^T A_0^{-1} R_0 \vec{r}_1^{\,(1)} \\
&= \vec{e}_1^{\,(1)} - R_0^T A_0^{-1} R_0 A_1 \vec{e}_1^{\,(1)} \\
&= \left(I_1 - \tilde{\Pi}_1\right) \vec{e}_1^{\,(1)}.
\end{aligned}$$

So

$$\vec{e}_1^{\,(2)} = \left(I_1 - \tilde{\Pi}_1\right) K_1^{m_1} \vec{e}_1^{\,k}.$$

Finally, from the post-smoothing step

$$\vec{e}_1^{\,k+1} = \vec{e}_1^{\,(3)} = \left(I_1 - S_1^T A_1\right)^{m_2} \vec{e}_1^{\,(2)}.$$

So

$$\vec{e}_1^{\,k+1} = (K_1{}^*)^{m_2}\left(I_1 - \tilde{\Pi}_1\right) K_1^{m_1} \vec{e}_1^{\,k},$$

as desired.

We leave the second part as an exercise. □

Remark 2.11. *In the case that $m_1 = m_2 = 1$, we have the following simple form for B_1:*

$$B_1 = \underbrace{S_1^T\left(S_1^{-T} + S_1^{-1} + A_1\right)S_1}_{\text{symmetric smoothing}} + \underbrace{K_1{}^* R_0^T A_0^{-1} R_0 K_1}_{\text{error correction}}.$$

In general, we can show that B_1 is symmetric with respect to

$$(\vec{u}_1, \vec{v}_1)_1 = \vec{u}_1^T \vec{v}_1,$$

iff $m_1 = m_2 = m$.

Theorem 2.12. *The error propagation matrix for the two-grid method, E_1, is symmetric with respect to $(\cdot,\cdot)_{A_1}$ iff $m_1 = m_2 = m$. Futhermore, if $m_1 = m_2 = m$ and if the Galerkin condition holds for A_0, then E_1 is Symmetric Positive Semi-Definite (SPSD) with respect to $(\cdot,\cdot)_{A_1}$.*

Proof. Let $\vec{u}_1, \vec{v}_1 \in \mathbb{R}^{n_1}$ be arbitrary. Then we have

$$\begin{aligned}
(E_1\vec{u}_1, \vec{v}_1)_{A_1} &= \left((K_1{}^*)^{m_2}\left(I_1 - \tilde{\Pi}_1\right)K_1^{m_1}\vec{u}_1, \vec{v}_1\right)_{A_1} \\
&= \left(\left(I_1 - \tilde{\Pi}_1\right)K_1^{m_1}\vec{u}_1, K_1^{m_2}\vec{v}_1\right)_{A_1} \\
&= \left(K_1^{m_1}\vec{u}_1, \left(I_1 - \tilde{\Pi}_1\right)^* K_1^{m_2}\vec{v}_1\right)_{A_1} \\
&= \left(\vec{u}_1, (K_1^*)^{m_1}\left(I_1 - \tilde{\Pi}_1\right)^* K_1^{m_2}\vec{v}_1\right)_{A_1}.
\end{aligned}$$

Observe that

$$\left(I_1 - \tilde{\Pi}_1\right)^* = \left(I_1 - \tilde{\Pi}_1\right).$$

Indeed,

$$\left(\left(I_1 - \tilde{\Pi}_1\right)\vec{u}_1, \vec{v}_1\right)_{A_1} = (\vec{u}_1, \vec{v}_1)_{A_1} - \left(\tilde{\Pi}_1\vec{u}_1, \vec{v}_1\right)_{A_1}$$

and

$$\begin{aligned}
\left(\tilde{\Pi}_1\vec{u}_1, \vec{v}_1\right)_{A_1} &= \left(R_0^T A_0^{-1} R_0 A_1\vec{u}_1, A_1\vec{v}_1\right)_1 \\
&= \left(A_1\vec{u}_1, \left(R_0^T A_0^{-1} R_0\right)^T A_1\vec{v}_1\right)_1 \\
&= \left(A_1\vec{u}_1, R_0^T A_0^{-1} R_0 A_1\vec{v}_1\right)_1 \\
&= \left(A_1\vec{u}_1, \tilde{\Pi}_1\vec{v}_1\right)_1 \\
&= \left(\vec{u}_1, \tilde{\Pi}_1\vec{v}_1\right)_{A_1}.
\end{aligned}$$

Therefore

$$\tilde{\Pi}_1^* = \tilde{\Pi}_1,$$

and

$$\left(I_1 - \tilde{\Pi}_1\right)^* = \left(I_1 - \tilde{\Pi}_1\right).$$

The symmetry of E_1 follows iff $m_1 = m_2$.

PSD-ness: let $\vec{u}_1 \in \mathbb{R}^{n_1}$ be arbitrary and suppose $m_1 = m_2 = m$. Since the Galerkin Condition holds, from Corollary.2.7, we have

$$\left(I_1 - \tilde{\Pi}_1\right)^2 = \left(I_1 - \tilde{\Pi}_1\right).$$

Now

$$\begin{aligned}
(E_1\vec{u}_1, \vec{u}_1)_{A_1} &= \left((K_1{}^*)^m \left(I_1 - \tilde{\Pi}_1\right) K_1^m \vec{u}_1, \vec{u}_1\right)_{A_1} \\
&= \left(\left(I_1 - \tilde{\Pi}_1\right) K_1^m \vec{u}_1, K_1^m \vec{u}_1\right)_{A_1} \\
&= \left(\left(I_1 - \tilde{\Pi}_1\right)^2 K_1^m \vec{u}_1, K_1^m \vec{u}_1\right)_{A_1} \\
&= \left(\left(I_1 - \tilde{\Pi}_1\right) K_1^m \vec{u}_1, \left(I_1 - \tilde{\Pi}_1\right)^* K_1^m \vec{u}_1\right)_{A_1} \\
&= \left(\left(I_1 - \tilde{\Pi}_1\right) K_1^m \vec{u}_1, \left(I_1 - \tilde{\Pi}_1\right) K_1^m \vec{u}_1\right)_{A_1} \\
&= \left\|\left(I_1 - \tilde{\Pi}_1\right) K_1^m \vec{u}_1\right\|_{A_1}^2 \geq 0.
\end{aligned}$$

□

Remark 2.13. *We remark that the Galerkin Condition is not required for the symmetry of E_1. All that was used to establish symmetry was the definition*

$$R_0^T A_0^{-1} R_0 A_1.$$

On the other hand, we used the Galerkin condition to establish that $\tilde{\Pi}_1$ and $\left(I_1 - \tilde{\Pi}_1\right)$ are projections and these fact played a role in showing that E_1 is SPSD.

It may be that the Galerkin Condition will fail in some multigrid application even where the definition of $\tilde{\Pi}_1$ above remains true.

3 Rigorous (Quantitative) Fourier Analysis of the Two-Grid Method

We will consider the Model problem:

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

Finite Difference Method:

Set

$$h := \frac{1}{n_1 + 1}, \quad n_1 \in \mathbb{Z}^+, \quad x_{1,i} = i \cdot h, \quad i = 0, 1, 2, \cdots, n_1 + 1.$$

Find $u_{1,1}, u_{1,2}, \cdots, u_{1,n_1}$, such that

$$\begin{cases} \frac{-u_{1,i-1} + 2u_{1,i} - u_{1,i+1}}{h^2} &= f(x_{1,i}) := f_{1,i}, \\ u_{1,0} = u_{1,n_1+1} &= 0. \end{cases} \tag{3.1}$$

Now set

$$\vec{u}_1^{\,FD} = \begin{bmatrix} u_{1,1} \\ u_{1,2} \\ \vdots \\ u_{1,n_1-1} \\ u_{1,n_1} \end{bmatrix}, \qquad \vec{f}_1^{\,FD} = \begin{bmatrix} hf_{1,1} \\ hf_{1,2} \\ \vdots \\ hf_{1,n_1-1} \\ hf_{1,n_1} \end{bmatrix} \in \mathbb{R}^{n_1},$$

and

$$A_1 := A_h = \begin{bmatrix} \frac{2}{h} & -\frac{1}{h} & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} \\ & & & -\frac{1}{h} & \frac{2}{h} \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}.$$

Then in matrix form, the Finite Difference Approximation is as follows: Find $\vec{u}_1^{FD} \in \mathbb{R}^{n_1}$, such that

$$A_1 \vec{u}_1^{FD} = \vec{f}_1^{FD}. \tag{3.2}$$

Finite Element Method:

We will use piecewise linear finite element weak formula: find $u \in H_0^1(0,1)$, such that

$$\left(\frac{du}{dx}, \frac{dv}{dx}\right)_{L^2(0,1)} = (f, v)_{L^2(0,1)} \quad \forall\, v \in H_0^1(0,1).$$

Define

$$V_h := \left\{ v \in C_0^0([0,1]) | \ v|_{T_{1,i}} \in \mathbb{P}_1(T_{1,i}), \ 1 \le i \le n_1 + 1 \right\}.$$

The n_1 hat functions form a basis function V_h.

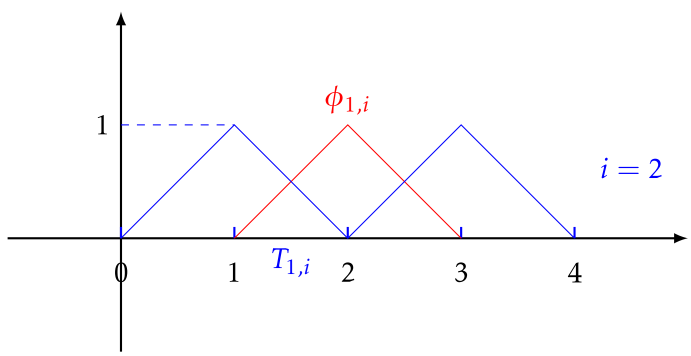


Figure 12: Piecewise linear basis function in 1D.

Then finite element approximation of the model problem is as follows: find $u_h V_h$, such that

$$\left(\frac{du_h}{dx}, \frac{d\phi_{1,i}}{dx}\right)_{L^2(0,1)} = (f, \phi_{1,i})_{L^2(0,1)},$$

for all $i = 1, 2, \cdots, n_1$.

Now set

$$\vec{f}_1^{FE} = \begin{bmatrix} (f, \phi_{1,1})_{L^2(0,1)} \\ \vdots \\ \vdots \\ \left(f, \phi_{1,n_1}\right)_{L^2(0,1)} \end{bmatrix} \in \mathbb{R}^{n_1}.$$

We expand u_h in the basis of hat functions.

$$u_h = \sum_{j=1}^{n_1} u_{1,j}\phi_{1,j}(x) \quad \in V_h.$$

Then, we set

$$\vec{u}_1^{\,FE} = \begin{bmatrix} u_{1,1} \\ \vdots \\ \vdots \\ u_{1,n_1} \end{bmatrix} \in \mathbb{R}^{n_1}.$$

Define the $n_1 \times n_1$ stiffness matrix A, via,

$$a_{1,i,j} := [A_1]_{ij} = \left(\frac{d\phi_{1,i}}{dx}, \frac{d\phi_{1,j}}{dx}\right)_{L^2(0,1)} = \left(\frac{d\phi_{1,i}}{dx}, \frac{d\phi_{1,j}}{dx}\right)_{L^2(0,1)}.$$

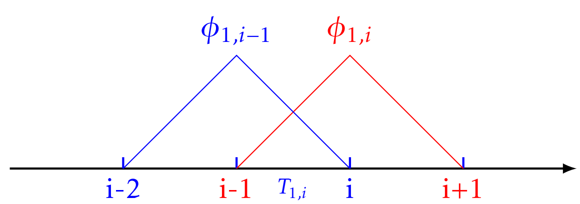


Figure 13: Piecewise linear basis function ϕ_i in 1D.

So,

$$\begin{aligned} a_{1,i,i-1} &= \left(\frac{d\phi_{1,i-1}}{dx}, \frac{d\phi_{1,i}}{dx}\right)_{L^2(0,1)} \\ &= \left(\frac{d\phi_{1,i-1}}{dx}, \frac{d\phi_{1,j}}{dx}\right)_{L^2(T_{1,i})} \\ &= \left(-\frac{1}{h}, \frac{1}{h}\right)_{L^2(T_{1,i})} \\ &= -\frac{1}{h}. \end{aligned}$$

Likewise,

$$a_{1,i,i} = \left(\frac{d\phi_{1,i}}{dx}, \frac{d\phi_{1,i}}{dx}\right)_{L^2(0,1)} = \frac{2}{h}.$$

Let, we have, as before,

$$A_1 := A_h = \begin{bmatrix} \frac{2}{h} & -\frac{1}{h} & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} \\ & & & -\frac{1}{h} & \frac{2}{h} \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}.$$

Therefore Finite Element Approximation matrix, is as follows: Find $\vec{u}_1^{FE} \in \mathbb{R}^{n_1}$, such that

$$A_1 \vec{u}_1^{FE} = \vec{f}_1^{FE}. \tag{3.3}$$

Theorem 3.1. *$A_1 \in \mathbb{R}^{n_1 \times n_1}$ is SPD. It's eigenvalue are*

$$\lambda_1^{(k)} = \frac{4}{h} \sin^2\left(\frac{k\pi h}{2}\right) = \frac{2}{h}(1 - \cos k\pi h). \tag{3.4}$$

for $k = 1, 2, \cdots, n_1$, and the corresponding eigenvectors are

$$\left[\vec{v}_1^{(k)}\right]_i = v_{1,i}^{(k)} = \sin k\pi x_{1,i}. \tag{3.5}$$

Proof. Let $1 \le i \le n_1$,

$$\begin{aligned} \left[A_1 \vec{v}_1^{(k)}\right]_i &= -\frac{1}{h}\sin(k\pi x_{1,i-1}) + \frac{1}{h}\sin(k\pi x_{1,i}) - \frac{1}{h}\sin(k\pi x_{1,i+1}) \\ &= \frac{2}{h}[1 - cos(k\pi h)]\sin(k\pi x_{1,i}) \end{aligned}$$

Thus

$$\left[A_1 \vec{v}_1^{(k)}\right]_i = \lambda_1^{(k)} v_{1,i}^{(k)},$$

for $1 \le i \le n_1$. □

Theorem 3.2. *The spectral condition number of A_1, i.e.*

$$\kappa_2(A_1) := \|A_1\|_2 \left\|A_1^{-1}\right\|_2,$$

satisfies the estimates

$$C_1 h^{-2} \le \kappa_2(A_1) \le C_2 h^{-2},$$

for some constants $0 < C_1 \le C_2$.

Proof. Since A_1 is SPD, it follows that

$$\kappa_2(A_1) = \frac{\lambda_1^{(n_1)}}{\lambda_1^{(1)}} = \frac{1 - \cos(n_1 \pi h)}{1 - \cos(\pi h)}.$$

Consider the function $f(x) = 1 - cos(\pi x)$ (Figure.15).

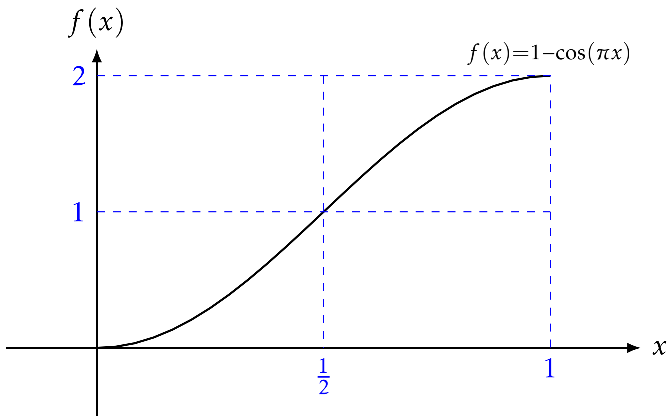


Figure 14: The value of $1-\cos(\pi x)$.

Now, observe that, if $n_1 \geq 1$, then

$$0 < h \leq \frac{1}{2}, \quad \left(0 < \pi h \leq \frac{\pi}{2}\right).$$

Since, we have $1 \leq 1-\cos(n_1\pi h) \leq 2$, it follows that

$$\frac{1}{1-\cos(\pi h)} \leq \kappa_2(A_1) \leq \frac{2}{1-\cos(\pi h)}.$$

Now, setting $x = \pi h$, we have, by Taylor Theorem (expanding at $x = 0$),

$$\cos(x) = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4\cos(\xi),$$

for some

$$0 < \xi < x = \pi h \leq \frac{\pi}{2}.$$

Hence

$$0 < \xi < \frac{\pi}{2},$$

and

$$0 < \cos(\xi) < 1.$$

So,

$$0 < \frac{1}{4!}x^4\cos(\xi) = \cos(x) - 1 + \frac{1}{2}x^2$$

or

$$1-\cos(\pi h) < \frac{\pi^2}{2}h^2, \quad 0 < \pi h \leq \frac{\pi}{2}.$$

We conclude the lower bound

$$\frac{2}{\pi^2 h^2} \leq \frac{1}{1-\cos(\pi h)}, \quad 0 < h \leq \frac{1}{2}.$$

The last bound may be viewed graphically as

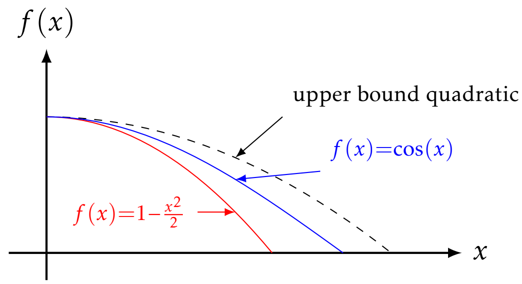


Figure 15: The upper and lower bound of $1-\cos(\pi x)$.

To go further, again by Taylor's Theorem, we have, for $0 < x \le \frac{\pi}{2}$,

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{6!}\cos(\xi),$$

for some $0 < \xi < \frac{\pi}{2}$. So

$$\cos(x) - 1 + \frac{x^2}{2} - \frac{x^4}{24} = -\frac{x^6}{6!}\cos(\xi) < 0,$$

and

$$\frac{x^2}{2} - \frac{x^4}{24} < 1 - \cos(x).$$

But

$$\frac{x^2}{2} - \frac{x^4}{24} \ge \frac{x^2}{4}, \quad 0 < x \le \frac{\pi}{2}.$$

Thus

$$\frac{x^2}{4} < 1 - \cos(x), \quad 0 < x \le \frac{\pi}{2}.$$

and

$$\frac{x^2}{2} - \frac{x^4}{24} < 1 - \cos(x), \quad 0 < x \le \frac{\pi}{2},$$

and

$$\frac{1}{1-\cos(\pi h)} < \frac{4}{\pi^2 h^2}, \quad 0 < \pi h \le \frac{\pi}{2}.$$

Thus

$$\frac{2}{\pi^2 h^2} \le \kappa_2(A_1) \le \frac{8}{\pi^2 h^2}.$$

□

3.1 The Damped Jacobi Smoother

To approximate the solution of

$$A_1 \vec{u}_1 = \vec{f}_1^{\square} \in \mathbb{R}^{n_1},$$

we apply the damped Jacobi method. This requires a splitting (1.8) of A_1, i.e.:

$$A_1 = D - U - L.$$

where

$$D = \begin{bmatrix} \frac{2}{h} & & & & \\ & \frac{2}{h} & & & \\ & & \ddots & & \\ & & & \frac{2}{h} & \\ & & & & \frac{2}{h} \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}, \quad U = \begin{bmatrix} 0 & -\frac{1}{h} & & & \\ & 0 & -\frac{1}{h} & & \\ & & \ddots & \ddots & \\ & & & 0 & -\frac{1}{h} \\ & & & & 0 \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}.$$

and $L = U^T$. The damped Jacobi method reads

$$\begin{cases} \vec{z}_1 & = D^{-1}(U + U^T)\vec{u}_1^{\,(\sigma)} + D^{-1}\vec{f}_1^{\,\square}, \\ \vec{u}_1^{\,(\sigma+1)} & = \omega \vec{z}_1 + (1-\omega)\vec{u}_1^{\,(\sigma)}, \end{cases} \tag{3.6}$$

where $0 < \omega \le 1$.

Theorem 3.3. *The eigenvectors of K_1 are the same as those for A_1, namely,*

$$\left[\vec{v}_1^{\,(k)}\right]_i = \vec{v}_{1,i}^{\,(k)} = \sin(k\pi x_{1,i})$$

for $k = 1, \cdots, n_1$. The eigenvalues of K_1 are

$$\mu_1^{(k)}(\omega) = \omega\cos(k\pi h) + 1 - \omega = 1 - 2\omega\sin^2\left(\frac{k\pi h}{2}\right), \quad 1 \le k \le n_1. \tag{3.7}$$

Proof.

$$\begin{aligned} K_1\vec{v}_1^{\,(k)} &= \vec{v}_1^{\,(k)} - \omega D^{-1}A_1\vec{v}_1^{\,(k)} \\ &= \vec{v}_1^{\,(k)} - \omega\frac{h}{2}\lambda_1^{(k)}\vec{v}_1^{\,(k)} \\ &= \left(1 - \omega\frac{h}{2}\frac{2}{h}(1-\cos(k\pi h))\right)\vec{v}_1^{\,(k)} \\ &= (1 - \omega(1-\cos(k\pi h)))\,\vec{v}_1^{\,(k)} \\ &= \left(1 - \omega 2\sin^2\left(\frac{k\pi h}{2}\right)\right)\vec{v}_1^{\,(k)} \end{aligned}$$

So,

$$\mu_1^{(k)}(\omega) = 1 - \omega + \omega\cos(k\pi h) = 1 - 2\omega\sin^2\left(\frac{k\pi h}{2}\right).$$

□

Remark 3.4. *In the multi grid setting, we want the Jacobi method (the smoother) to have a "smooth" effect on the error. In other words, we want to damper high-frequency modes of the error faster than those of low-frequency. Recall, for the damped Jacobi method*

$$\vec{e}_1^{\,(\sigma+1)} = K_1\vec{e}_1^{\,(\sigma)},$$

where

$$K_1 = I_1 - \frac{\omega h}{2}A_1.$$

Now, expand $\vec{e}_1^{\,(\sigma)}$ *in the basis of eigenvectors* $\left\{\vec{v}_1^{\,(k)}\right\}_{k=1}^{n_1}$*: There exist unique number*

$$\epsilon_k^{(\sigma)} \in \mathbb{R}, \;\; k = 1, 2 \cdots, n.$$

such that

$$\vec{e}_1^{\,(\sigma)} = \sum_{k=1}^{n_1} \epsilon_k^{(\sigma)} \vec{v}_1^{\,(k)}.$$

Then

$$\vec{e}_1^{\,(\sigma+1)} = \sum_{k=1}^{n_1} \mu_1^{(k)}(\omega) \epsilon_k^{(\sigma)} \vec{v}_1^{\,(k)}. \tag{3.8}$$

3.2 A Basic Mesh Assumption

We assume that $n_1 + 1 \geq 2$ is even. In this case, we can define

$$n_0 := \frac{n_1 + 1}{2} - 1.$$

For example, suppose $n_1 = 3$. Then

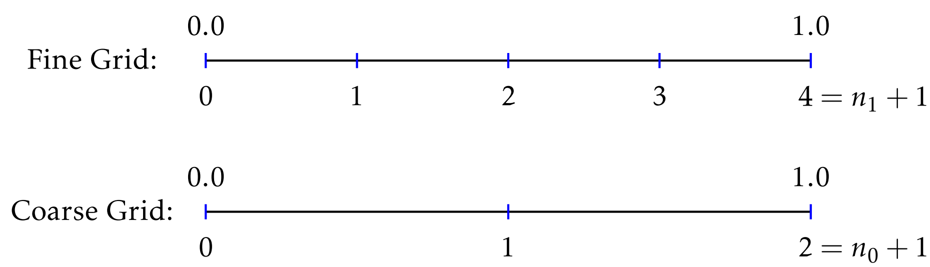


Figure 16: Fine and coarse grid for $n_1 = 3$.

Definition 3.5. (*High- and low- Frequency*) *Suppose* $n_1 + 1$ *is even, we say mode* $k, 1 \leq k \leq n_1$*, is of high frequency iff*

$$n_0 + 1 = \frac{n_1 + 1}{2} \leq k \leq n_1.$$

Otherwise, it is of low frequency.

Theorem 3.6. *The quantity*

$$S(\omega) = \max_{\frac{n_1+1}{2} \le k \le n_1} \left| \mu_1^{(k)}(\omega) \right|,$$

is minimized by

$$\omega = \omega_0 = \frac{2}{3},$$

in which case

$$\left| \mu_1^{(k)}(\omega_0) \right| \le \frac{1}{3} \tag{3.9}$$

for all $\frac{n_1+1}{2} \le k \le n_1$. *More generally, if* $0 < \omega \le 1$, *then*

$$\left| \mu_1^{(k)}(\omega) \right| < 1. \tag{3.10}$$

Proof. Proof by plots of $\mu_1^{(k)}(\omega) = 1 - 2\omega \sin^2\left(\frac{k\pi h}{2}\right)$(Figure. 17):

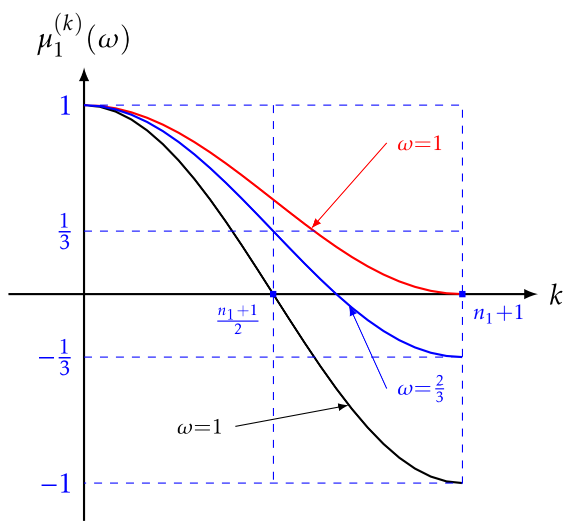


Figure 17: The eigenvalue of $\mu_1^{(k)}(\omega) = 1 - 2\omega \sin^2\left(\frac{k\pi h}{2}\right)$.

□

Remark 3.7. *Recall, with* $\omega = \omega_0$, *we have*

$$\vec{e}_1^{\,(\sigma+1)} = \sum_{k=1}^{n_1} \mu_1^{(k)}(\omega_0)\epsilon_k^{(\sigma)} \vec{v}_1^{\,(k)}.$$

Thus high-frequency modes will be damped faster than those of low-frequency. In fact the modes $\frac{n_1+1}{2} \le k \le n_1$ *will be reduced by at least* $\frac{1}{3}$ *after a single smoothing iteration.*

1^{st} Multigrid Principle: Many classical iterative methods have error smoothing property (High-frequency modes are damped more rapidly than those of low-frequency), but converge very slowly, especially as

$h \to 0$.

2^{nd} Multigrid Principle: Low-frequency information(modes) is well approximated on a coarse grid.

Theorem 3.8. *Let $K_1 = K(\omega) = I_1 - \omega D^{-1} A_1$ be the error propagation matrix for damped Jacobi method for the model problem (3.2). Then*

$$\rho(K_1) = \mu_1^{(1)}(\omega) = 1 - \mathcal{O}(h^2),$$

for all $0 < \omega \le 1$, i.e., there exist $0 < C_1 \le C_2$ such that

$$0 \le C_1 h^2 \le 1 - \rho(K_1) \le C_2 h^2.$$

Proof. Since $\mu_1^{(k)}(\omega) = 1 - 2\omega \sin^2\left(\frac{k\pi h}{2}\right)$, then

$$\rho(K_1) = \mu_1^{(1)}(\omega) = 1 - 2\omega \sin^2\left(\frac{\pi h}{2}\right).$$

By Taylor expansion, we have

$$\sin^2(x) = x^2 - \frac{1}{3}x^4 + \mathcal{O}(x^6).$$

Then

$$\rho(K_1) = \mu_1^{(1)}(\omega) = 1 - 2\omega\left(\frac{\pi^2}{4}h^2 - \frac{1}{3}\frac{\pi^4}{2^4}h^4 + \mathcal{O}(h^6)\right) = 1 - \mathcal{O}(h^2).$$

□

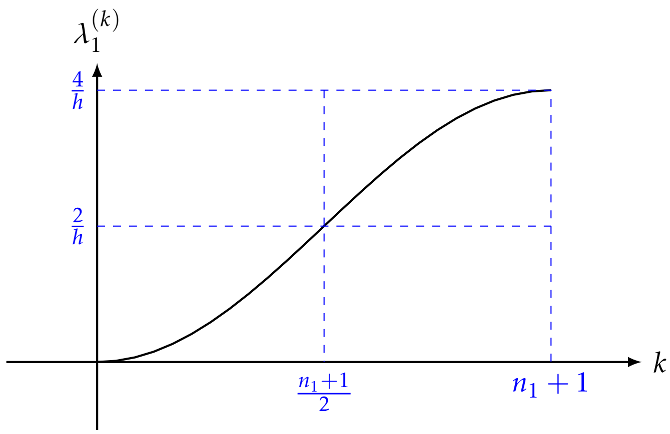
Remark 3.9. *We observed that a smoothed error is well approximated on a coarse grid. We can show that if the error is smooth, then the residual is also smooth.*

Recall (3.8) with $\omega = \omega_0$,

$$\vec{e}_1^{\,(\sigma+1)} = \sum_{k=1}^{n_1} \mu_1^{(k)}(\omega_0)\epsilon_k^{(\sigma)} \vec{v}_1^{\,(k)}.$$

Since $\vec{r}_1^{\,(\sigma+1)} = A_1 \vec{e}_1^{\,(\sigma+1)}$,

$$\vec{r}_1^{\,(\sigma+1)} = A_1 \vec{e}_1^{\,(\sigma+1)} = \sum_{k=1}^{n_1} \mu_1^{(k)}(\omega_0)\epsilon_k^{(\sigma)} A_1 \vec{v}_1^{\,(k)} = \sum_{k=1}^{n_1} \mu_1^{(k)}(\omega_0)\epsilon_k^{(\sigma)} \lambda_1^{(k)} \vec{v}_1^{\,(k)}. \tag{3.11}$$



$$\lambda_1^{(k)} = \frac{2}{h}(1-\cos(k\pi h)),$$
$$h = \frac{1}{n_1+1}.$$

Thus $\vec{r}_1^{\,(\sigma+1)}$ will be well approximated on the coarse grid iff $\vec{e}_1^{\,(\sigma+1)}$ has this property.

3.3 Prolongation and Restriction Operators

We first define the prolongation matrix P_0, then we set $R_0 = P_0^T$.

We approach this from the FEM point of view. Suppose $u_H \in V_H$ is piecewise linear.

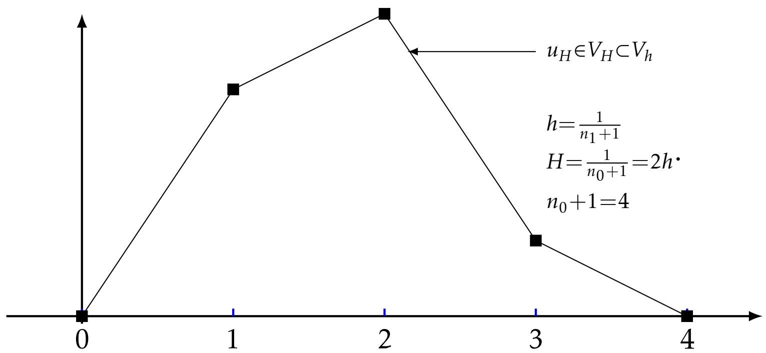


Figure 18: Coarse grid $n_0 = 3$ in 1D.

Recall, based on our basic mesh assumption

$$n_0 = \frac{n_1+1}{2} - 1 \Leftrightarrow n_1 = 2(n_0+1) - 1.$$

In this case $n_1 = 7$.

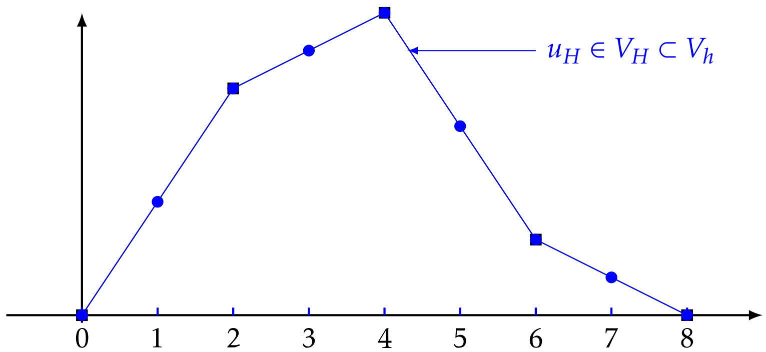


Figure 19: Fine grid $n_1 = 7$ in 1D.

Find a matrix that maps the 3 coarse grid DOF's into the 7 fine grid DOF's, i.e. find a matrix $P_0 \in \mathbb{R}^{n_1 \times n_0}$ such that

$$P_0 \vec{u}_0 = \vec{u}_1 \in \mathbb{R}^{n_1},$$

where $\vec{u}_0 \in \mathbb{R}^{n_0}$ is coordinate representation of $u_H = u_0 \in V_H = V_0$ and $\vec{u}_1 \in \mathbb{R}^{n_1}$ is the representation of $u_h = u_1 \in V_h = V_1$ such that

$$u_h = u_H. \quad \text{(same piece-wise linear function)}$$

For this example ($n_0 = 3, n_1 = 7$)

$$P_0 = \begin{bmatrix} \frac{1}{2} & & \\ 1 & & \\ \frac{1}{2} & \frac{1}{2} & \\ & 1 & \\ & \frac{1}{2} & \frac{1}{2} \\ & & 1 \\ & & \frac{1}{2} \end{bmatrix} \in \mathbb{R}^{7\times 3}.$$

Therefore,

$$R_0 = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} & & & & \\ & & \frac{1}{2} & 1 & \frac{1}{2} & & \\ & & & & \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix} \in \mathbb{R}^{3\times 7}.$$

Definition 3.10. *Suppose that the positive integers n_0 and n_1 satisfy*

$$n_1 = 2(n_0+1)-1,$$

let $\vec{u}_0 \in \mathbb{R}^{n_0}$. The action of $P_0 \in \mathbb{R}^{n_1\times n_0}$ on $\vec{u}_0$ is defined as follows

$$\begin{aligned}
[P_0\vec{u}_0]_1 : &= \frac{1}{2}u_{0,1}, \\
[P_0\vec{u}_0]_{n_1} : &= \frac{1}{2}u_{0,n_0}, \\
[P_0\vec{u}_0]_{2i} : &= u_{0,i}, \;\; 1 \le i \le n_0 \\
[P_0\vec{u}_0]_{2i+1} : &= \frac{1}{2}\left(u_{0,i}+u_{0,i+1}\right), \;\; 1 \le i \le n_0-1.
\end{aligned}$$

We define

$$R_0 = P_0^T \in \mathbb{R}^{n_0\times n_1}. \tag{3.12}$$

Theorem 3.11. *Let P_0, R_0 be defined as above, with A_1 defined as in (3.3). Suppose $A_0 \in \mathbb{R}^{n_0\times n_0}$ satisfies the Galerkin condition, i.e.*

$$A_0 := R_0A_1P_0.$$

Then

$$A_0 = \begin{bmatrix} \frac{2}{H} & -\frac{1}{H} & & & \\ -\frac{1}{H} & \frac{2}{H} & -\frac{1}{H} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{H} & \frac{2}{H} & -\frac{1}{H} \\ & & & -\frac{1}{H} & \frac{2}{H} \end{bmatrix} \in \mathbb{R}^{n_0\times n_0},$$

where

$$H = 2h = \frac{1}{n_0+1},$$

A_0 is clearly SPD and has the eigen-pairs

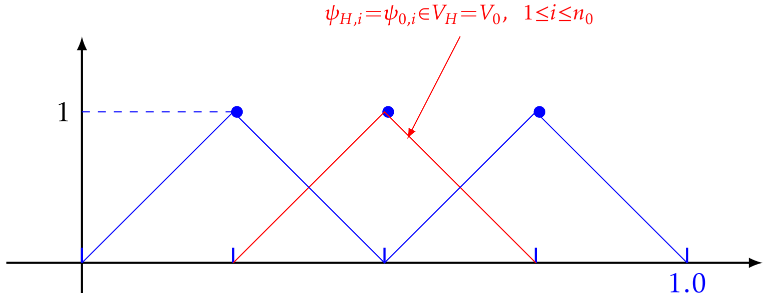
$$\begin{aligned}
\left[\vec{v}_0^{\,(k)}\right]_i &= \sin(k\pi x_{0,i}), \;\; 1 \le i \le n_0, \qquad (3.13)\\
\lambda_0^{(k)} &= \frac{2}{H}\left(1-\cos(k\pi H)\right),
\end{aligned}$$

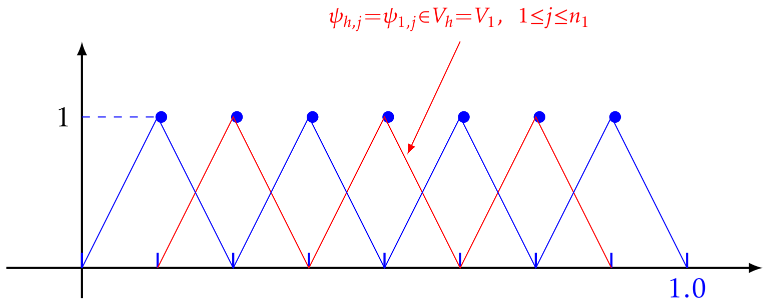
for $k = 1, \cdots, n_0$, where

$$x_{0,i} = iH, \;\; 0 \le i \le n_0+1.$$

Proof. Here we give a proof that works in much more general settings. It is based on the FEM picture using piecewise linear elements.

Consider the Lagrange nodel basis for the coarse space





Since $V_0 \subset V_1$, each $\psi_{0,i} \in V_1$. Therefore, since $B_1 = \left\{\psi_{1,j}\right\}_{j=1}^{n_1}$ is a basis for V_1, $\exists!$ numbers

$$p_{0,k,i} \in \mathbb{R}, \quad 1 \le k \le n_1,$$

such that

$$\psi_{0,i} = \sum_{k=1}^{n_1} p_{0,k,i} \psi_{1,k}. \tag{3.14}$$

Observe that, in present case where P_0 is as defined in Definition.3.10,

$$[P_0]_{k,i} = p_{0,k,i}.$$

In fact, we could have used (3.14) in Definition.3.10. For instance, it should be clear that

$$\psi_{0,1} = \frac{1}{2}\psi_{1,1} + \psi_{1,2} + \frac{1}{2}\psi_{1,3}.$$

Now, as usual in the FEM setting, we define stiffness matrices via

$$[A_1]_{i,j} := [A_h]_{i,j} = a\left(\psi_{h,j}, \psi_{h,i}\right) = a\left(\psi_{1,j}, \psi_{1,i}\right).$$

On the coarse grid, we proceed similarly

$$\begin{aligned}
[A_0]_{i,j} := [A_H]_{i,j} &= a\left(\psi_{H,j}, \psi_{H,i}\right) \\
&= a\left(\psi_{0,j}, \psi_{0,i}\right) \\
&\overset{(3.14)}{=} a\left(\sum_{k=1}^{n_1} p_{0,k,i}\psi_{1,k}, \sum_{\ell=1}^{n_1} p_{0,\ell,j}\psi_{1,\ell}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n_1}\sum_{\ell=1}^{n_1} p_{0,k,i}\, a\left(\psi_{1,k},\psi_{1,\ell}\right) p_{0,\ell,j} \qquad (3.15)\\
&= P_0^T A_1 P_0\\
&= R_0 A_1 P_0.
\end{aligned}
$$

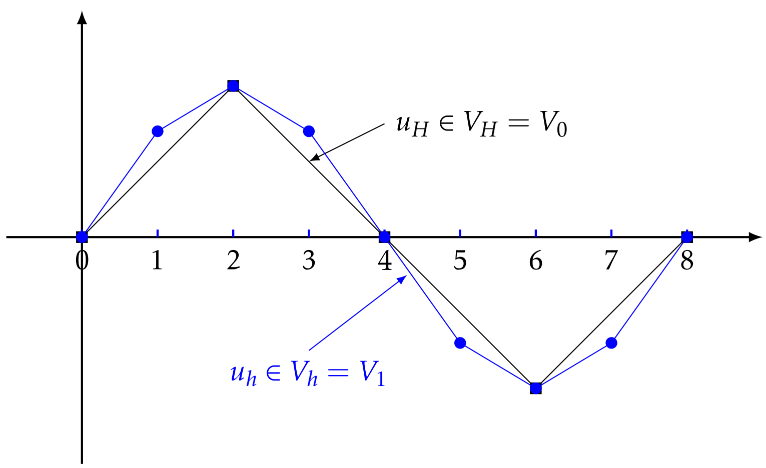
But this is the same as the matrix A_0 which is defined to satisfy the Galerkin condition, i.e.

$$
A_0 := R_0 A_1 P_0 \overset{(3.15)}{=} A_H := \begin{bmatrix} \frac{2}{H} & -\frac{1}{H} & & & \\ -\frac{1}{H} & \frac{2}{H} & -\frac{1}{H} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{H} & \frac{2}{H} & -\frac{1}{H} \\ & & & -\frac{1}{H} & \frac{2}{H} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0}.
$$

□

3.4 Restriction Operator

Consider the fine grid function



Suppose $\vec{u}_1 \in \mathbb{R}^{n_1}$ is the coordinate vector of $u_h \in V_h = V_1$ w.r.t the lagrange nodal basis B_1. What is the action of R_0 on $\vec{u}_1$?

$$
R_0 \vec{u}_1 = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} & & & & \\ & & \frac{1}{2} & 1 & \frac{1}{2} & & \\ & & & & \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{1,2} \\ u_{1,3} \\ u_{1,4} \\ u_{1,5} \\ u_{1,6} \\ u_{1,7} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}u_{1,1} + u_{1,2} + \frac{1}{2}u_{1,3} \\ \frac{1}{2}u_{1,3} + u_{1,4} + \frac{1}{2}u_{1,5} \\ \frac{1}{2}u_{1,5} + u_{1,6} + \frac{1}{2}u_{1,7}. \end{bmatrix}
$$

Definition 3.12. *The matrix $\frac{1}{2}R_0$, where R_0 is defined as above, is the matrix representing the full weighting restriction operator.*

Remark 3.13. *$\frac{1}{2}R_0$ gives a more geometrically meaningful result, its a proper weighted average.*

3.5 Quantitative Error Analysis

Recall the result of Theorem.2.10 concerning the two grid algorithm:

$$\vec{e}_1^{\,k+1} = E_1 \vec{e}_1^{\,k},$$

where

$$E_1 = (K_1^*)^{m_2}\left(I_1 - \tilde{\Pi}_1\right)(K_1)^{m_2}.$$

Recall that, for the present case,

$$K_1 = I_1 - \omega_0 \frac{h}{2} A_1.$$

$\tilde{\Pi}_1$ is the coarse grid Ritz projection matrix

$$\tilde{\Pi}_1 = R_0^T A_0^{-1} R_0 A_1.$$

Since our A_0 satisfies the Galerkin condition

$$\tilde{\Pi}_1^2 = \tilde{\Pi}_1.$$

3.6 Some Technical Results

Proposition 3.14. *With the basic mesh assumption in place, et cetera,*

$$R_0 \vec{v}_1^{\,(k)} = \begin{cases} 2\cos^2\left(\frac{k\pi h}{2}\right)\vec{v}_0^{\,(k)}, & 1 \le k \le n_0, \\ -2\sin^2\left(\frac{(n_1+1-k)k\pi h}{2}\right)\vec{v}_0^{\,(n_1+1-k)}, & n_0+1 \le k \le n_1. \end{cases} \tag{3.16}$$

Proof. Consider for $1 \le i \le n_0$, $1 \le k \le n_0$,

$$\begin{aligned}
\left[R_0 \vec{v}_1^{\,(k)}\right]_i &= \frac{1}{2}\left\{\sin(k\pi x_{1,2i-1}) + 2\sin(k\pi x_{1,2i}) + \sin(k\pi x_{1,2i+1})\right\} \\
&= \frac{1}{2}\left\{\sin(k\pi x_{1,2i} - k\pi h) + 2\sin(k\pi x_{1,2i}) + \sin(k\pi x_{1,2i} + k\pi h)\right\} \\
&= \frac{1}{2}\left\{\sin(k\pi x_{1,2i})\cos(k\pi h) - \cos(k\pi x_{1,2i})\sin(k\pi h)\right. \\
&+ 2\sin(k\pi x_{1,2i}) \\
&+ \left.\sin(k\pi x_{1,2i})\cos(k\pi h) + \cos(k\pi x_{1,2i})\sin(k\pi h)\right\} \\
&= \frac{1}{2}\left\{2\sin(k\pi x_{1,2i})\cos(k\pi h) + 2\sin(k\pi x_{1,2i})\right\} \\
&= \sin(k\pi x_{1,2i})(\cos(k\pi h) + 1) \\
&= 2\cos^2\left(\frac{k\pi h}{2}\right)\sin(k\pi x_{1,2i}) \\
&= 2\cos^2\left(\frac{k\pi h}{2}\right)\left[\vec{v}_0^{\,(k)}\right]_i, \quad 1 \le k \le n_0.
\end{aligned}$$

This complete the first part. For the second part, assume $n_0 + 1 \le k \le n_1$. Now

$$\left[\vec{v}_0^{\,(n_1+1-k)}\right]_i = \sin\left((n_1+1-k)\pi i H\right)$$

$$\begin{aligned}
&= \sin\left((n_1+1-k)\pi\frac{2i}{n_1+1}\right) \\
&= \sin(2\pi i - k\pi x_{0,i}) \\
&= \sin(-k\pi x_{0,i}) \\
&= -\sin(k\pi x_{0,i}) \\
&= -\left[\vec{v}_0^{(k)}\right]_i .
\end{aligned}$$

Similarly,

$$\cos((n_1+1-k)\pi h) = -\cos(k\pi h).$$

The first calculation is still valid for $n_0+1 \le k \le n_1$.

$$\begin{aligned}
\left[R_0\vec{v}_1^{(k)}\right]_i &= \frac{1}{2}\left\{\sin(k\pi x_{1,2i-1}) + 2\sin(k\pi x_{1,2i}) + \sin(k\pi x_{1,2i+1})\right\} \\
&= \frac{1}{2}\left\{\sin(k\pi x_{1,2i} - k\pi h) + 2\sin(k\pi x_{1,2i}) + \sin(k\pi x_{1,2i} + k\pi h)\right\} \\
&= \frac{1}{2}\left\{\sin(k\pi x_{1,2i})\cos(k\pi h) - \cos(k\pi x_{1,2i})\sin(k\pi h)\right. \\
&+ 2\sin(k\pi x_{1,2i}) \\
&+ \left.\sin(k\pi x_{1,2i})\cos(k\pi h) + \cos(k\pi x_{1,2i})\sin(k\pi h)\right\} \\
&= \frac{1}{2}\left\{2\sin(k\pi x_{1,2i})\cos(k\pi h) + 2\sin(k\pi x_{1,2i})\right\} \\
&= \sin(k\pi x_{1,2i})(\cos(k\pi h)+1) \\
&= (\cos(k\pi h)+1)\left[\vec{v}_0^{(k)}\right]_i \\
&= -(1-\cos((n_1+1-k)\pi h))\left[\vec{v}_0^{(n_1+1-k)}\right]_i \\
&= -2\sin^2\left(\frac{(n_1+1-k)\pi h}{2}\right)\left[\vec{v}_0^{(n_1+1-k)}\right]_i .
\end{aligned}$$

□

Proposition 3.15. *With the usual assumption, for* $1 \le k \le n_0$,

$$P_0\vec{v}_0^{(k)} = \cos^2\left(\frac{k\pi h}{2}\right)\vec{v}_1^{(k)} - \sin^2\left(\frac{k\pi h}{2}\right)\vec{v}_1^{(n_1+1-k)}. \tag{3.17}$$

Proof. Observe, for $1 \le k \le n_0$, a simple calculation shows

$$\left[\vec{v}_1^{(n_1+1-k)}\right]_i = (-1)^{i+1}\left[\vec{v}_1^{(k)}\right]_i .$$

Now, set

$$\begin{aligned}
C_k : &= \cos^2\left(\frac{k\pi h}{2}\right) = \frac{1}{2}(1+\cos(k\pi h)), \\
S_k : &= \sin^2\left(\frac{k\pi h}{2}\right) = \frac{1}{2}(1-\cos(k\pi h)).
\end{aligned}$$

Then, for $1 \le i \le n_1$, it follows that

$$
\begin{aligned}
& C_k \left[\vec{v}_1^{\,(k)}\right]_i - S_k \left[\vec{v}_1^{\,(n_1+1-k)}\right]_i \\
&= \begin{cases} \cos(k\pi h)\sin(k\pi x_{1,i}), & i \text{ odd}, \\ \sin(k\pi x_{1,i}), & i \text{ even}, \end{cases} \\
&= \begin{cases} \cos(k\pi h)\sin(k\pi x_{1,2j-1}), & 1 \le j \le n_0+1, \\ \sin(k\pi x_{1,2j}), & 1 \le j \le n_0, \end{cases} \\
&= \begin{cases} \frac{1}{2}\sin(k\pi h + k\pi x_{1,2j-1}) + \frac{1}{2}\sin(k\pi x_{1,2j-1} - k\pi h), & 1 \le j \le n_0+1, \\ \sin(k\pi x_{1,2j}), & 1 \le j \le n_0, \end{cases} \\
&= \begin{cases} \frac{1}{2}\sin(k\pi x_{1,2j}) + \frac{1}{2}\sin(k\pi x_{1,2j-2}), & 1 \le j \le n_0+1, \\ \sin(k\pi x_{1,2j}), & 1 \le j \le n_0, \end{cases} \\
&= \begin{cases} \frac{1}{2}\sin(k\pi x_{0,j}) + \frac{1}{2}\sin(k\pi x_{0,j-1}), & 1 \le j \le n_0+1, \\ \sin(k\pi x_{0,j}), & 1 \le j \le n_0, \end{cases} \\
&= \begin{cases} \left[P_0 \vec{v}_0^{\,(k)}\right]_{2j-1}, & 1 \le j \le n_0+1, \\ \left[P_0 \vec{v}_0^{\,(k)}\right]_{2j}, & 1 \le j \le n_0, \end{cases}
\end{aligned}
$$

□

Proposition 3.16. *With the usual assumption,*

$$(I_1 - \tilde{\Pi}_1)\vec{v}_1^{\,(k)} = S_k \vec{v}_1^{\,(k)} + S_k \vec{v}_1^{\,(n_1+1-k)}, \tag{3.18}$$

for $1 \le k \le n_0$. Similarly,

$$(I_1 - \tilde{\Pi}_1)\vec{v}_1^{\,(n_1+1-k)} = C_k \vec{v}_1^{\,(k)} + C_k \vec{v}_1^{\,(n_1+1-k)}, \tag{3.19}$$

for $1 \le k \le n_0 + 1$, where, as before

$$
\begin{aligned}
C_k : &= \cos^2\left(\frac{k\pi h}{2}\right) = \frac{1}{2}(1 + \cos(k\pi h)), \\
S_k : &= \sin^2\left(\frac{k\pi h}{2}\right) = \frac{1}{2}(1 - \cos(k\pi h)).
\end{aligned}
$$

Proof. Recall that

$$\tilde{\Pi}_1 = P_0 A_0^{-1} R_0 A_1 \in \mathbb{R}^{n_1 \times n_1} \quad \text{Ritz Projection},$$

where

$$A_0 = R_0 A_1 P_0 \in \mathbb{R}^{n_0 \times n_0} \quad \text{Galerkin Condition}.$$

Then, for $1 \le k \le n_0$,

$$
\begin{aligned}
\tilde{\Pi}_1 \vec{v}_1^{\,(k)} &= P_0 A_0^{-1} R_0 A_1 \vec{v}_1^{\,(k)} \\
&= \lambda_1^{(k)} P_0 A_0^{-1} R_0 \vec{v}_1^{\,(k)}
\end{aligned}
$$

$$\begin{aligned}
&\overset{(3.16)}{=} 2C_k\lambda_1^{(k)}P_0A_0^{-1}\vec{v}_0^{\,(k)} \\
&= \frac{2C_k\lambda_1^{(k)}}{\lambda_0^{(k)}}P_0\vec{v}_0^{\,(k)} \\
&\overset{(3.17)}{=} \frac{2C_k\lambda_1^{(k)}}{\lambda_0^{(k)}}\left\{C_k\vec{v}_1^{\,(k)} - S_k\vec{v}_1^{\,(n_1+1-k)}\right\}.
\end{aligned}$$

Observe,

$$\begin{aligned}
\frac{\lambda_1^{(k)}}{\lambda_0^{(k)}} &= \frac{\frac{2}{h}(1-\cos(k\pi h))}{\frac{2}{H}(1-\cos(k\pi H))} \\
&= \frac{H}{h}\frac{\sin^2\left(\frac{k\pi h}{2}\right)}{\sin^2(k\pi h)} \\
&= \frac{2S_k}{\sin^2(k\pi h)}.
\end{aligned}$$

Thus

$$\tilde{\Pi}_1\vec{v}_1^{\,(k)} = \frac{4C_kS_k}{\sin^2(k\pi h)}\left\{C_k\vec{v}_1^{\,(k)} - S_k\vec{v}_1^{\,(n_1+1-k)}\right\}.$$

But

$$C_kS_k = \frac{1}{4}\sin^2(k\pi h),$$

as the reader can easily check. We have

$$\tilde{\Pi}_1\vec{v}_1^{\,(k)} = C_k\vec{v}_1^{\,(k)} - S_k\vec{v}_1^{\,(n_1+1-k)}, \quad 1 \le k \le n_0. \tag{3.20}$$

Therefore, for $1 \le k \le n_0$,

$$\begin{aligned}
(I_1 - \tilde{\Pi}_1)\vec{v}_1^{\,(k)} &= (1-C_k)\vec{v}_1^{\,(k)} - S_k\vec{v}_1^{\,(n_1+1-k)} \\
&= S_k\left(\vec{v}_1^{\,(k)} - \vec{v}_1^{\,(n_1+1-k)}\right),
\end{aligned}$$

which is (3.18). Equation (3.19) is established in an analogous way. □

Theorem 3.17. *(Convergence of the one-sided two-grid method in $\|\cdot\|_2$) Suppose $m_2 = 0$ and $\omega_0 = \frac{2}{3}$, with $m_1 \ge 1$. Then*

$$\left\|\vec{e}_1^{\,\ell+1}\right\|_2^2 \le \left(\frac{1}{2} + \frac{1}{3^{m_1}}\right)\left\|\vec{e}_1^{\,\ell}\right\|_2^2.$$

Proof. We begin with the basis expansion

$$\vec{e}_1^{\,\ell} = \sum_{k=1}^{n_1}\epsilon_k\vec{v}_1^{\,(k)}.$$

The error propagation matrix, E_1, for the two-grid scheme, in this case, satisfies

$$\vec{e}_1^{\,\ell+1} = E_1\vec{e}_1^{\,\ell},$$

where

$$E_1 = (I_1 - \tilde{\Pi}_1)K_1^{m_1}.$$

Hence,

$$\vec{e}_1^{\,\ell+1} = \sum_{k=1}^{n_1} \epsilon_k (I_1 - \tilde{\Pi}_1)K_1^{m_1}\vec{v}_1^{\,(k)}.$$

By Theorem.3.3,

$$\left|\mu_1^{(k)}(\omega_0)\right| \le \frac{1}{3}, \quad n_0+1 \le k \le n_1,$$

and

$$\left|\mu_1^{(k)}(\omega)\right| < 1, \quad 1 \le k \le n_1.$$

By proposition.3.19,

$$(I_1 - \tilde{\Pi}_1)K_1^{m_1}\vec{v}_1^{\,(k)} = \alpha_k \left\{\vec{v}_1^{\,(k)} + \vec{v}_1^{\,(n_1+1-k)}\right\}, \tag{3.21}$$

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

$$(I_1 - \tilde{\Pi}_1)K_1^{m_1}\vec{v}_1^{\,(n_1+1-k)} = \beta_k \left\{\vec{v}_1^{\,(k)} + \vec{v}_1^{\,(n_1+1-k)}\right\}, \tag{3.22}$$

where

$$\alpha_k := \left(\mu_1^{(k)}(\omega_0)\right)^{m_1} S_k, \quad 1 \le k \le n_0+1,$$

and

$$\beta_k := \left(\mu_1^{(n_1+1-k)}(\omega_0)\right)^{m_1} C_k, \quad 1 \le k \le n_0+1.$$

Observe that we have extended the upper limit of the index k in (3.21) by 1, up to $k = n_0+1$. The result is still valid in this case. Further, observe that

$$\alpha_{n_0+1} = \beta_{n_0+1},$$

as the reader can easily check.

Now, notice that the eigenvalue satisfy

$$\begin{aligned} \left(\vec{v}_1^{\,(i)}, \vec{v}_1^{\,(j)}\right)_1 &= \sum_{m=1}^{n_1} \sin(i\pi x_{1,m} \sin(j\pi x_{1,m})) \\ &= \delta_{ij}\frac{n_1+1}{2}. \end{aligned}$$

Now, we can estimate α_k, β_k as follows:

$$|\alpha_k| \le 1^{m_1} S_k \le 1 \cdot \frac{1}{2}, \quad 1 \le k \le n_0+1,$$

and

$$\left|\beta_k\right| \le \left(\frac{1}{2}\right)^{m_1} C_k \le \left(\frac{1}{3}\right) \cdot 1, \quad 1 \le k \le n_0+1.$$

We can represent the error as

$$\begin{aligned} \vec{e}_1^{\,\ell+1} &= E_1 \vec{e}_1^{\,\ell} \\ &= \sum_{k=1}^{n_0+1} \delta_k \left(\epsilon_k \alpha_k + \epsilon_{n_1+1-k}\beta_k\right)\left(\vec{v}_1^{\,(k)} + \vec{v}_1^{\,(n_1+1-k)}\right), \end{aligned}$$

where

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

Therefore,

$$\begin{aligned}
\left\|\vec{e}_1^{\,\ell+1}\right\|_2^2 &= \left(\vec{e}_1^{\,\ell+1}, \vec{e}_1^{\,\ell+1}\right)_2 \\
&= (n_1+1)\sum_{k=1}^{n_0+1} \delta_k \left(\epsilon_k^2 \alpha_k^2 + \epsilon_{n_1+1-k}^2 \beta_k^2 + 2\epsilon_k \epsilon_{n_1+1-k} \alpha_k \beta_k\right) \\
&\overset{a.g.m.i}{\leq} (n_1+1)\sum_{k=1}^{n_0+1} \delta_k \left(\epsilon_k^2 \alpha_k^2 + \epsilon_{n_1+1-k}^2 \beta_k^2 + (\epsilon_k^2 + \epsilon_{n_1+1-k}^2)\left|\alpha_k\right|\left|\beta_k\right|\right) \\
&\leq (n_1+1)\sum_{k=1}^{n_0+1} \delta_k \left\{\epsilon_k^2 \left(\frac{1}{2}\right)^2 + \epsilon_{n_1+1-k}^2 \left(\frac{1}{3}\right)^{2m_1} + (\epsilon_k^2 + \epsilon_{n_1+1-k}^2)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{m_1}\right\}.
\end{aligned}$$

Since,

$$\left(\frac{1}{3}\right)^{2m_1} \leq \left(\frac{1}{2}\right)^2,$$

we have

$$\begin{aligned}
\left\|\vec{e}_1^{\,\ell+1}\right\|_2^2 &\leq \frac{n_1+1}{2}\left\{\frac{1}{2} + \left(\frac{1}{3}\right)^{2m_1}\right\}\sum_{k=1}^{n_0+1} \delta_k \left(\epsilon_k^2 + \epsilon_{n_1+1-k}^2\right) \\
&= \left\{\frac{1}{2} + \left(\frac{1}{3}\right)^{2m_1}\right\}\left\|\vec{e}_1^{\,\ell}\right\|_2^2.
\end{aligned}$$

Finally,

$$\left\|\vec{e}_1^{\,\ell+1}\right\|_2 \leq \left\{\frac{1}{2} + \left(\frac{1}{3}\right)^{2m_1}\right\}^{1/2}\left\|\vec{e}_1^{\,\ell}\right\|_2 := \gamma(m_1)\left\|\vec{e}_1^{\,\ell}\right\|_2.$$

□

m_1	1	2	3	$\cdots$	∞
$\gamma(m_1)$	0.91287	0.78176	0.73283	$\cdots$	0.70710

Remark 3.18. *The key observation is that $\gamma(m_1)$ is h-independent.*

3.7 Qualitative Two-Grid Method Convergence

Recall, for any $\vec{v}_1 \in \mathbb{R}^{n_1}$,

$$\begin{aligned}
\left\|\vec{v}_1\right\|_{A_1} : &= \sqrt{(\vec{v}_1, \vec{v}_1)_{A_1}} \\
&= \sqrt{(\vec{v}_1, A_1\vec{v}_1)_1}
\end{aligned}$$

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

In a similar way, we can define

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

Lemma 3.19. *With the usual assumptions in this section, there exists a positive constant $C_1 > 0$ such that*

$$
\left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_1 \le C_1\sqrt{\Lambda_1^{-1}}\left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_{A_1}, \tag{3.23}
$$

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

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

Here $h_1 = h, h_0 = H = 2h$.

Proof. Exercise. □

Remark 3.20. *We will prove this later in the general FEM framework. Question: What is the smallest (optimal) value of C_1 in the present setting ?*

Theorem 3.21. *(Approximation property) For any $\vec{v}_1 \in \mathbb{R}^{n_1}$,*

$$
\left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_{A_1} \le C_1\sqrt{\Lambda_1^{-1}}\left\|\vec{v}_1\right\|_{A_1^2}, \tag{3.24}
$$

where

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

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

$$
\begin{aligned}
\left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_{A_1}^2 &= \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1(I_1 - \tilde{\Pi}_1)\vec{v}_1\right)_1 \\
&= \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1\vec{v}_1\right)_1 - \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1\tilde{\Pi}_1\vec{v}_1\right)_1 .
\end{aligned}
$$

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

$$
\begin{aligned}
\left((I_1 - \tilde{\Pi}_1)\vec{v}_1, \tilde{\Pi}_1\vec{v}_1\right)_{A_1} &= \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1\tilde{\Pi}_1\vec{v}_1\right)_1 \\
&= \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1(P_0A_0^{-1}R_0A_1)\vec{v}_1\right)_1 \\
&= \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, (A_1P_0A_0^{-1}R_0)A_1\vec{v}_1\right)_1 \\
&= \left((A_1P_0A_0^{-1}R_0)^T(I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1\vec{v}_1\right)_1 \\
&= \left(\tilde{\Pi}_1(I_1 - \tilde{\Pi}_1)\vec{v}_1, \vec{v}_1\right)_{A_1}
\end{aligned}
$$

$$= \quad 0.$$

Since, $\tilde{\Pi}_1(I_1 - \tilde{\Pi}_1) = \tilde{\Pi}_1 - \tilde{\Pi}_1^2 = \tilde{\Pi}_1 - \tilde{\Pi}_1 = 0$. Therefore,

$$\begin{aligned}
\left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_{A_1}^2 &= \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, A_1\vec{v}_1\right)_1 \\
&\overset{C.S.}{\leq} \left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_1 \left\|A_1\vec{v}_1\right\|_1 \\
&= \left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_1 \left\|\vec{v}_1\right\|_{A_1^2} \\
&\overset{(3.23)}{\leq} C_1\sqrt{\Lambda_1^{-1}}\left\|(I_1 - \tilde{\Pi}_1)\vec{v}_1\right\|_{A_1} \left\|\vec{v}_1\right\|_{A_1^2}.
\end{aligned}$$

The result follows. □

Corollary 3.22. *Suppose that the usual assumptions from this section are in place, in particular, the Galerkin condition is satisfies in the definition of A_0. Then*

$$\left((I_1 - \tilde{\Pi}_1)\vec{v}_1, (I_1 - \tilde{\Pi}_1)\vec{v}_1\right)_{A_1} = \left((I_1 - \tilde{\Pi}_1)\vec{v}_1, \vec{v}_1\right)_{A_1},$$

or, equivalently ,

$$\left((I_1 - \tilde{\Pi}_1)\vec{v}_1, \tilde{\Pi}_1\vec{v}_1\right)_{A_1} = 0.$$

Remark 3.23. *Estimate (3.24) is called an approximation property (or condition) with constant C_1. Of course this estimate was directly required by estimate (3.23).*

3.8 Smoothing Revisited

In this discussion, we want to change our smooth a bit. Let's use Richardson's "smoothing" method.

$$K_1 = I_1 - \frac{1}{\Lambda_1}A_1, \quad \Lambda_1 = \frac{4}{h_1}.$$

Recall

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

So,

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

Λ_1 is almost the spectral radices of A_1, the bound is asymptotically sharp.

For damped Jacobi, recall that

$$\begin{aligned}
K_1 = K_1(\omega) &= I_1 - \omega D^{-1}A_1 \\
&= I_1 - \omega\frac{h_1}{2}A_1.
\end{aligned}$$

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

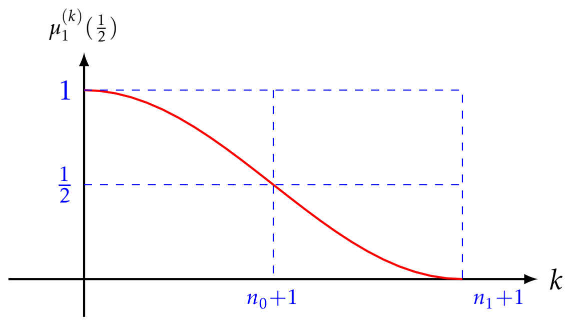
The eigenvalues of

$$K_1 = I_1 - \Lambda_1^{-1}A_1 = I_1 - \frac{h_1}{4}A_1,$$

are simply.

$$\mu_1^{(k)}\left(\frac{1}{2}\right) = \omega\cos k\pi h_1 + 1 - \omega$$

$$= \quad 1 - \frac{\lambda_1^{(k)}}{\Lambda_1}.$$



So

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

Lemma 3.24. *(Smoothing property for Richardson's smoother) There is some constant $C_2 > 0$, such that*

$$\left\|K_1^{m_1}\vec{v}_1\right\|_{A_1^2} \leq C_2\sqrt{\Lambda_1}\, m_1^{-\frac{1}{2}} \left\|\vec{v}_1\right\|_{A_1}. \tag{3.25}$$

for all $\vec{v}_1 \in \mathbb{R}^{n_1}$. ($\Lambda_1 = \frac{4}{h_1}$).

Proof.

$$\begin{aligned}\left\|K_1^{m_1}\vec{v}_1\right\|_{A_1^2} &= \left\|A_1K_1^{m_1}\vec{v}_1\right\|_1^2\\ &= \left(A_1K_1^{m_1}\vec{v}_1, A_1K_1^{m_1}\vec{v}_1\right)_1.\end{aligned}$$

Let us write

$$\vec{v}_1 = \sum_{k=1}^{n_1} \nu_k \vec{v}_1^{(k)}.$$

Then

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

where

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

Observe that, upon resealing

$$G(m) \leq \sup_{0\leq x\leq 1} x(1-x)^{2m}.$$

Set

$$f(x) = x(1-x)^{2m} = x(x-1)^{2m}.$$

$$\begin{aligned}
& 0 = f'(x_0) = (x_0-1)^{2m} + x_0(2m)(x_0-1)^{2m-1} \\
\Leftrightarrow\ & x-0-1+x_0(2m) = 0 \\
\Leftrightarrow\ & (1+2m)x_0 = 1 \\
\Leftrightarrow\ & x_0 = \frac{1}{2m+1} \\
\Leftrightarrow\ & f(x_0) = \frac{1}{2m+1}\left(\frac{2m}{2m+1}\right)^{2m}.
\end{aligned}$$

Thus

$$G(m) \leq \frac{1}{2m+1}\left(\frac{2m}{2m+1}\right)^{2m} \leq \frac{1}{2m+1} \leq \frac{1}{2}m^{-1}.$$

Therefore

$$\left\|K_1^{m_1}\vec{v}_1\right\|_{A_1^2} \leq \sqrt{\frac{1}{2}}\sqrt{\Lambda_1}m_1^{-\frac{1}{2}}\left\|\vec{v}_1\right\|_{A_1}. \tag{3.26}$$

and $C_2 = \sqrt{\frac{1}{2}}$. □

Theorem 3.25. *(Convergence of the one-sided Two-grid Method with Richardson Smoothing) Suppose $m_2 = 00$, and with $\omega = \frac{1}{2}$(Richardson). Then the two grid method obtained in this section converges provided m_1 is sufficiently large, and we have the qualitative error estimate*

$$\left\|\vec{e}_1^{\,\ell+1}\right\|_{A_1} \leq C_1C_2m_1^{-1/2}\left\|\vec{e}_1^{\,\ell}\right\|_{A_1}.$$

Where $C_1, C_2 > 0$ are as given in Theorem.3.21 and Lemma.3.24, respectively.

Proof. Recall, the error propagation matrix in this case is

$$E_1 = (I_1 - \tilde{\Pi}_1)K_1^{m_1}.$$

So

$$\begin{aligned}
\left\|\vec{e}_1^{\,\ell+1}\right\|_{A_1} &= \left\|(I_1-\tilde{\Pi}_1)K_1^{m_1}\vec{e}_1^{\,\ell}\right\|_{A_1} \\
&\overset{(3.24)}{\leq} C_1\sqrt{\Lambda_1^{-1}}\left\|K_1^{m_1}\vec{e}_1^{\,\ell}\right\|_{A_1^2} \\
&\overset{(3.25)}{\leq} C_1\sqrt{\Lambda_1^{-1}}C_2\sqrt{\Lambda_1}m_1^{-\frac{1}{2}}\left\|\vec{e}_1^{\,\ell}\right\|_{A_1} \\
&= C_1C_2m_1^{-\frac{1}{2}}\left\|\vec{e}_1^{\,\ell}\right\|_{A_1}.
\end{aligned}$$

□

Remark 3.26. *The method converges at a uniform, h-independent rate provide $m_1 \geq 1$ is large enough so that*

$$0 < C_1C_2m_1^{-\frac{1}{2}} < 1.$$

We have shown that $C_2 = \sqrt{\frac{1}{2}}$ works, but since we do not have a quantitative estimate of C_1 (yet), the convergence is just "qualitative".

4 Multigrid

The idea behind multigrid is to replace the exact solution in the coarse-grid connection by a recursive application of a "two-grid" method.

Suppose

$$1 \le n_0 < n_1 < \cdots < n_\ell < \cdots < n_L \in \mathbb{Z}.$$

Suppose

$$R_{\ell-1} \in \mathbb{R}^{n_{\ell-1}\times n_\ell}, \quad 1 \le \ell \le L,$$

is full rank, i.e.

$$\operatorname{rank}(R_{\ell-1}) = n_{\ell-1}.$$

Set

$$P_{\ell-1} = R_{\ell-1}^T \in \mathbb{R}^{n_\ell \times n_{\ell-1}}, \quad 1 \le \ell \le L.$$

Assume that we have a family of SPD matrices

$$A_\ell \in \mathbb{R}^{n_\ell \times n_\ell}, \quad 0 \le \ell \le L.$$

Symmetry here is understood with respect to the canonical product $(\cdot, cdot)_\ell : \mathbb{R}^{n_\ell} \times \mathbb{R}^{n_\ell} \to \mathbb{R}$, which is defined as

$$(\vec{u}_\ell, \vec{v}_\ell)_\ell = \vec{u}_1^T \vec{v}_\ell = \sum_{j=1}^{n_\ell} u_{\ell,j} v_{\ell,j},$$

for all $\vec{u}_\ell, \vec{v}_\ell \in \mathbb{R}^{n_\ell}$. We define $(\cdot,\cdot)_{A_\ell}$ via

$$(\vec{u}_\ell, \vec{v}_\ell)_{A_\ell} = (\vec{u}_\ell, A_\ell \vec{v}_\ell)_\ell .$$

Our goal is to find an efficient iterative solver for the equation

$$A_L \vec{u}_L^{\,E} = \vec{f}_L,$$

given $\vec{f}_L \in \mathbb{R}^{n_L}$. As usual, we define

$$\vec{e}_L^{\,\square} = \vec{u}_L^{\,E} - \vec{u}_L^{\,\square},$$

where $\vec{u}_L^{\,\square} \in \mathbb{R}^{n_L}$ is an approximate solution. Similarly,

$$\vec{r}_L^{\,\square} := \vec{f}_L - A_L \vec{u}_L^{\,\square} = A_L \vec{e}_L^{\,\square}.$$

Definition 4.1. *We say that the Galerkin Condition holds iff*

$$A_{\ell-1} = R_{\ell-1} A_\ell P_{\ell-1},$$

for all $1 \le \ell \le L$. Otherwise, the Galerkin Condition fails.

Definition 4.2. *Suppose $0 \le \ell \le L$ and $\vec{g}_\ell, \vec{u}_\ell^{\,(0)} \in \mathbb{R}^{n_\ell}$ are given. The recursive multi grid operator MG is defined as follows:*

$$\vec{u}_\ell^{\,(3)} := MG\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right)$$

- *If $(\ell = 0)$ then*

$$\vec{u}_\ell^{\,(3)} := A_0^{-1}\vec{g}_0$$

- *else if $(1 \le \ell \le L)$ then*
  - *Pre-Smoothing:*
    - $\vec{u}_\ell^{\,(1,0)} := \vec{u}_\ell^{\,(0)}$
    - $\vec{u}_\ell^{\,(1,\sigma+1)} = \vec{u}_\ell^{\,(1,\sigma)} + S_\ell\left(\vec{g}_\ell - A_\ell \vec{u}_\ell^{\,(1,\sigma)}\right), \quad 0 \le \sigma \le m_1 - 1$
    - $\vec{u}_\ell^{\,(1)} := \vec{u}_\ell^{\,(1,m_1)}$
  - *Coarse-Grid Correction:*
    - $\vec{r}_\ell^{\,(1)} := \vec{g}_\ell - A_\ell \vec{u}_\ell^{\,(1)}$
    - $\vec{r}_{\ell-1}^{\,(1)} := R_{\ell-1}\vec{r}_\ell^{\,(1)}$
    - $\vec{q}_{\ell-1}^{\,(1,0)} := \vec{0}$
    - $\vec{q}_{\ell-1}^{\,(1,\sigma+1)} := \mathrm{MG}\left(\vec{r}_{\ell-1}^{\,(1)}, \ell-1, \vec{q}_{\ell-1}^{\,(1,\sigma)}\right), \quad 0 \le \sigma \le p-1$
    - $\vec{q}_{\ell-1}^{\,(1)} := \vec{q}_{\ell-1}^{\,(1,p)}$
    - $\vec{u}_\ell^{\,(2)} := \vec{u}_\ell^{\,(1)} + P_{\ell-1}\vec{q}_{\ell-1}^{\,(1)}$
  - *Post-Smoothing:*
    - $\vec{u}_\ell^{\,(3,0)} := \vec{u}_\ell^{\,(2)}$
    - $\vec{u}_\ell^{\,(3,\sigma+1)} := \vec{u}_\ell^{\,(3,\sigma)} + S_1^T\left(\vec{g}_\ell - A_\ell \vec{u}_\ell^{\,(3,\sigma)}\right), \quad 0 \le \sigma \le m_2 - 1$
- $\vec{u}_\ell^{\,(3)} = \vec{u}_\ell^{\,(3,m_2)}$

Algorithm 4.3. *Let m_1, m_2, p be nonnegative integers. Let $\vec{u}_L^{\,k} \in \mathbb{R}^{n_L}$ be given. Then*

$$\vec{u}_L^{\,k+1} = \mathrm{MG}\left(\vec{f}_L, L, \vec{u}_L^{\,k}\right)$$

defines the genetic multigrid iteration method for solving

$$A_L \vec{u}_L^{\,E} = \vec{f}_L.$$

Definition 4.4. *(one-sided multigrid method) The multigrid (multilevel) algorithm is called one-sided iff $m_2 = 0$ and $m_1 \ge 1$. The algorithm is called a W-cycle iff $p = 2$, and a V-cycle iff $p = 1$. The algorithm is symmetric iff $m_1 = m_2 = m$.*

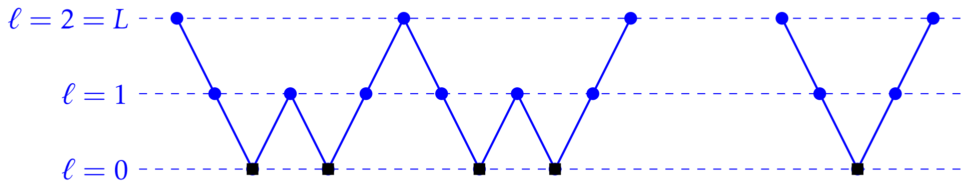


Figure 20: One full W- and V- Cycle for three levels. Left One full W-Cycle, Right: One full V- Cycle .

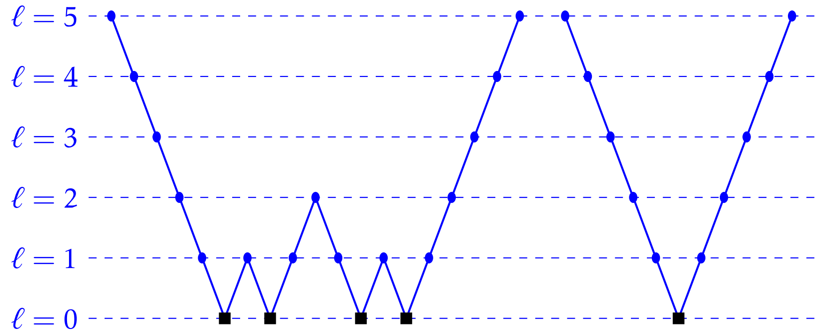


Figure 21: One full W- and V- Cycle for six levels. Left One full W-Cycle, Right: One full V- Cycle .

Remark 4.5. *In the case $L = 0$, we just do a direct solve. When $L = 1$, we recover the two-grid (two-level) method.*

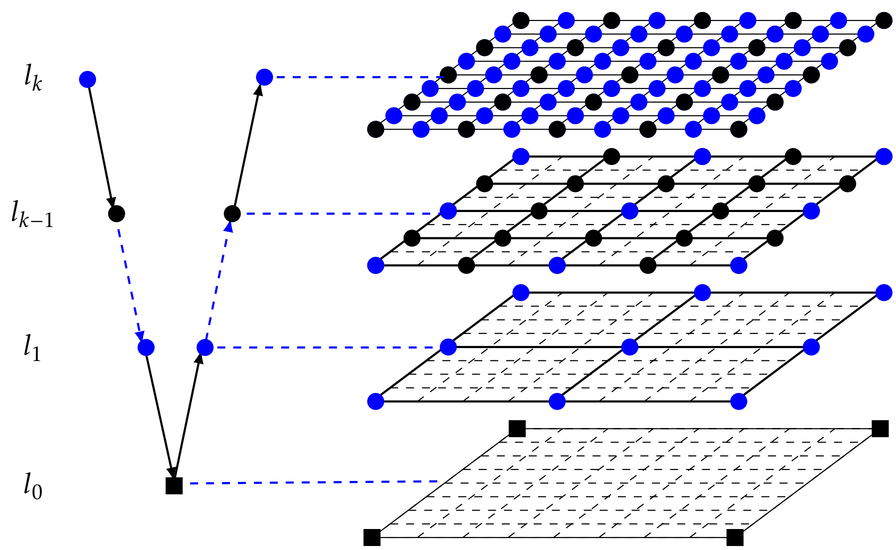


Figure 22: Mesh on the fine and coarse grid of the multigrid method in 2D.

We now want to find the error propagation matrix for the general multilevel algorithm.

Definition 4.6. *Let $\ell \geq 0$. Define for $\ell = 0$*

$$E_0 := 0 \in \mathbb{R}^{n_0 \times n_0},$$

and, for $\ell \geq 1$,

$$E_\ell := \left(K_\ell^*\right)^{m_2} \left(I_\ell - P_{\ell-1}\left(I_{\ell-1} - E_{\ell-1}^p\right)\Pi_{\ell-1}\right) K_\ell^{m_1} \in \mathbb{R}^{n_\ell \times n_\ell},$$

where

$$\left.\begin{aligned} K_\ell &:= I_\ell - S_\ell A_\ell \\ K_\ell^* &= I_\ell - S_\ell^T A_\ell \end{aligned}\right\} \in \mathbb{R}^{n_\ell \times n_\ell},$$

and

$$\Pi_{\ell-1} := A_{\ell-1}^{-1} R_{\ell-1} \in \mathbb{R}^{n_{\ell-1} \times n_\ell}.$$

For further use, let us also define

$$\tilde{\Pi}_\ell := P_{\ell-1}\Pi_{\ell-1} \in \mathbb{R}^{n_\ell \times n_\ell}.$$

Remark 4.7. *Observe that, for any $\vec{u}_\ell, \vec{v}_\ell \in \mathbb{R}^{n_\ell}$,*

$$(\vec{u}_\ell, K_\ell \vec{v}_\ell)_{A_\ell} = \left(K_\ell^* \vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell}.$$

More generally, we use "" to denote the adjoint w.r.t $(\cdot,\cdot)_{A_\ell}$.*

Theorem 4.8. *Suppose that $\vec{u}_\ell^{\,E}, \vec{g}_\ell \in \mathbb{R}^{n_\ell}$ satisfy*

$$A_\ell \vec{u}_\ell^{\,E} = \vec{g}_\ell.$$

Then, given $\vec{u}_\ell^{\,(0)} \in \mathbb{R}^{n_\ell}$,

$$\vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right) = E_\ell\left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right),$$

where E_ℓ is the recursively -defined matrix from Definition.4.6. In particular

$$\vec{e}_L^{\,k+1} = E_L \vec{e}_L^{\,k}.$$

Proof. The proof is by induction. Cases $\ell = 0$ and $\ell = 1$ (two-grid method) are clear.

Induction Hypothesis: Assume that the result is true for level $\ell - 1$. Suppose that $\vec{q}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}}$ solves

$$A_{\ell-1}\vec{q}_{\ell-1} = \vec{r}_{\ell-1}^{\,(1)}.$$

Then

$$\vec{q}_{\ell-1} - \mathrm{MG}\left(\vec{r}_{\ell-1}^{\,(1)}, \ell-1, \vec{0}\right) = E_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{0}\right) = E_{\ell-1}\vec{q}_{\ell-1}.$$

Written in the language of Algorithm.4.3,

$$\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1,1)} = E_{\ell-1}\vec{q}_{\ell-1}.$$

So

$$\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1,p)} = E_{\ell-1}^p \vec{q}_{\ell-1},$$

or

$$\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)} = E_{\ell-1}^p \vec{q}_{\ell-1}.$$

Therefore

$$\vec{q}^{\,(1)}_{\ell-1} = \left(I_{\ell-1} - E^{p}_{\ell-1}\right)\vec{q}_{\ell-1}.$$

Now,

$$\begin{aligned}
\vec{q}_{\ell-1} &:= A^{-1}_{\ell-1}\vec{r}^{\,(1)}_{\ell-1} \\
&= A^{-1}_{\ell-1}R_{\ell-1}\left(\vec{g}_{\ell} - A_{\ell}\vec{u}^{\,(1)}_{\ell}\right) \\
&= A^{-1}_{\ell-1}R_{\ell-1}A_{\ell}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell}\right) \\
&= \Pi_{\ell-1}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell}\right).
\end{aligned}$$

Hence

$$\vec{q}^{\,(1)}_{\ell-1} = \left(I_{\ell-1} - E^{p}_{\ell-1}\right)\Pi_{\ell-1}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell}\right).$$

Putting it all together,

$$\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(3)}_{\ell} = \left(K^{*}_{\ell}\right)^{m_2}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(2)}_{\ell}\right),$$

$$\begin{aligned}
\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(2)}_{\ell} &= \vec{u}^{\,E}_{\ell} - \left(\vec{u}^{\,(1)}_{\ell} + P_{\ell-1}\vec{q}^{\,(1)}_{\ell-1}\right) \\
&= \vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell} - P_{\ell-1}\left(I_{\ell-1} - E^{p}_{\ell-1}\right)\Pi_{\ell-1}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell}\right) \\
&= \left(I_{\ell} - P_{\ell-1}\left(I_{\ell-1} - E^{p}_{\ell-1}\right)\Pi_{\ell-1}\right)\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell}\right), \\
\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(1)}_{\ell} &= K^{m_1}_{\ell}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(0)}_{\ell}\right),
\end{aligned}$$

we have

$$\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(3)}_{\ell} = E_{\ell-1}\left(\vec{u}^{\,E}_{\ell} - \vec{u}^{\,(0)}_{\ell}\right).$$

□

Definition 4.9. *(Assumption (A0) and Assumption (A1)) We say that the stiffness matrices satisfy the Galerkin Condition, Assumption (A0), iff*

$$A_{\ell-1} = R_{\ell-1}A_{\ell}P_{\ell-1}, \quad 1 \le \ell \le L. \tag{4.1}$$

We say that the Weakened Galerkin Condition, Assumption (A1), holds iff

$$(\vec{v}_{\ell}, \vec{v}_{\ell})_{A_{\ell}} \ge (\Pi_{\ell-1}\vec{v}_{\ell}, \Pi_{\ell-1}\vec{v}_{\ell})_{A_{\ell-1}}, \tag{4.2}$$

for all $\vec{v}_{\ell} \in \mathbb{R}^{n_{\ell}}$ and all $1 \le \ell \le L$.

We will show that $A0$ $A1$. To do this, we first need the following:

Lemma 4.10. *Assumption (A0), the (strong) Galerkin condition, implies that $\tilde{\Pi}_{\ell} = \tilde{\Pi}^{2}_{\ell}$. But $\tilde{\Pi}^{*}_{\ell} = \tilde{\Pi}_{\ell}$ holds even without (A0).*

Proof. If the Galerkin Condition holds then

$$A_{\ell-1} = R_{\ell-1}A_{\ell}P_{\ell-1}, \quad 1 \le \ell \le L.$$

Recall that

$$\tilde{\Pi}_\ell = P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell, \quad 1 \le \ell \le L.$$

Consequently,

$$\begin{aligned}
\tilde{\Pi}_\ell^2 &= P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell \\
&= P_{\ell-1}A_{\ell-1}^{-1}\left(R_{\ell-1}A_\ell P_{\ell-1}\right)A_{\ell-1}^{-1}R_{\ell-1}A_\ell \\
&= P_{\ell-1}A_{\ell-1}^{-1}A_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell \\
&= P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell \\
&= \tilde{\Pi}_\ell.
\end{aligned}$$

Next, let $\vec{u}_\ell, \vec{v}_\ell \in \mathbb{R}^{n_\ell}$ be arbitrary. Then just using the definition of $\Pi_{\ell-1}, \tilde{\Pi}_\ell$, which hold indpendent of (A0),

$$\begin{aligned}
\left(\vec{u}_\ell, \tilde{\Pi}_\ell \vec{v}_\ell\right)_{A_\ell} &= \left(\vec{u}_\ell, A_\ell \tilde{\Pi}_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\vec{u}_\ell, A_\ell P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\vec{u}_\ell, \left(A_\ell P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}\right)A_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\left(A_\ell P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}\right)^T \vec{u}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\tilde{\Pi}_\ell \vec{u}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\tilde{\Pi}_\ell \vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}$$

So,

$$\tilde{\Pi}_\ell^* = \tilde{\Pi}_\ell.$$

□

Corollary 4.11. $\left(I_\ell - \tilde{\Pi}_\ell\right)^2 = I_\ell - \tilde{\Pi}_\ell$ *and* $\left(I_\ell - \tilde{\Pi}_\ell\right)^* = I_\ell - \tilde{\Pi}_\ell$.

Lemma 4.12. *(Assumption (A0) and Assumption (A1)) Assumption (A0) implies Assumption (A1).*

Proof. First, for any $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$, consider

$$\begin{aligned}
(\Pi_{\ell-1}\vec{v}_\ell, \Pi_{\ell-1}\vec{v}_\ell)_{A_{\ell-1}} &= (\Pi_{\ell-1}\vec{v}_\ell, A_{\ell-1}\Pi_{\ell-1}\vec{v}_\ell)_{\ell-1} \\
&= \left(\Pi_{\ell-1}\vec{v}_\ell, A_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell\vec{v}_\ell\right)_{\ell-1} \\
&= (\Pi_{\ell-1}\vec{v}_\ell, R_{\ell-1}A_\ell\vec{v}_\ell)_{\ell-1} \\
&= \left(R_{\ell-1}^T\Pi_{\ell-1}\vec{v}_\ell, A_\ell\vec{v}_\ell\right)_\ell \\
&= (P_{\ell-1}\Pi_{\ell-1}\vec{v}_\ell, A_\ell\vec{v}_\ell)_\ell \\
&= (P_{\ell-1}\Pi_{\ell-1}\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} \\
&= \left(\tilde{\Pi}_\ell\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}$$

Using the last calculation

$$\begin{aligned}
& (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - (\Pi_{\ell-1}\vec{v}_\ell, \Pi_{\ell-1}\vec{v}_\ell)_{A_{\ell-1}} \\
=\; & (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - \left(\tilde{\Pi}_\ell\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}
\end{aligned}$$

$$\begin{aligned}
&= \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} \\
&\overset{Cor.(4.11)}{=} \left(\left(I_\ell - \tilde{\Pi}_\ell\right)^2 \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} \\
&\overset{Cor.(4.11)}{=} \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{v}_\ell, \left(I_\ell - \tilde{\Pi}_\ell\right)\vec{v}_\ell\right)_{A_\ell} \\
&= \left\|\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{v}_\ell\right\|_{A_\ell} \geq 0.
\end{aligned}$$

Thus (A0) implies (A1). □

Corollary 4.13. *(Assumption (A2)) Assumption (A1) is equivalent to the following statement (Assumption (A2)): for all* $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$ *and* $1 \leq \ell \leq L$

$$\left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \geq 0. \tag{4.3}$$

Proof. We showed in the proof of lemma.4.12, using only the definition of $\Pi_{\ell-1}, \tilde{\Pi}_\ell$, that

$$\begin{aligned}
&(\vec{u}_\ell, \vec{u}_\ell)_{A_\ell} - (\Pi_{\ell-1}\vec{u}_\ell, \Pi_{\ell-1}\vec{u}_\ell)_{A_{\ell-1}} \\
= \;& (\vec{u}_\ell, \vec{v}_\ell)_{A_\ell} - \left(\tilde{\Pi}_\ell \vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \\
= \;& \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \\
= \;& \left(\vec{u}_\ell, \left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell\right)_{A_\ell}.
\end{aligned}$$

So (A1) holds iff statement (4.3) holds. □

Definition 4.14. *We say that Assumption (A2) holds for a multilevel method iff statement (4.3) holds.*

Theorem 4.15. *Suppose that the weakened Galerkin Condition, Assumption (A1), holds, or equivalently that Assumption (A2) holds. Then if* $m_1 = m_2 = m$, *for all* $0 \leq \ell \leq L$

$$E_\ell = E_\ell^*,$$

and

$$(E_\ell \vec{u}_\ell, \vec{u}_\ell)_{A_\ell} \geq 0, \quad \forall \vec{u}_\ell \in \mathbb{R}^{n_\ell}.$$

Assumption (A1) is required only for the second statement.

Proof. By induction: The case $\ell = 0$ is trivial. For $\ell = 1$ (the two-grid algorithm) the proof follows as in the proof of Theorem.(2.12).

Induction Hypothesis: Assume that

$$(E_{\ell-1}\vec{u}_{\ell-1}, \vec{v}_{\ell-1})_{A_{\ell-1}} = (\vec{u}_{\ell-1}, E_{\ell-1}\vec{v}_{\ell-1})_{A_{\ell-1}},$$

and

$$(E_{\ell-1}\vec{u}_{\ell-1}, \vec{v}_{\ell-1})_{A_{\ell-1}} \geq 0$$

for all $\vec{u}_{\ell-1}, \vec{v}_{\ell-1} \in \mathbb{R}^{n_\ell}$. We will make use of the definition of $\Pi_{\ell-1}, \tilde{\Pi}_\ell$, a number of times:

$$\Pi_{\ell-1} = A_{\ell-1}^{-1} R_{\ell-1} A_\ell.$$

So

$$A_{\ell-1}\Pi_{\ell-1} = R_{\ell-1}A_\ell.$$

And

$$\tilde{\Pi}_\ell = P_{\ell-1}\Pi_{\ell-1} = P_{\ell-1}A_{\ell-1}^{-1}R_{\ell-1}A_\ell.$$

We have

$$\begin{aligned}
& \left(E_\ell\vec{u}_\ell,\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\left(K_\ell^*\right)^{m_2}\left(I_\ell - P_{\ell-1}\left(I_{\ell-1}-E_{\ell-1}^p\right)\Pi_{\ell-1}\right)K_\ell^{m_1}\vec{u}_\ell,\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\left(K_\ell^*\right)^{m}\left(I_\ell - P_{\ell-1}\left(I_{\ell-1}-E_{\ell-1}^p\right)\Pi_{\ell-1}\right)K_\ell^{m}\vec{u}_\ell,\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\left(I_\ell - P_{\ell-1}\left(I_{\ell-1}-E_{\ell-1}^p\right)\Pi_{\ell-1}\right)K_\ell^{m}\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\left(I_\ell - P_{\ell-1}\Pi_{\ell-1}+P_{\ell-1}E_{\ell-1}^p\Pi_{\ell-1}\right)K_\ell^{m}\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\left(I_\ell - P_{\ell-1}\Pi_{\ell-1}\right)K_\ell^{m}\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(P_{\ell-1}E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^{m}\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(P_{\ell-1}E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(P_{\ell-1}E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,K_\ell^m\vec{v}_\ell\right)_{A_\ell} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(P_{\ell-1}E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,A_\ell K_\ell^m\vec{v}_\ell\right)_{\ell} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,P_{\ell-1}^T A_\ell K_\ell^m\vec{v}_\ell\right)_{\ell-1} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,R_{\ell-1} A_\ell K_\ell^m\vec{v}_\ell\right)_{\ell-1} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,A_{\ell-1}\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{\ell-1} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{A_{\ell-1}} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,E_{\ell-1}^p\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{A_{\ell-1}} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(A_{\ell-1}\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,E_{\ell-1}^p\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{\ell-1} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(R_{\ell-1}A_\ell K_\ell^m\vec{u}_\ell,E_{\ell-1}^p\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{\ell-1} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(A_\ell K_\ell^m\vec{u}_\ell,R_{\ell-1}^T E_{\ell-1}^p\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{\ell} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(A_\ell K_\ell^m\vec{u}_\ell,P_{\ell-1} E_{\ell-1}^p\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{\ell} \\
= & \left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{v}_\ell\right)_{A_\ell} + \left(K_\ell^m\vec{u}_\ell,P_{\ell-1} E_{\ell-1}^p\Pi_{\ell-1} K_\ell^m\vec{v}_\ell\right)_{A_\ell} \\
= & \left(\vec{u}_\ell,E_\ell\vec{v}_\ell\right)_{A_\ell}.
\end{aligned}$$

Symmetry is proven. Notice we have not yet used Assumption (A1), only the definition of $\Pi_{\ell-1},\tilde{\Pi}_\ell$, which are assumed to always hold.

Now, we setting $\vec{v}_\ell = \vec{u}_\ell$ in the last calculation, we have

$$\left(\vec{u}_\ell,E_\ell\vec{u}_\ell\right)_{A_\ell} = \underbrace{\left(K_\ell^{m}\vec{u}_\ell,\left(I_\ell - \tilde{\Pi}_\ell\right)K_\ell^m\vec{u}_\ell\right)_{A_\ell}}_{\geq 0,\text{by (A2)}} + \underbrace{\left(E_{\ell-1}^p\Pi_{\ell-1}K_\ell^m\vec{u}_\ell,\Pi_{\ell-1} K_\ell^m\vec{u}_\ell\right)_{A_{\ell-1}}}_{\geq 0,\text{by Induction Hypothesis}} \geq 0.$$

for any $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$. The result is proven. □

Remark 4.16. *In the proof above we used the fact that any positive integer power of a self-adjoint positive semi-definite matrix is also positive semi-definite. This is easy to prove and is left as an exercise.*

Definition 4.17. *(Assumption (A3): Strong approximation property and Assumption (A4): weakened approximation property) We say that the multilevel algorithm satisfies the Strong approximation property , or Assumption (A3), iff for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$ and $1 \le \ell \le L$*

$$\left\| \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right\|_\ell^2 \le C_3^2 \rho_\ell^{-1} \left\| \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right\|_{A_\ell}^2, \tag{4.4}$$

for some $C_3 > 0$ that is independent of ℓ, where $\rho_\ell = \rho(A_\ell)$. The multilevel algorithm satisfies the Weakened approximation property , or Assumption (A4), iff for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$ and $1 \le \ell \le L$

$$\left(\left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell, \vec{u}_\ell \right)_{A_\ell} \le C_4^2 \rho_\ell^{-1} \left\| A_\ell \vec{u}_\ell \right\|_\ell^2, \tag{4.5}$$

for some $C_4 > 0$ that is independent of ℓ.

Theorem 4.18. *If the Galerkin condition, Assumption (A0), holds , then (A3) implies (A4).*

Proof. Since (A0) holds

$$\left(I_\ell - \tilde{\Pi}_\ell \right)^2 = I_\ell - \tilde{\Pi}_\ell.$$

Therefore

$$\begin{aligned}
\left\| \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right\|_{A_\ell}^2 &= \left(\left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell, \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right)_{A_\ell} \\
&= \left(\left(I_\ell - \tilde{\Pi}_\ell \right)^2 \vec{u}_\ell, \vec{u}_\ell \right)_{A_\ell} \\
&= \left(\left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell, \vec{u}_\ell \right)_{A_\ell} \\
&= \left(\left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell, A_\ell \vec{u}_\ell \right)_\ell \\
&\overset{c.s.}{\le} \left\| \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right\|_\ell \left\| A_\ell \vec{u}_\ell \right\|_\ell \\
&\overset{(A3)}{\le} C_3 \rho_\ell^{-1/2} \left\| \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right\|_{A_\ell} \left\| A_\ell \vec{u}_\ell \right\|_\ell.
\end{aligned}$$

Thus

$$\left\| \left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell \right\|_{A_\ell} \le C_3 \rho_\ell^{-1/2} \left\| A_\ell \vec{u}_\ell \right\|_\ell. \tag{4.6}$$

Squaring, we have

$$\left(\left(I_\ell - \tilde{\Pi}_\ell \right) \vec{u}_\ell, \vec{u}_\ell \right)_{A_\ell} \le C_3^2 \rho_\ell^{-1} \left\| A_\ell \vec{u}_\ell \right\|_\ell^2,$$

which is the desired result with $C_4 = C_3$. □

4.1 Richardson's Smoother

Definition 4.19. *Recall that Richardson's method is defined via*

$$\vec{u}_\ell^{\,(\sigma+1)} = \vec{u}_\ell^{\,(\sigma)} + \omega^{-1}\left(\vec{g}_\ell - A_\ell \vec{u}_\ell^{\,(\sigma)}\right),$$

where ω us a parameter, In this case

$$K_\ell = I_\ell - \omega^{-1} A_\ell = K_\ell^*.$$

Choosing

$$C_s \rho_\ell \geq \omega := \Lambda_\ell \geq \rho_\ell = \rho(A_\ell), \quad 1 \leq \ell \leq L, \exists C_s \geq 1,$$

we obtain Richardson' s smoother. In this case

$$K_\ell = I_\ell - \Lambda_\ell^{-1} A_\ell = K_\ell^*.$$

Definition 4.20. *(Assumption (A5): First Smoothing Property) We say that the multilevel scheme satisfies the First Smoothing Property , Assumption (A5), iff*

$$\left\| K_\ell^m \vec{u}_\ell \right\|_{A_\ell^2} \leq C_5 \rho_\ell^{1/2} m^{-1/2} \left\| \vec{u}_\ell \right\|_{A_\ell}, \tag{4.7}$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$ and $1 \leq \ell \leq L$, for some $C_5 > 0$ that is independent of ℓ.

Theorem 4.21. *Richardson's smoother satisfies that first smoothing property, Assumption (A5).*

Proof. The proof is similar to the proof of Lemma.3.24 . □

4.2 Convergence of the Two-level Method Revisited

Theorem 4.22. *Suppose that $L = 1$ (Two-level) $m_1 = m \geq 1$ and $m_2 = 0$ (once-sided). Suppose that Assumption (A0) (Galerkin Condition), Assumption (A3) (Strong Approximation property) and (A5) (1^{st} smoothing property) all hold. Then*

$$\left\| \vec{u}_1^{\,E} - \mathrm{TG}_1\left(\vec{f}_1, \vec{u}_1^{\,(0)}\right) \right\|_{A_1} \leq C_3 C_5 m^{-1/2} \left\| \vec{u}_1^{\,E} - \vec{u}_1^{\,(0)} \right\|_{A_1}$$

where

$$A_1 \vec{u}_1^{\,E} = \vec{f}_1.$$

Written another way,

$$\left\| \vec{e}_1^{\,k+1} \right\|_{A_1} \leq C_3 C_5 m^{-1/2} \left\| \vec{e}_1^{\,k} \right\|_{A_1}$$

Proof. Recall that, in the present case,

$$E_1 = \left(I_1 - \tilde{\Pi}_1\right) K_1^m,$$

and

$$\vec{e}_1^{\,k+1} = E_1 \vec{e}_1^{\,k},$$

or, equivalently

$$\vec{u}_1^{\,E} - \mathrm{TG}_1\left(\vec{f}_1, \vec{u}_1^{\,(0)}\right) = E_1\left(\vec{u}_1^{\,E} - \vec{u}_1^{\,(0)}\right).$$

We showed in the proof of Theorem.4.18 that Assumption (A0) and (A3) imply that

$$\left\|\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell\right\|_{A_\ell} \leq C_3\rho_\ell^{-1/2}\left\|A_\ell\vec{u}_\ell\right\|_\ell, \tag{4.8}$$

for any $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$. Applying (4.8) (with $\ell = 1$), and using Assumption (A5), we have

$$\begin{aligned}
\left\|\vec{e}_1^{\,k+1}\right\|_{A_1} &= \left\|\left(I_1 - \tilde{\Pi}_1\right)K_1^m\vec{e}_1^{\,k}\right\|_{A_1} \\
&\overset{(4.8)}{\leq} C_3\rho_1^{-1/2}\left\|A_\ell K_1^m\vec{e}_1^{\,k}\right\|_1 \\
&\overset{(A5)}{\leq} C_3\rho_1^{-1/2}C_5\rho_1^{1/2}m^{-1/2}\left\|\vec{e}_1^{\,k}\right\|_{A_1} \\
&= C_3C_5m^{-1/2}\left\|\vec{e}_1^{\,k}\right\|_{A_1}.
\end{aligned}$$

□

Before we go to the next result, we need a technical lemma:

Lemma 4.23. *For Richardson's smoother we have*

$$\left\|K_\ell\vec{v}_\ell\right\|_{A_\ell} \leq \left\|\vec{v}_\ell\right\|_{A_\ell} \tag{4.9}$$

$$(K_\ell\vec{v}_\ell, \vec{v}_\ell)_\ell \leq (\vec{v}_\ell, \vec{v}_\ell)_\ell. \tag{4.10}$$

for all $\vec{v}_\ell \in \mathbb{R}^{n_\ell}, \ell \geq 0$

Proof. 1. Let $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$ be arbitrary. Suppose $\left\{\vec{\omega}_\ell^{(1)}\right\}, \left\{\vec{\omega}_\ell^{(2)}, \cdots, \left\{\vec{\omega}_\ell^{(n_\ell)}\right\}\right\}$ is an orthonormal basis of eigenvalue of A_ℓ w.r.t $\|\cdot,\cdot\|_\ell$. Then there exist unique $\alpha_1, \alpha_2, \cdots, \alpha_{n_\ell} \in \mathbb{R}$ such that

$$\vec{v}_\ell = \sum_{k=1}^{n_\ell}\alpha_k\vec{\omega}_\ell^{\,(k)}.$$

Recall

$$K_\ell = I_\ell - \Lambda_\ell^{-1}A_\ell,$$

with

$$\rho(A_\ell) =: \rho_\ell \leq \Lambda_\ell \leq C_s\rho_\ell,$$

where $C_s \geq 1$ is independent of ℓ. Then

$$K_\ell\vec{\omega}_\ell^{\,(k)} = \left(1 - \frac{\lambda_\ell^{(k)}}{\Lambda_\ell}\right)\vec{\omega}_\ell^{\,(k)} =: \mu_\ell^{(k)}\vec{\omega}_\ell^{\,(k)}.$$

Thus

$$\begin{aligned}
\left\|K_\ell\vec{v}_\ell\right\|_{A_\ell}^2 &= (K_\ell\vec{v}_\ell, A_\ell K_\ell\vec{v}_\ell)_\ell \\
&= \sum_{i=1}^{n_\ell}\left(\mu_\ell^{(i)}\right)^2\lambda_\ell^{(i)}\alpha_i^2.
\end{aligned}$$

Since

$$0 \leq \left(\mu_\ell^{(i)}\right)^2 = \left(1 - \frac{\lambda_\ell^{(k)}}{\Lambda_\ell}\right)^2 \leq 1,$$

we have

$$\left\|K_\ell \vec{v}_\ell\right\|_{A_\ell}^2 \leq \left\|\vec{v}_\ell\right\|_{A_\ell}^2.$$

Hence

$$\left\|K_\ell \vec{v}_\ell\right\|_{A_\ell} \leq \left\|\vec{v}_\ell\right\|_{A_\ell}.$$

2.

$$\begin{aligned}
(K_\ell \vec{v}_\ell, \vec{v}_\ell)_\ell &= \left(K_\ell \sum_{k=1}^{n_\ell} \alpha_k \vec{\omega}_\ell^{(k)}, \sum_{k=1}^{n_\ell} \alpha_k \vec{\omega}_\ell^{(k)}\right)_\ell \\
&= \left(\sum_{k=1}^{n_\ell} \alpha_k K_\ell \vec{\omega}_\ell^{(k)}, \sum_{k=1}^{n_\ell} \alpha_k \vec{\omega}_\ell^{(k)}\right)_\ell \\
&= \left(\sum_{k=1}^{n_\ell} \alpha_k \mu_\ell^{(k)} \vec{\omega}_\ell^{(k)}, \sum_{k=1}^{n_\ell} \alpha_k \vec{\omega}_\ell^{(k)}\right)_\ell \\
&= \sum_{k=1}^{n_\ell} \alpha_k^2 \mu_\ell^{(k)} \\
&\leq \sum_{k=1}^{n_\ell} \alpha_k^2 \\
&= (\vec{v}_\ell, \vec{v}_\ell)_\ell.
\end{aligned}$$

□

Theorem 4.24. *(Convergence of the one-sided W-cycle with Richardson Smoothing) Suppose that $p \geq 2$ (Two-level) $m_1 = m \geq 1$ and $m_2 = 0$ (once-sided). Suppose, further that Assumption (A0) (Galerkin Condition), Assumption (A3) (Strong Approximation property) hold and the smoothing is done by Richardson's method. Then for any $0 < \gamma < 1$, m can be chosen large enough so that*

$$\left\|\vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right)\right\|_{A_\ell} \leq \gamma \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell}$$

for any $\ell \geq 0$, where

$$A_\ell \vec{u}_\ell^{\,E} = \vec{g}_\ell.$$

Proof. First, observe that Richardson's smoother satisfies Assumption (A5) (Theorem.4.21).

The proof is by induction. The cases $\ell = 0$ and $\ell = 1$ (Theorem.4.22) are clearly true. Now, define $\vec{q}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}}$ such that

$$A_{\ell-1} \vec{q}_{\ell-1} = \vec{r}_{\ell-1}^{\,(1)}.$$

Then

$$\begin{aligned}
\vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right) &= \vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(2)} \\
&= \vec{u}_\ell^{\,E} - \left\{\vec{u}_\ell^{\,(1)} + P_{\ell-1} \vec{q}_{\ell-1}^{\,(1)}\right\} \\
&= \vec{u}_\ell^{\,E} - \left\{\vec{u}_\ell^{\,(1)} + P_{\ell-1} \vec{q}_{\ell-1} - P_{\ell-1} \vec{q}_{\ell-1} + P_{\ell-1} \vec{q}_{\ell-1}^{\,(1)}\right\} \\
&= \vec{u}_\ell^{\,E} - \left(\vec{u}_\ell^{\,(1)} + P_{\ell-1} \vec{q}_{\ell-1}\right) + P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right)
\end{aligned}$$

$$= \vec{u}_\ell^{\,E} - \mathrm{TG}_\ell\left(\vec{g}_\ell, \vec{u}_\ell^{\,(0)}\right) + P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right).$$

Suppose that m satisfies

$$m \geq \left(\frac{C_3 C_5}{\gamma - \gamma^p}\right)^2.$$

Then

$$\begin{aligned}
\left\|\vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right)\right\|_{A_\ell} &\leq \left\|\vec{u}_\ell^{\,E} - \mathrm{TG}_\ell\left(\vec{g}_\ell, \vec{u}_\ell^{\,(0)}\right)\right\|_{A_\ell} + \left\|P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right)\right\|_{A_\ell} \\
&\overset{Thm. 4.22}{\leq} C_3 C_5 m^{-1/2}\left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell} + \left\|P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right)\right\|_{A_\ell}. \qquad (4.11)
\end{aligned}$$

Now, observe that, for any $\vec{\omega}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}}$,

$$\begin{aligned}
\left\|P_{\ell-1}\vec{\omega}_{\ell-1}\right\|_{A_\ell}^2 &= (P_{\ell-1}\vec{\omega}_{\ell-1}, P_{\ell-1}\vec{\omega}_{\ell-1})_{A_\ell} \\
&= (P_{\ell-1}\vec{\omega}_{\ell-1}, A_\ell P_{\ell-1}\vec{\omega}_{\ell-1})_\ell \\
&= \left(\vec{\omega}_{\ell-1}, P_{\ell-1}^T A_\ell P_{\ell-1}\vec{\omega}_{\ell-1}\right)_{\ell-1} \\
&= (\vec{\omega}_{\ell-1}, R_{\ell-1} A_\ell P_{\ell-1}\vec{\omega}_{\ell-1})_{\ell-1} \\
&\overset{(A0)}{=} (\vec{\omega}_{\ell-1}, A_{\ell-1}\vec{\omega}_{\ell-1})_{\ell-1} \\
&= (\vec{\omega}_{\ell-1}, \vec{\omega}_{\ell-1})_{A_{\ell-1}} \\
&= \left\|\vec{\omega}_{\ell-1}\right\|_{A_{\ell-1}}^2.
\end{aligned}$$

In the proof of Theorem.4.8, we showed that

$$\begin{aligned}
\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)} &= E_{\ell-1}^p \vec{q}_{\ell-1} \\
&= E_{\ell-1}^p \Pi_{\ell-1}\left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(1)}\right) \\
&= E_{\ell-1}^p \Pi_{\ell-1} K_\ell^m \left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right).
\end{aligned}$$

Induction Hypothesis: Assume

$$\left\|E_{\ell-1}\vec{\omega}_{\ell-1}\right\|_{A_{\ell-1}} \leq \gamma \left\|\vec{\omega}_{\ell-1}\right\|_{A_{\ell-1}}$$

is true for any $\vec{\omega}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}}$. Therefore

$$\begin{aligned}
\left\|P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right)\right\|_{A_\ell} &= \left\|\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right\|_{A_{\ell-1}} \\
&= \left\|E_{\ell-1}^p \Pi_{\ell-1} K_\ell^m \left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right)\right\|_{A_{\ell-1}} \\
&\overset{ind.hyp.}{\leq} \gamma^p \left\|\Pi_{\ell-1} K_\ell^m \left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right)\right\|_{A_{\ell-1}}.
\end{aligned}$$

Since we are assuming the Galerkin Condition, it follows that

$$\left\|\Pi_{\ell-1}\vec{\omega}_\ell\right\|_{A_{\ell-1}} = \left\|\tilde{\Pi}_\ell \vec{\omega}_\ell\right\|_{A_\ell}.$$

Furthermore,

$$\left\|\tilde{\Pi}_\ell \vec{\omega}_\ell\right\|_{A_\ell}^2 = \left(\tilde{\Pi}_\ell \vec{\omega}_\ell, \tilde{\Pi}_\ell \vec{\omega}_\ell\right)_{A_\ell}$$

$$
\begin{aligned}
&= \left(\tilde{\Pi}_\ell^2 \vec{\omega}_\ell, \vec{\omega}_\ell\right)_{A_\ell} \\
&\stackrel{(A0)}{=} \left(\tilde{\Pi}_\ell \vec{\omega}_\ell, \vec{\omega}_\ell\right)_{A_\ell} \\
&\leq \left\|\tilde{\Pi}_\ell \vec{\omega}_\ell\right\|_{A_\ell} \left\|\vec{\omega}_\ell\right\|_{A_\ell}.
\end{aligned}
$$

So

$$
\left\|\tilde{\Pi}_\ell \vec{\omega}_\ell\right\|_{A_\ell} \leq \left\|\vec{\omega}_\ell\right\|_{A_\ell}. \quad \text{stability} \tag{4.12}
$$

Therefore,

$$
\begin{aligned}
\left\|P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right)\right\|_{A_\ell} &\leq \gamma^p \left\|\Pi_{\ell-1} K_\ell^m \left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right)\right\|_{A_{\ell-1}} \\
&= \gamma^p \left\|\tilde{\Pi}_\ell K_\ell^m \left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right)\right\|_{A_\ell} \\
&\stackrel{(4.12)}{\leq} \gamma^p \left\|K_\ell^m \left(\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right)\right\|_{A_\ell} \\
&\stackrel{Lem. 4.23}{\leq} \gamma^p \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell}
\end{aligned}
$$

Combining this with estimate (4.11), we have

$$
\begin{aligned}
\left\|\vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right)\right\|_{A_\ell} &\leq C_3 C_5 m^{-1/2} \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell} + \left\|P_{\ell-1}\left(\vec{q}_{\ell-1} - \vec{q}_{\ell-1}^{\,(1)}\right)\right\|_{A_\ell} \\
&\leq C_3 C_5 m^{-1/2} \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell} + \gamma^p \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell} \\
&\leq \left(C_3 C_5 m^{-1/2} + \gamma^p\right) \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell} \\
&\leq \left(C_3 C_5 \left(\left(\frac{C_3 C_5}{\gamma - \gamma^p}\right)^2\right)^{-1/2} + \gamma^p\right) \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_\ell} \\
&= \gamma \left\|\vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)}\right\|_{A_{\ell-1}}.
\end{aligned}
$$

□

4.3 Convergence of the V-cycle Algorithm

Definition 4.25. *(Assumption (A6): Second Smoothing Property) We say that the multilevel scheme satisfies the Second Smoothing Property , Assumption (A6), iff there is some $C_6 > 0$ such that*

$$
\left\|\vec{v}_\ell\right\|_\ell^2 \leq \rho_\ell C_6^2 \left(\overline{K}_\ell \vec{v}_\ell, \vec{v}_\ell\right)_\ell, \tag{4.13}
$$

for all $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$ and $\ell \geq 1$, where

$$
\overline{K}_\ell := \left(I_\ell - K_\ell^* K_\ell\right) A_\ell^{-1}.
$$

Lemma 4.26. *Let $J_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ and $J_\ell = J_\ell^*$. Then*

$$(J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell} - \left(J_\ell^2 \vec{v}_\ell, J_\ell^2 \vec{v}_\ell\right)_{A_\ell} \le (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - (J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell}, \tag{4.14}$$

for any $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$

Proof. Since A_ℓ is SPD.

$$\begin{aligned}
0 &\le \left\|\left(I_\ell - J_\ell^2\right)\vec{v}_\ell\right\|_{A_\ell}^2 \\
&= \left(\left(I_\ell - J_\ell^2\right)\vec{v}_\ell, \left(I_\ell - J_\ell^2\right)\vec{v}_\ell\right)_{A_\ell} \\
&= \left(I_\ell \vec{v}_\ell - J_\ell^2 \vec{v}_\ell, I_\ell \vec{v}_\ell - J_\ell^2 \vec{v}_\ell\right)_{A_\ell} \\
&= (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - \left(J_\ell^2 \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} - \left(\vec{v}_\ell, J_\ell^2 \vec{v}_\ell\right)_{A_\ell} + \left(J_\ell^2 \vec{v}_\ell, J_\ell^2 \vec{v}_\ell\right)_{A_\ell} \\
&= (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - (J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell} - (J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell} + \left(J_\ell^2 \vec{v}_\ell, J_\ell^2 \vec{v}_\ell\right)_{A_\ell} \\
&= (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - 2\,(J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell} + \left(J_\ell^2 \vec{v}_\ell, J_\ell^2 \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}$$

So

$$(J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell} - \left(J_\ell^2 \vec{v}_\ell, J_\ell^2 \vec{v}_\ell\right)_{A_\ell} \le (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - (J_\ell \vec{v}_\ell, J_\ell \vec{v}_\ell)_{A_\ell}.$$

□

Lemma 4.27. *For any $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$*

$$(\Pi_{\ell-1} \vec{v}_\ell, \Pi_{\ell-1} \vec{v}_\ell)_{A_{\ell-1}} = (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}. \tag{4.15}$$

Proof. Recall that we always have (Note: this is not from Galerkin Condition, but the definition of $\Pi_{\ell-1}$.)

$$R_{\ell-1} A_\ell = A_{\ell-1} \Pi_{\ell-1},$$

and

$$\tilde{\Pi}_\ell = P_{\ell-1} A_{\ell-1}^{-1} R_{\ell-1} A_\ell = P_{\ell-1} \Pi_{\ell-1}.$$

Then

$$\begin{aligned}
(\Pi_{\ell-1} \vec{v}_\ell, \Pi_{\ell-1} \vec{v}_\ell)_{A_{\ell-1}} &= (\Pi_{\ell-1} \vec{v}_\ell, A_{\ell-1} \Pi_{\ell-1} \vec{v}_\ell)_{\ell-1} \\
&= (\Pi_{\ell-1} \vec{v}_\ell, R_{\ell-1} A_\ell \vec{v}_\ell)_{\ell-1} \\
&= \left(R_{\ell-1}^T \Pi_{\ell-1} \vec{v}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\
&= (P_{\ell-1} \Pi_{\ell-1} \vec{v}_\ell, A_\ell \vec{v}_\ell)_\ell \\
&= \left(\tilde{\Pi}_\ell \vec{v}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\tilde{\Pi}_\ell \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} \\
&= (\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} - \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}$$

□

Theorem 4.28. *Suppose that Assumption Weakened Galerkin Condition (A1), Weakened approximation property (A4) and Second smoothing property (A6) hold. Suppose that $p=1$, $m_1=m_2=m=1$ and $S_\ell = S_\ell^T$. Then*

$$0 \le (E_\ell \vec{u}_\ell, \vec{u}_\ell)_{A_\ell} \le \frac{C_4^2 C_6^2}{C_4^2 C_6^2+1} (\vec{u}_\ell, \vec{u}_\ell)_{A_\ell}$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$.

Proof. Recall, since $P=1$, $m_1=m_2=m=1$ and $S_\ell = S_\ell^T$,

$$E_\ell = K_\ell \left(I_\ell - \tilde{\Pi}_\ell\right) K_\ell + K_\ell P_{\ell-1} E_{\ell-1} \Pi_{\ell-1} K_\ell.$$

In particular, notice that

$$K_\ell^* = I_\ell - S_\ell^T = I_\ell - S_\ell A_\ell = K_\ell.$$

Now, set

$$T_1 := \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{\omega}_\ell, \vec{\omega}_\ell\right)_{A_\ell},$$

and

$$T_2 := (P_{\ell-1} E_{\ell-1} \Pi_{\ell-1} \vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell},$$

where

$$\vec{\omega}_\ell = K_\ell \vec{u}_\ell.$$

Then

$$(E_\ell \vec{u}_\ell, \vec{u}_\ell)_{A_\ell} = T_1 + T_2.$$

Let us first consider T_1:

$$\begin{aligned}
T_1 &= \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{\omega}_\ell, \vec{\omega}_\ell\right)_{A_\ell} \\
&\overset{(A4)}{\le} C_4^2 \rho_\ell^{-1} \left\| A_\ell \vec{\omega}_\ell \right\|_\ell^2 \\
&= C_4^2 \rho_\ell^{-1} \left\| A_\ell K_\ell \vec{u}_\ell \right\|_\ell^2 \\
&\overset{(A6)}{\le} C_4^2 \rho_\ell^{-1} C_6^2 \rho_\ell \left(\overline{K}_\ell A_\ell K_\ell \vec{u}_\ell, A_\ell K_\ell \vec{u}_\ell\right)_\ell \qquad (4.16) \\
&= C_4^2 C_6^2 \left(\left(I_\ell - K_\ell^* K_\ell\right) A_\ell^{-1} A_\ell K_\ell \vec{u}_\ell, A_\ell K_\ell \vec{u}_\ell\right)_\ell \\
&= C_4^2 C_6^2 \left(\left(I_\ell - K_\ell^* K_\ell\right) K_\ell \vec{u}_\ell, A_\ell K_\ell \vec{u}_\ell\right)_\ell \\
&= C_4^2 C_6^2 \left\{ (K_\ell \vec{u}_\ell, A_\ell K_\ell \vec{u}_\ell)_\ell - \left(K_\ell^* K_\ell K_\ell \vec{u}_\ell, A_\ell K_\ell \vec{u}_\ell\right)_\ell \right\} \\
&= C_4^2 C_6^2 \left\{ (K_\ell \vec{u}_\ell, K_\ell \vec{u}_\ell)_{A_\ell} - \left(K_\ell^* K_\ell K_\ell \vec{u}_\ell, K_\ell \vec{u}_\ell\right)_{A_\ell} \right\} \\
&= C_4^2 C_6^2 \left\{ (K_\ell \vec{u}_\ell, K_\ell \vec{u}_\ell)_{A_\ell} - \left(K_\ell^2 \vec{u}_\ell, K_\ell^2 \vec{u}_\ell\right)_{A_\ell} \right\} \\
&\overset{(4.14)}{\le} C_4^2 C_6^2 \left\{ (\vec{u}_\ell, \vec{u}_\ell)_{A_\ell} - (K_\ell \vec{u}_\ell, K_\ell \vec{u}_\ell)_{A_\ell} \right\}.
\end{aligned}$$

Now, we turn to the bound for T_2:

$$\begin{aligned}
T_2 &= (E_{\ell-1} \Pi_{\ell-1} \vec{\omega}_\ell, R_{\ell-1} A_\ell \vec{\omega}_\ell)_{\ell-1} \\
&= (E_{\ell-1} \Pi_{\ell-1} \vec{\omega}_\ell, A_{\ell-1} \Pi_{\ell-1} \vec{\omega}_\ell)_{\ell-1} \\
&= (E_{\ell-1} \Pi_{\ell-1} \vec{\omega}_\ell, \Pi_{\ell-1} \vec{\omega}_\ell)_{A_{\ell-1}}.
\end{aligned}$$

Set

$$\gamma = \frac{C_4^2 C_6^2}{C_4^2 C_6^2 + 1}.$$

The proof proceeds by induction:

$$\begin{aligned}
T_2 &= (E_{\ell-1}\Pi_{\ell-1}\vec{\omega}_\ell, \Pi_{\ell-1}\vec{\omega}_\ell)_{A_{\ell-1}} \\
&\overset{ind.hyp.}{\leq} \gamma\,(\Pi_{\ell-1}\vec{\omega}_\ell, \Pi_{\ell-1}\vec{\omega}_\ell)_{A_{\ell-1}} \\
&\overset{4.15}{=} \gamma\left\{(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} - \left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{\omega}_\ell,\vec{\omega}_\ell\right)_{A_\ell}\right\} \\
&= \gamma\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} - \gamma T_1.
\end{aligned} \tag{4.17}$$

To finish up,

$$\begin{aligned}
(E_\ell\vec{u}_\ell,\vec{u}_\ell)_{A_\ell} &= T_1 + T_2 \\
&= (1-\gamma)T_1 + \gamma T_1 + T_2 \\
&\overset{(4.17)}{\leq} (1-\gamma)T_1 + \gamma T_1 + \gamma\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} - \gamma T_1 \\
&= (1-\gamma)T_1 + \gamma\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} \\
&\overset{(4.16)}{\leq} (1-\gamma)C_4^2C_6^2\left\{(\vec{u}_\ell,\vec{u}_\ell)_{A_\ell} - (\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell}\right\} + \gamma\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} \\
&= \Big(1 - \frac{C_4^2C_6^2}{C_4^2C_6^2+1}\Big)C_4^2C_6^2\left\{(\vec{u}_\ell,\vec{u}_\ell)_{A_\ell} - (\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell}\right\} + \frac{C_4^2C_6^2}{C_4^2C_6^2+1}\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} \\
&= \frac{C_4^2C_6^2}{C_4^2C_6^2+1}\left\{(\vec{u}_\ell,\vec{u}_\ell)_{A_\ell} - (\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell}\right\} + \frac{C_4^2C_6^2}{C_4^2C_6^2+1}\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} \\
&= \gamma\left\{(\vec{u}_\ell,\vec{u}_\ell)_{A_\ell} - (\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell}\right\} + \gamma\,(\vec{\omega}_\ell,\vec{\omega}_\ell)_{A_\ell} \\
&= \gamma\,(\vec{u}_\ell,\vec{u}_\ell)_{A_\ell}.
\end{aligned}$$

□

Lemma 4.29. *Richardson's smoother satisfies Assumption (A6) with $S_\ell = \Lambda_\ell^{-1} I_\ell = S_\ell^T$.*

Proof. Recall

$$\rho_\ell \leq \Lambda_\ell \leq C_R\rho_\ell, \quad \exists C_R \geq 1$$

Then

$$\begin{aligned}
\overline{K}_\ell &= (I_\ell - K_\ell^* K_\ell)A_\ell^{-1} \\
&= \left\{I_\ell - \left(I_\ell - \Lambda_\ell^{-1}A_\ell\right)\left(I_\ell - \Lambda_\ell^{-1}A_\ell\right)\right\}A_\ell^{-1} \\
&= \left(I_\ell - \left\{I_\ell - 2\Lambda_\ell^{-1}A_\ell + \Lambda_\ell^{-2}A_\ell^2\right\}\right)A_\ell^{-1} \\
&= 2\Lambda_\ell^{-1}I_\ell - \Lambda_\ell^{-2}A_\ell.
\end{aligned}$$

Define

$$J_\ell := \rho_\ell C_R \overline{K}_\ell - I_\ell.$$

If we can show that J_ℓ is SPSD w.r.t. $(\cdot,\cdot)_\ell$ then we get (A6) with $C_6^2 = C_R$.

J_ℓ is clearly symmetric w.r.t. $(\cdot,\cdot)_\ell$. Now let $\left\{\vec{v}_\ell^{(1)}, \vec{v}_\ell^{(2)}, \cdots, \vec{v}_\ell^{(n_\ell)}\right\}$ be the orthonormal eigenvector of A_ℓ (w.r.t. $(\cdot,\cdot)_\ell$). Then

$$\begin{aligned}
J_\ell \vec{v}_\ell^{(k)} &= \rho_\ell C_R \overline{K}_\ell \vec{v}_\ell^{(k)} - \vec{v}_\ell^{(k)} \\
&= \rho_\ell C_R \left(2\Lambda_\ell^{-1} I_\ell - \Lambda_\ell^{-2} A_\ell\right) \vec{v}_\ell^{(k)} - \vec{v}_\ell^{(k)} \\
&= 2\rho_\ell C_R \Lambda_\ell^{-1} \vec{v}_\ell^{(k)} - \rho C_R \Lambda_\ell^{-2} \lambda_\ell^{(k)} \vec{v}_\ell^{(k)} - \vec{v}_\ell^{(k)} \\
&= \left(2\rho_\ell C_R \Lambda_\ell^{-1} - \rho_\ell C_R \Lambda_\ell^{-2} \lambda_\ell^{(k)} - 1\right) \vec{v}_\ell^{(k)}.
\end{aligned}$$

Set

$$\eta_\ell^{(k)} = 2\rho_\ell C_R \Lambda_\ell^{-1} - \rho_\ell C_R \Lambda_\ell^{-2} \lambda_\ell^{(k)} - 1.$$

We want to show that $\eta_\ell^{(k)} \geq 0$ for all $1 \leq k \leq n_\ell$.

$$\begin{aligned}
\eta_\ell^{(k)} &= 2C_R \frac{\rho_\ell}{\Lambda_\ell} - C_R \frac{\rho_\ell \lambda_\ell^{(k)}}{\Lambda_\ell^2} - 1 \\
&\geq 2C_R \frac{\rho_\ell}{\Lambda_\ell} - C_R \frac{\rho_\ell}{\Lambda_\ell} - 1 \qquad (-\lambda_\ell^{(k)} \geq -\Lambda_\ell) \\
&= C_R \frac{\rho_\ell}{\Lambda_\ell} - 1 \\
&\geq 1 - 1 = 0. \qquad (C_R \rho_\ell \geq \Lambda_\ell)
\end{aligned}$$

Thus the eigenvalues of J_ℓ, $\eta_\ell^{(k)}$, are all non-negative and J_ℓ is SPSD. This implies

$$0 \leq (J_\ell \vec{v}_\ell, \vec{v}_\ell)_\ell = \rho_\ell C_R \left(\overline{K}_\ell \vec{v}_\ell, \vec{v}_\ell\right)_\ell - (\vec{v}_\ell, \vec{v}_\ell)_\ell,$$

and (A6) follows with $C_6^2 = C_R$. □

4.4 Convergence of the V-Cycle Algorithm

Lemma 4.30. *Suppose that smoothing is done with Richardson's Method, i.e.*

$$S_\ell = \Lambda_\ell^{-1},$$

and

$$\rho_\ell \leq \Lambda_\ell \leq C_R \rho_\ell, \;\; \exists C_R \geq 1.$$

$$\left(\left(I_\ell - K_\ell\right) K_\ell^{2m} \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} \leq \frac{1}{2m} \left(\left(I_\ell - K_\ell^{2m}\right) \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} \tag{4.18}$$

for any $m \geq 1$ and $\ell \geq 1$.

Proof. Suppose $i, j \in \mathbb{Z}$ with $0 \leq j \leq i$. Then

$$\begin{aligned}
\left((I_\ell - K_\ell) K_\ell^i \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} &= \left(A_\ell (I_\ell - K_\ell) K_\ell^i \vec{v}_\ell, \vec{v}_\ell\right)_\ell \\
&= \Lambda_\ell^{-1} \left(A_\ell^2 K_\ell^i \vec{v}_\ell, \vec{v}_\ell\right)_\ell \\
&= \Lambda_\ell^{-1} \left(K_\ell^i \vec{v}_\ell, A_\ell \vec{v}_\ell\right)_\ell
\end{aligned}$$

$$
\begin{aligned}
&\overset{4.10}{\leq} \Lambda_\ell^{-1}\left(K_\ell^j \vec{v}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\
&= \left((I_\ell - K_\ell)K_\ell^j \vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& 2m\left((I_\ell - K_\ell)K_\ell^{2m}\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} \\
= \;& \underbrace{\left((I_\ell - K_\ell)K_\ell^{2m}\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell} + \cdots + \left((I_\ell - K_\ell)K_\ell^{2m}\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}}_{2m} \\
\overset{\text{(above)}}{\leq} \;& \underbrace{\left((I_\ell - K_\ell)K_\ell^{0}\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}}_{(j=0)} + \underbrace{\left((I_\ell - K_\ell)K_\ell^{1}\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}}_{(j=1)} + \cdots + \underbrace{\left((I_\ell - K_\ell)K_\ell^{2m-1}\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}}_{(j=2m-1)} \\
= \;& \left((I_\ell - K_\ell^{2m})\vec{v}_\ell, \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}
$$

The last equality follows since the sum telescopes. □

Theorem 4.31. *Suppose that Assumptions (A1) and (A4) hold. Suppose $p = 1, m_1 = m_2 = m \geq 1$ and smoothing is done with Richardson's method. Then*

$$
0 \leq (E_\ell \vec{u}_\ell, \vec{u}_\ell)_{A_\ell} \leq \frac{M}{M+m} (\vec{u}_\ell, \vec{u}_\ell)_{A_\ell},
$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$, where

$$
M := \frac{C_4^2 C_R}{2}.
$$

Proof. The proof is similar to that of Theorem.4.28. We begin with an expression for the error propagation matrix :

$$
E_\ell = K_\ell^m (I_\ell - \tilde{\Pi}_\ell) K_\ell^m + K_\ell P_{\ell-1} E_{\ell-1} P_{\ell-1} \Pi_{\ell-1} K_\ell^m,
$$

where

$$
K_\ell = I_\ell - \Lambda_\ell^{-1} A_\ell = K_\ell^*,
$$

and

$$
\rho_\ell \leq \Lambda_\ell \leq C_R \rho_\ell, \quad \exists C_R \geq 1.
$$

As before, set

$$
T_1 := \left((I_\ell - \tilde{\Pi}_\ell)\vec{\omega}_\ell, \vec{\omega}_\ell\right)_{A_\ell},
$$

and

$$
T_2 := (P_{\ell-1} E_{\ell-1} P_{\ell-1} \vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell},
$$

where

$$
\vec{\omega}_\ell = K_\ell^m \vec{u}_\ell.
$$

Then

$$
(E_\ell \vec{u}_\ell, \vec{u}_\ell)_{A_\ell} = T_1 + T_2.
$$

We first estimate T_1:

$$
T_1 = \left((I_\ell - \tilde{\Pi}_\ell)\vec{\omega}_\ell, \vec{\omega}_\ell\right)_{A_\ell}
$$

$$\begin{aligned}
&\overset{A4}{\leq} && C_4^2\rho_\ell^{-1}\left\|A_\ell\vec{\omega}_\ell\right\|_\ell^2 \\
&= && C_4^2\rho_\ell^{-1}\left\|A_\ell K_\ell^m\vec{u}_\ell\right\|_\ell \\
&= && C_4^2\rho_\ell^{-1}\left(A_\ell K_\ell^m\vec{u}_\ell, A_\ell K_\ell^m\vec{u}_\ell\right)_\ell \\
&= && C_4^2\rho_\ell^{-1}\left(A_\ell^2 K_\ell^m\vec{u}_\ell, K_\ell^m\vec{u}_\ell\right)_\ell \qquad (4.19)\\
&= && C_4^2\rho_\ell^{-1}\left(A_\ell K_\ell^m\vec{u}_\ell, K_\ell^m\vec{u}_\ell\right)_{A_\ell} \\
&= && C_4^2\rho_\ell^{-1}\Lambda_\ell\left((I_\ell - K_\ell)K_\ell^m\vec{u}_\ell, K_\ell^m\vec{u}_\ell\right)_{A_\ell} \\
&= && C_4^2\rho_\ell^{-1}\Lambda_\ell\left((I_\ell - K_\ell)K_\ell^{2m}\vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \\
&\overset{(4.18)}{\leq} && \frac{C_4^2\rho_\ell^{-1}\Lambda_\ell}{2m}\left((I_\ell - K_\ell^{2m})\vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \\
&\leq && \frac{C_4^2 C_R}{2m}\left((I_\ell - K_\ell^{2m})\vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \\
&\leq && \frac{M}{m}\left\{(\vec{u}_\ell, \vec{u}_\ell)_{A_\ell} - (\vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell}\right\}.
\end{aligned}$$

Set

$$\gamma := \frac{M}{M+m} \quad (<1).$$

Exactly as before, the induction step yields

$$T_2 \leq \gamma\,(\vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell} - \gamma T_1. \qquad (4.20)$$

Therefore,

$$\begin{aligned}
(E_\ell\vec{u}_\ell, \vec{u}_\ell)_{A_\ell} &= T_1 + T_2 \\
&= (1-\gamma)T_1 + \gamma T_1 + T_2 \\
&\overset{4.20}{\leq} (1-\gamma)T_1 + \gamma T_1 + \gamma\,(\vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell} - \gamma T_1 \\
&\overset{4.19}{\leq} (1-\gamma)\frac{M}{m}\left\{(\vec{u}_\ell, \vec{u}_\ell)_{A_\ell} - (\vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell}\right\} + \gamma\,(\vec{\omega}_\ell, \vec{\omega}_\ell)_{A_\ell} \\
&= \gamma\,(\vec{u}_\ell, \vec{u}_\ell)_{A_\ell}.
\end{aligned}$$

□

Theorem 4.32. *Suppose that hypotheses of either Theorem.4.28 or 4.31 hold, as appropriate. Suppose that $\vec{u}_\ell^{\,E}, \vec{g}_\ell \in \mathbb{R}^{n_\ell}$ satisfy*

$$A_\ell \vec{u}_\ell^{\,E} = \vec{g}_\ell.$$

Then, given any $\vec{u}_\ell^{\,(0)} \in \mathbb{R}^{n_\ell}$,

$$\begin{aligned} \left\| \vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(3)} \right\|_{A_\ell} &= \left\| \vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)} \right) \right\|_{A_\ell} \\ &\le \frac{M}{M+m} \left\| \vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)} \right\|_{A_\ell}, \end{aligned}$$

where

$$M = C_4^2 C_6^2, M = 1, \quad (4.28)$$

or

$$M = \frac{C_4^2 C_R}{2}, M \ge 1, \quad (4.31).$$

Proof. We need only show that

$$\left\| E_\ell \vec{v}_\ell \right\|_{A_\ell} \le \frac{M}{M+m} \left\| \vec{v}_\ell \right\|_{A_\ell},$$

is true for any $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$. Since E_ℓ is SPSD w.r.t. $(\cdot,\cdot)_{A_\ell}$, for $\ell \ge 1$, there is an orthonormal basis of eigenvectors of E_ℓ, $\left\{ \vec{\omega}_\ell^{(1)}, \cdots, \vec{\omega}_\ell^{(n_\ell)} \right\}$, such that

$$E_\ell \vec{\omega}_\ell^{(j)},$$

$$\left(\vec{\omega}_\ell^{(i)}, \vec{\omega}_\ell^{(j)} \right)_{A_\ell} = \delta_{ij},$$

and

$$0 \le \epsilon_\ell^{(1)} \le \epsilon_\ell^{(2)} \le \cdots \le \epsilon_\ell^{(n_\ell)}.$$

Suppose

$$\vec{v}_\ell = \sum_{k=1}^{n_\ell} C_k \vec{\omega}_\ell^{(k)}.$$

Then

$$(E_\ell \vec{v}_\ell, \vec{v}_\ell)_{A_\ell} = \sum_{k=1}^{n_\ell} C_k^2 \epsilon_\ell^{(k)},$$

and

$$(\vec{v}_\ell, \vec{v}_\ell)_{A_\ell} = \sum_{k=1}^{n_\ell} C_k^2.$$

Theorem. 4.31 implies that

$$\sum_{k=1}^{n_\ell} C_k^2 \epsilon_\ell^{(k)} \le \frac{M}{M+m} \sum_{k=1}^{n_\ell} C_k^2,$$

for any $C_1, \cdots, C_{n_\ell} \mathbb{R}$. This implies that

$$0 \le \epsilon_\ell^{(k)} \le \frac{M}{M+m}, \quad 1 \le k \le n_\ell.$$

Therefore

$$
\begin{aligned}
(E_\ell \vec{v}_\ell)_{A_\ell}^2 &= (E_\ell \vec{v}_\ell, E_\ell \vec{v}_\ell)_{A_\ell} \\
&= \sum_{k=1}^{n_\ell} C_k^2 \left(\epsilon_\ell^{(k)}\right)^2 \\
&= \left(\frac{M}{M+m}\right)^2 \sum_{k=1}^{n_\ell} C_k^2 \\
&= \left(\frac{M}{M+m}\right)^2 \left\|\vec{v}_\ell\right\|_{A_\ell}^2 .
\end{aligned}
$$

□

5 Conforming Finite Element Method

Consider the following model problem: Let $\Omega \subset \mathbb{R}^d$, $d = 1$ or $d = 2$, be an open polygonal domain. (Often, we will also assume that Ω is convex.) Given $F \in L^2(\Omega)$, find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = (f,v)_{L^2(\Omega)} \tag{5.1}$$

where

$$a(u,v) := (\nabla u, \nabla v)_{L^2(\Omega)}.$$

Let $\mathcal{T}_0$ be a triangulation of Ω. We define $\mathcal{T}_1$ to the triangulation of Ω that results from bisecting (d=1) or quadrisecting (d=2) the triangulation $\mathcal{T}_0$. For $d = 2$, we connect the edge midpoints.

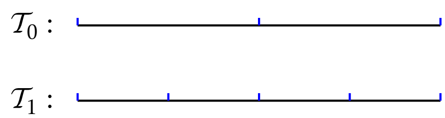


Figure 23: Bisecting the triangulation $\mathcal{T}_0$ in 1D.

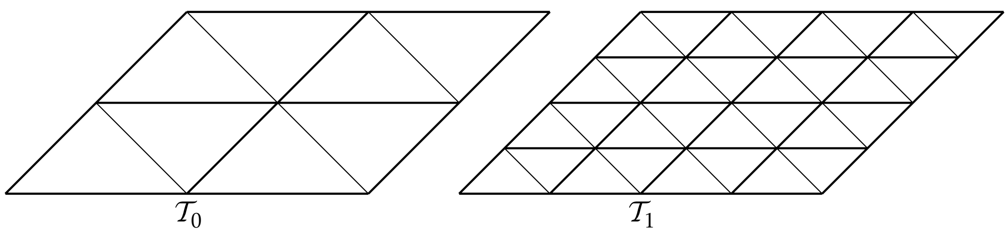


Figure 24: Quadrisecting the triangulation $\mathcal{T}_0$ in 2D.

Observe that the daughter triangles are similar to the mother. Continuing, we can recursively define a family of the triangulations, indexed by ℓ,

$$\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \cdots, \mathcal{T}_\ell, \cdots, \mathcal{T}_L.$$

Definition 5.1. *For $0 \le \ell \le L$, define*

$$h_\ell := \max_{K \in \mathcal{T}_\ell} \operatorname{diam}(K).$$

Subordinate to $\mathcal{T}_\ell$, define

$$V_\ell := \left\{ v_\ell \in C^0(\overline{\Omega} | v_\ell|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_\ell, v_\ell|_{\partial\Omega} \equiv 0 \right\}.$$

Define

$$n_\ell := \dim V_\ell$$

Lemma 5.2. *For $0 \le \ell \le L$ and $v_\ell, \mathcal{T}_\ell$ defined as above, we have*

$$V_1 \subset V_2 \subset \cdots V_\ell \subset V_L \subset H_0^1(\Omega),$$

and

$$0 \le n_0 < n_1 < n_2 < \cdots < n_\ell < \cdots < n_L < \infty.$$

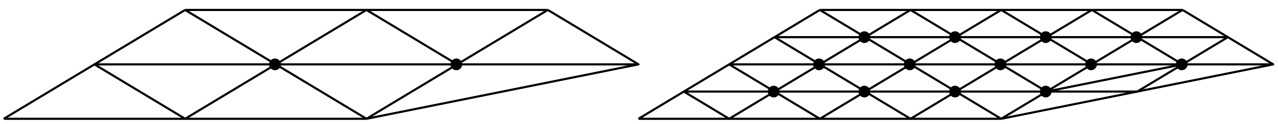
Furthermore, n_ℓ is precisely the number of interior vertices of the triangulation $\mathcal{T}_\ell$. Finally,

$$h_{\ell-1} = 2h_\ell, \quad 1 \le \ell \le L.$$

Proof. Exercise. □

Remark 5.3. *These ideas can be extended to $d = 3$ as well, but we neglect this case for simplicity.*

Example. *The Recursively Mesh in 1D and 2D.*



$\mathcal{T}_0 : n_0 = 2$ $\mathcal{T}_1 : n_1 = 13.$

Figure 25: Recursively Mesh in 2D.

Remark 5.4. *We will assume that $n_0 > 0$.*

Definition 5.5. *Let $0 \le \ell \le L$ and $V_\ell, \mathcal{T}_\ell$ be defined as above. By B_ℓ we denote the Lagrange Nodal basis for V_ℓ, i.e.*

$$B_\ell := \left\{\psi_{\ell,i}\right\}_{i=1}^{n_\ell},$$

where $\psi_{\ell,i} \in V_\ell$ is the unique function with the property that

$$\psi_{\ell,i}(\vec{N}_{\ell,j}) = \delta_{ij},$$

and $\left\{N_{\ell,j}\right\}_{j=1}^{n_\ell}$ is the set of interior vertices of the triangulation $\mathcal{T}_\ell$.

Example. *Example of $N_{\ell,j}$ and $\psi_{\ell,i}$ in 2D.*

Lemma 5.6. *B_ℓ us a bona fide basis for V_ℓ, $0 \le \ell \le L$. And for every $1\ell \le L$, there exist unique numbers*

$$p_{\ell-1,j,i} \in \mathbb{R}, \ 1 \le j \le n_\ell, \ 1 \le i \le n_{\ell-1}$$

with the property that

$$\psi_{\ell-1,i} = \sum_{j=1}^{n_\ell} p_{\ell-1,j,i}\psi_{\ell,j}, \tag{5.2}$$

for each $1 \le i \le n_{\ell-1}$.

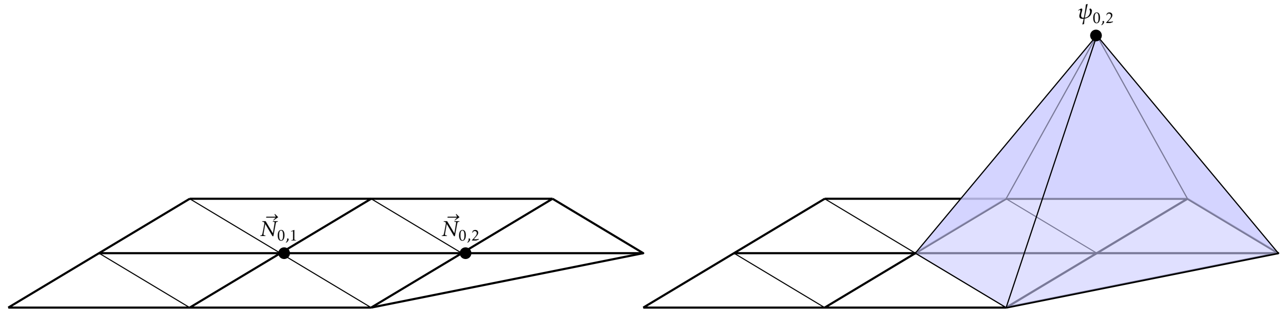


Figure 26: Example of $N_{\ell,j}$ and $\psi_{\ell,i}$ in 2D.

Proof. Since $V_{\ell-1}$ is a linear subspace of V_ℓ and B_ℓ is a basis for the latter, for every $\psi_{\ell-1,i} \in B_{\ell-1} \subset V_{\ell-1} \subset V_\ell$, there exists unique coefficients

$$p_{\ell-1,j,i} \in \mathbb{R}, \ \ 1 \le j \le n_\ell,$$

such that

$$\psi_{\ell-1,i} = \sum_{j=1}^{n_\ell} p_{\ell-1,j,i}\psi_{\ell,j}.$$

Recall that basis representation are unique. □

Definition 5.7. *For $0 \le \ell \le L$, define the prolongation (or injection) matrix*

$$P_{\ell-1} \in \mathbb{R}^{n_\ell \times n_{\ell-1}},$$

via

$$[P_{\ell-1}]_{i,j} := p_{\ell-1,i,j}$$

Lemma 5.8. *Let $1 \le \ell \le L$ and $v_{\ell-1} \in V_{\ell-1}$ be arbitrary. Suppose $\vec{v}_{\ell-1}$ is the coordinate vector of $v_{\ell-1}$ in the Lagrange Nodal basis $B_{\ell-1}$. Suppose $\vec{v}_\ell$ is the coordinate vector of $v_{\ell-1}$ in B_ℓ. (Recall $v_{\ell-1} \in v_\ell$.) Then*

$$v_\ell = P_{\ell-1}v_{\ell-1}.$$

Proof. We write

$$[\vec{v}_{\ell-1}]_i = v_{\ell-1,i}, \ \ 1 \le i \le n_{\ell-1},$$

and

$$[\vec{v}_\ell]_i = v_{\ell,i}, \ \ 1 \le i \le n_\ell.$$

Thus

$$\begin{aligned}
v_{\ell-1} &= \sum_{i=1}^{n_{\ell-1}} v_{\ell-1,i}\psi_{\ell-1,i} \\
&\stackrel{5.2}{=} \sum_{i=1}^{n_{\ell-1}} v_{\ell-1,i} \sum_{j=1}^{n_\ell} p_{\ell-1,j,i}\psi_{\ell,j} \\
&= \sum_{j=1}^{n_\ell} \left\{ \sum_{i=1}^{n_{\ell-1}} p_{\ell-1,j,i}v_{\ell-1,i} \right\} \psi_{\ell,j}
\end{aligned}$$

$$= \sum_{j=1}^{n_\ell} [P_{\ell-1}\vec{v}_{\ell-1}]_j \, \psi_{\ell,j}.$$

But

$$v_{\ell-1} = \sum_{j=1}^{n_\ell} [\vec{v}_\ell]_j \, \psi_{\ell,j}.$$

Since basis representations are unique,

$$\vec{v}_\ell = P_{\ell-1}\vec{v}_{\ell-1}.$$

□

Definition 5.9. *For $1 \le \ell \le L$, define the Restriction matrix as*

$$R_{\ell-1} := P_{\ell-1}^T \in \mathbb{R}^{n_{\ell-1}\times n_\ell}.$$

Lemma 5.10. *For $1 \le \ell \le L$ and suppose $R_{\ell-1}$ and $P_{\ell-1}$ are defined as above. Then*

$$\mathrm{rank}(P_{\ell-1}) = \mathrm{rank}(R_{\ell-1}) = n_{\ell-1}.$$

Proof. Suppose

$$P_{\ell-1}\vec{v}_{\ell-1} = \vec{0} \in \mathbb{R}^{n_\ell}.$$

This represents a linear combination of the $n_{\ell-1}$ columns of $P_{\ell-1}$. Using the notation from the last lemma and its proof, we have

$$\vec{v}_\ell = \vec{0} \in \mathbb{R}^{n_\ell},$$

where $\vec{v}_{\ell-1}$ and $\vec{v}_\ell$ are coordinate vectors of some function $\vec{v}_{\ell-1} \in V_{\ell-1}$ in the basis $B_{\ell-1}$ and B_ℓ, respectively.

The only way that $\vec{v}_\ell = \vec{0}$ is if $v_{\ell-1} \equiv 0$ in $V_{\ell-1}$. But then $\vec{v}_{\ell-1} = \vec{0}$. Thus, the columns of $P_{\ell-1}$ are linearly independent and

$$\mathrm{rank}(P_{\ell-1}) = n_{\ell-1}.$$

□

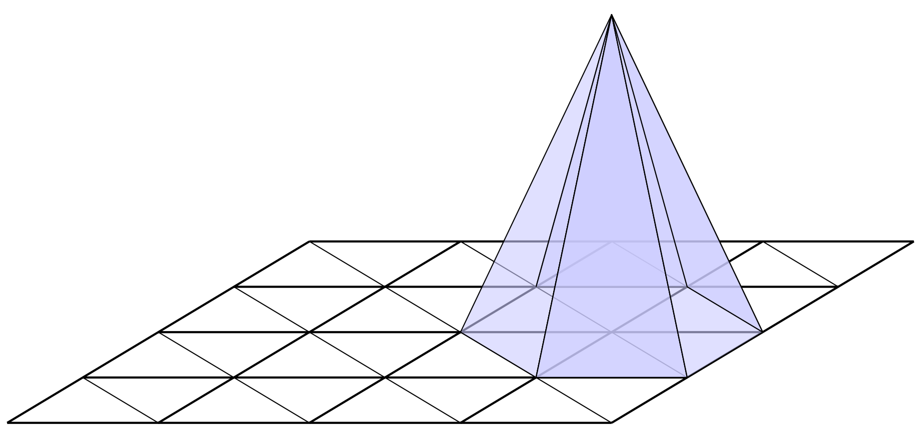


Figure 27: Mesh on the fine and coarse grid of the multigrid method in 2D.

5.1 The Stiffness Matrices

Definition 5.11. *For $\ell \geq 0$ and $V_\ell, \mathcal{T}_\ell$ defined as before, we define the Stiffness Matrices $A_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ via*

$$[A_\ell]_{ij} := a\left(\psi_{\ell,j}, \psi_{\ell,i}\right)$$

for all $1 \leq i, j \leq n_\ell$.

Lemma 5.12. *The Stiffness Matrices $A_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ are all SPD, and , moreover, for $1 \leq \ell \leq L$,*

$$A_{\ell-1} = R_{\ell-1} A_\ell P_{\ell-1}. \tag{5.3}$$

for all $1 \leq i, j \leq n_\ell$.

Proof. 1. Symmetric-ness:

$$\begin{aligned} [A_\ell]_{ij} &= a\left(\psi_{\ell,j}, \psi_{\ell,i}\right) \\ &= a\left(\psi_{\ell,i}, \psi_{\ell,j}\right) \\ &= [A_\ell]_{ji}. \end{aligned}$$

2. PD-ness: Let $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$ be arbitrary and suppose $v_\ell V_\ell$ is the unique function with coordinations $\vec{v}_\ell$. Then, using the Poincare inequality, there is a $C > 0$ such that

$$\begin{aligned} 0 \leq C \|v_\ell\|^2_{L^2(\Omega)} &\leq a(v_\ell, v_\ell) \\ &= a\left(\sum_{j=1}^{n_\ell} v_{\ell,j}\psi_{\ell,j}, \sum_{j=1}^{n_\ell} v_{\ell,i}\psi_{\ell,i}\right) \\ &= \sum_{i,j=1}^{n_\ell} v_{\ell,i}\, a\left(\psi_{\ell,j}, \psi_{\ell,i}\right) v_{\ell,j} \\ &= \vec{v}_\ell^T A_\ell \vec{v}_\ell. \end{aligned}$$

But $v_\ell \equiv 0$ iff $\vec{v}_\ell = \vec{0}$. Hence A_ℓ is SPD.

3. Galerkin Condition: By defintion

$$\begin{aligned} [A_{\ell-1}]_{ij} &= a\left(\psi_{\ell-1,j}, \psi_{\ell-1,i}\right) \\ &\overset{5.2}{=} a\left(\sum_{J=1}^{n_\ell} p_{\ell-1,J,j}\psi_{\ell,J}, \sum_{I=1}^{n_\ell} p_{\ell-1,I,i}\psi_{\ell,I}\right) \\ &= \sum_{I,J=1}^{n_\ell} p_{\ell-1,I,i}\, a(\psi_{\ell,J}, \psi_{\ell,I})\, p_{\ell-1,J,j} \\ &= \sum_{I,J=1}^{n_\ell} [R_{\ell-1}]_{iI} [A_\ell]_{I,J} [P_{\ell-1}]_{J,j}. \end{aligned}$$

Thus

$$A_{\ell-1} = R_{\ell-1} A_\ell P_{\ell-1}.$$

□

Remark 5.13. *Observe that this last result guarantees that $P_{\ell-1}$ and $R_{\ell-1}$ are of full rank. Otherwise $A_{\ell-1}$ could not be positive definite.*

Next we estimate the size of the condition number of A_ℓ. To do this we need a couple of results whose proof can be found in Finite Element books.

Lemma 5.14. *Suppose $d = 1, 2, or, 3$. let $v_\ell \in V_\ell$ be arbitrary and $\vec{v}_\ell \in \mathbb{R}^{n_\ell}$ be its unique coordinate vector in the basis B_ℓ. Then there are constants $C_2 \geq C_1 > 0$, independent of $\ell \geq 0$ and v_ℓ, such that*

$$C_1 h_\ell^d \left\|\vec{v}_\ell\right\|_2^2 \leq \|v_\ell\|_{L^2(\Omega)}^2 \leq C_2 h_\ell^d \left\|\vec{v}_\ell\right\|_2^2. \tag{5.4}$$

Lemma 5.15. *There is a constant $C > 0$ independent of $\ell \geq 0$ such that*

$$a\left(v_\ell, v_\ell\right) \leq C_3 h_\ell^{-2} \|v_\ell\|_{L^2}^2. \tag{5.5}$$

And, as a consequence of (5.4),

$$a\left(v_\ell, v_\ell\right) \leq C_2 C_3 h_\ell^{d-2} \left\|\vec{v}_\ell\right\|_2^2. \tag{5.6}$$

Remark 5.16. *These results require some conditions on the underlying family of conforming meshes. such as global quasi-uniformity and shape regularity, which hold thanks to our construction of the family $\mathcal{T}_\ell$, $\ell \geq 0$.*

Estimate (5.4) is a norm equivalence, (5.5) is called inverse estimate.

Lemma 5.17. *(Asymptotic Sharpness of (5.5)) Let $V_\ell, \mathcal{T}_\ell$ be defined as usual. There exist a constant $C_4 > 0$, independent of ℓ , and functions $v'_\ell \in V_\ell$, such that, for every $0 \leq \ell \leq L$*

$$C_4 h_\ell^{-2} \left\|v'_\ell\right\|_{L^2}^2 \leq a\left(v'_\ell, v'_\ell\right). \tag{5.7}$$

Moreover, the Poincare inequality is asymptotic sharp in a similar sense: There exists a constant $C_5 > 0$, independent of ℓ, and there exist functions $v''_\ell \in V_\ell$, such that,for every $0 \leq \ell \leq L$

$$a\left(v''_\ell, v''_\ell\right) \leq C_5 \left\|\vec{v}''_\ell\right\|_{L^2}^2. \tag{5.8}$$

Remark 5.18. *Proofs of these results can be found in Brass's Finite Element book [1].*

Theorem 5.19. *Let $d = 1, 2, or 3$. There exist constants $C_7 \geq C_6 > 0$, independent of $\ell \geq 0$, such that*

$$C_6 h_\ell^{-2} \leq \kappa_2(A_\ell) = \frac{\lambda_\ell^{(n_\ell)}}{\lambda_\ell^{(1)}} \leq C_7 h_\ell^{-2}. \tag{5.9}$$

In particular, there are constant $C_7^{(i)}, C_6^{(i)} > 0$ for $i = 1, n_\ell$, such that

$$C_6^{(n_\ell)} h_\ell^{d-2} \leq \lambda_\ell^{(n_\ell)} \leq C_7^{(n_\ell)} h_\ell^{d-2}.$$
$$C_7^{(1)} h_\ell^{d} \leq \lambda_\ell^{(1)} \leq C_6^{(1)} h_\ell^{-d}.$$

Proof. First we recall some basis facts for the Rayleigh Quotation.

$$\lambda_\ell^{(1)} = R(\vec{v}_\ell^{(1)}) = \min_{\vec{v}_\ell} R(\vec{v}_\ell) > 0,$$

and

$$\lambda_\ell^{(n_\ell)} = R(\vec{v}_\ell^{(n_\ell)}) = \max_{\vec{v}_\ell} R(\vec{v}_\ell),$$

where

$$A_\ell \vec{v}_\ell^{\,(k)} = \lambda_\ell^{(k)} \vec{v}_\ell^{\,(k)}, \ \ 1 \le k \le n_\ell.$$

Upper bound in (5.9): As usual let $v_\ell \in V_\ell, v_\ell \in \mathbb{R}^{n_\ell}$. Then, for arbitrary $v_\ell \in V_\ell$,

$$\begin{aligned} R(\vec{v}_\ell) : &= \frac{\vec{v}_\ell^T A_\ell \vec{v}_\ell}{\vec{v}_\ell^T \vec{v}_\ell} = \frac{a(v_\ell, v_\ell)}{\|\vec{v}_\ell\|_2^2} \\ &\overset{5.6}{\le} \frac{C_2 C_3 h_\ell^{d-2} \|\vec{v}_\ell\|_2^2}{\|\vec{v}_\ell\|_2^2} \\ &=: C_7^{(n_\ell)} h_\ell^{d-2}. \end{aligned}$$

This implies

$$\lambda_\ell^{(n_\ell)} \le C_7^{(n_\ell)} h_\ell^{d-2}.$$

Similarly,

$$\begin{aligned} R(\vec{v}_\ell) : &= \frac{a(v_\ell, v_\ell)}{\|\vec{v}_\ell\|_2^2} \\ &\overset{\text{Poincare}}{\ge} \frac{C_p \|\vec{v}_\ell\|_{L^2(\Omega)}^2}{\|\vec{v}_\ell\|_2^2} \\ &\overset{5.4}{\ge} \frac{C_p C_1 h_\ell^d \|\vec{v}_\ell\|_2^2}{\|\vec{v}_\ell\|_2^2} \\ &=: C_7^{(1)} h_\ell^d. \end{aligned}$$

Therefore,

$$\lambda_\ell^{(1)} \ge C_7^{(1)} h_\ell^d.$$

It follows that

$$\begin{aligned} \lambda_\ell^{(n_\ell)} &= R(\vec{v}_\ell^{(n_\ell)}) \ge R(\vec{v}_\ell') \\ &= \frac{a\left(v_\ell', v_\ell'\right)}{\|\vec{v}\,'_\ell\|_2^2} \\ &\overset{5.7}{\ge} \frac{C_4 h_\ell^{-2} \|v_\ell'\|_{L^2}^2}{\|\vec{v}\,'_\ell\|_2^2} \\ &\overset{5.4}{\ge} \frac{C_1 C_4 h_\ell^{d-2} \|v_\ell'\|_2^2}{\|\vec{v}\,'_\ell\|_2^2} \\ &=: C_6^{(n_\ell)} h_\ell^{d-2}. \end{aligned}$$

Likewise

$$\lambda_\ell^{(1)} = R(\vec{v}_\ell^{(1)}) \ge R(\vec{v}\,''_\ell)$$

$$
\begin{aligned}
&= \frac{a\left(v_\ell'', v_\ell''\right)}{\left\|\vec{v}''_\ell\right\|_2^2} \\
&\overset{5.8}{\leq} \frac{C_5 \left\|v_\ell''\right\|_{L^2}^2}{\left\|\vec{v}''_\ell\right\|_2^2} \\
&\overset{5.4}{\leq} \frac{C_2 C_5 h_\ell^d \left\|v_\ell''\right\|_2^2}{\left\|\vec{v}''_\ell\right\|_2^2} \\
&=: C_6^{(1)} h_\ell^d.
\end{aligned}
$$

It then follows that

$$
\kappa_2(A_\ell) \geq \frac{C_6^{(n_\ell)} h_\ell^{d-2}}{C_6^{(1)} h_\ell^d} = C_6 h_\ell^{-2}.
$$

□

5.2 Strong Approximation Property

Now, we want to show that the strong approximate property, Assumption (A3), holds for the correct situation: There is some C_{A3} such that

$$
\left\|\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell\right\|_\ell^2 \leq C_{A3}^2 \rho_\ell^{-1} \left\|\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell\right\|_{A_\ell}^2. \tag{5.10}
$$

We need a bot of PDE and FE theory first.

Let $f \in H^{-1}(\Omega) = \left(H_0^1(\Omega)\right)'$ be given. Find $u \in H_0^1(\Omega)$ such that

$$
a(u,v) = (f,v) =: \langle f,v \rangle, \quad \forall \; v \in H_0^1(\Omega).
$$

It is well-known that a unique solution $u \in H_0^1(\Omega)$ exists.

A conforming FE approximation of the problem may be written as follows: Find $\vec{u}_\ell \in V_\ell$ such that

$$
a(\vec{u}_\ell, \vec{v}_\ell) = f(v_\ell), \quad \forall \vec{v}_\ell \in V_\ell, \tag{5.11}
$$

where V_ℓ is the family of nested conforming FE subspaces of $H_0^1(\Omega)$ that we constructed easier.

Observe that every $f \in L^2(\Omega)$ gives figure out what's at here to an $L_f \in H^{-1}$ in a nature way:

$$
L_f(v) := (f,v)_{L^2(\Omega)}, \quad \forall \; v \in H_0^1(\Omega).
$$

In this case we write

$$
\langle f,v \rangle = (f,v)_{L^2(\Omega)},
$$

and

$$
L_f = f.
$$

Definition 5.20. *We say that the model problem satisfies the standard regularity condition iff when $f \in L^2(\Omega) \cap H^{-1}(\Omega)$, then $u \in H_0^1(\Omega) \cap H^2(\Omega)$ and*

$$
|u|_{H^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}, \tag{5.12}
$$

for some universal (regularity) constant $C > 0$.

Theorem 5.21. *If Ω is convex and polyhedral, then the standard regularity condition holds.*

Theorem 5.22. *Let $\Omega \subset \mathbb{R}^d, d = 1,2,$or 3, be an open polyhedral domain and suppose $\mathcal{T}_h$ is a shape regular family of triangulation of Ω parameterized by*

$$h := \max_{K \in \mathcal{T}_h} \mathrm{diam}(K).$$

Then, if Ω is convex (so that the standard regularity condition holds) and $f \in H^{-1}(\Omega)$, there is a constant $C > 0$, such that

$$\|u - u_h\|_{L^2(\Omega)} \leq ch|u - u_h|_{H^1(\Omega)}, \tag{5.13}$$

where $u_h \in V_h = \left\{ v \in C^0(\Omega) | \quad v|_K \in \mathbb{P}_1(K), K \in \mathcal{T}_h, v|_{\partial\Omega \equiv 0} \right\}$ is the unique solution to

$$a(u_h, v_h) = f(v_h), \quad v_h \in V_h.$$

If, in addition, it is known that $f \in L^2(\Omega)$, so that $u \in H_0^1(\Omega) \cap H^2(\Omega)$, then

$$|u - u_h|_{H^1(\Omega)} \leq ch|u|_{H^2(\Omega)}, \tag{5.14}$$

for some $C > 0$. And, all together,

$$\|u - u_h\|_{L^2(\Omega)} \leq ch^2|u|_{H^2(\Omega)}. \tag{5.15}$$

Definition 5.23. *Let $\mathcal{T}_h$ and V_h be as in the last theorem. Let $u \in H_0^1(\Omega)$ be arbitrary. Define the Ritz projection as follows: $\mathcal{R}_h u \in V_h$ is the unique solution to*

$$a(\mathcal{R}_h u, v_h) = a(u, v_h), \ \forall \ v_h \in V_h.$$

in the case where $V_h = V_\ell, \mathcal{T}_h = \mathcal{T}_\ell$, we write $\mathcal{R}_h =: \mathcal{R}_\ell$:

$$a(\mathcal{R}_\ell u, v_\ell) = a(u, v_\ell), \ \forall \ v_\ell \in V_h.$$

Remark 5.24.

$$a(u, v) = (-\Delta u, v)_{L^2}, \ \forall \ v \in H_0^1(\Omega).$$

In any case, it should be clear that $\mathcal{R}_h u \in V_h$ is just the FE approximation of u.

Lemma 5.25. *Let $\mathcal{T}_\ell$ and V_ℓ as usual, and suppose $\vec{u}_\ell \in V_\ell$ is given. Then, Ω is convex, as follows: $\mathcal{R}_h u \in V_h$ is the unique solution to*

$$\left\| \vec{u}_\ell - \mathcal{R}_{\ell-1} \vec{u}_\ell \right\|_{L^2} \leq Ch_\ell \left| \vec{u}_\ell - \mathcal{R}_{\ell-1} \vec{u}_\ell \right|_{H^1}, \tag{5.16}$$

for some constant $C > 0$ that is independent of $\ell \geq 0$.

Proof. Observe that $\vec{u}_\ell \in V_\ell \subset H_0^1(\Omega)$. But $\vec{u}_\ell \notin H^2(\Omega)$. $\vec{u}_\ell$ plays the role of the exact PDE solution in Theorem.(5.22), but it is not H^2-regular. But this does not matter. We may still apply (5.13), since Ω is convex, to conclude

$$\left\| \vec{u}_\ell - \mathcal{R}_{\ell-1} \vec{u}_\ell \right\|_{L^2} \leq Ch_{\ell-1} \left| \vec{u}_\ell - \mathcal{R}_{\ell-1} \vec{u}_\ell \right|_{H^1}$$

for some $C > 0$ that is independent of ℓ. Now, note that

$$h_{\ell-1} = 2h_\ell,$$

and the result follows. □

Theorem 5.26. *Let $\mathcal{T}_\ell$ and V_ℓ as usual, and suppose that Ω is convex polyhedral. Then the strong approximate property is satisfied. In particular, there is some $C_{A3} > 0$, independent of ℓ, such that*

$$\left\|\vec{u}_\ell - \tilde{\Pi}_\ell \vec{u}_\ell\right\|_\ell^2 \leq C_{A3}^2 \rho_\ell^{-1} \left\|\vec{u}_\ell - \tilde{\Pi}_\ell \vec{u}_\ell\right\|_{A_\ell}^2 \tag{5.17}$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$.

Proof. Let $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$ be arbitrary. Suppose $\vec{u}_\ell \in V_\ell$ is the unique element whose coordinate vector is $\vec{u}_\ell$ with basis $\mathcal{B}_\ell$.

$$\vec{u}_\ell \in V_\ell \overset{\mathcal{B}_\ell}{\leftrightarrow} \vec{u}_\ell \mathbb{R}^{n_\ell}.$$

Referring to (5.16)

$$\left|\vec{u}_\ell - \mathcal{R}_{\ell-1}\vec{u}_\ell\right|_{H^1}^2 = a\left(\vec{u}_\ell - \mathcal{R}_{\ell-1}\vec{u}_\ell, \vec{u}_\ell - \mathcal{R}_{\ell-1}\vec{u}_\ell\right),$$

let $\vec{w}_\ell \in \mathbb{R}^{n_\ell}$ be the unique coordinate vector of

$$\underset{\in V_\ell}{\vec{u}_\ell} - \underset{\in V_{\ell-1}\subset V_\ell}{\mathcal{R}_{\ell-1}\vec{u}_\ell} \in V_\ell,$$

w.r.t. the Lagrange nodal basis $\mathcal{B}_\ell$. We want to show that

$$\vec{w}_\ell = \vec{u}_\ell - \tilde{\Pi}_\ell \vec{u}_\ell = \vec{u}_\ell - P_{\ell-1} A_{\ell-1}^{-1} R_{\ell-1} A_\ell \vec{u}_\ell.$$

We begin with the definition of $\mathcal{R}_{\ell-1}$:

$$a\left(\mathcal{R}_{\ell-1}\vec{u}_\ell, v_{\ell-1}\right) = a\left(\vec{u}_\ell, v_{\ell-1}\right), \quad \forall v_{\ell-1} \in V_{\ell-1}.$$

Set $u'_{\ell-1} := \mathcal{R}_{\ell-1}\vec{u}_\ell \in V_{\ell-1}$

$$\vec{u}'_{\ell-1} \in \mathbb{R}^{n_{\ell-1}} \overset{\mathcal{B}_{\ell-1}}{\leftrightarrow} \vec{u}'_{\ell-1} \in V_{\ell-1}.$$
$$\vec{v}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}} \leftrightarrow \vec{v}_{\ell-1} \in V_{\ell-1}.$$

Then

$$\begin{aligned}\left(\vec{u}'_{\ell-1}, \vec{v}_{\ell-1}\right)_{A_{\ell-1}} &= \left(\vec{u}_\ell, P_{\ell-1}\vec{v}_{\ell-1}\right)_{A_\ell} \\ &= \left(A_\ell\vec{u}_\ell, P_{\ell-1}\vec{v}_{\ell-1}\right)_\ell \\ &= \left(R_{\ell-1}A_\ell\vec{u}_\ell, \vec{v}_{\ell-1}\right)_{\ell-1}.\end{aligned}$$

But

$$\left(\vec{u}'_{\ell-1}, \vec{v}_{\ell-1}\right)_{A_{\ell-1}} = \left(A_{\ell-1}\vec{u}'_{\ell-1}, \vec{v}_{\ell-1}\right)_{\ell-1}.$$

So, it follows that

$$A_{\ell-1}\vec{u}'_{\ell-1} = R_{\ell-1}A_\ell\vec{u}_\ell.$$

Thus

$$\vec{u}'_{\ell-1} = A_{\ell-1}^{-1}R_{\ell-1}A_\ell\vec{u}_\ell = \Pi_{\ell-1}\vec{u}_\ell.$$

Therefore

$$\begin{aligned}\vec{w}_\ell &= \vec{u}_\ell - P_{\ell-1}\vec{u}'_{\ell-1} \\ &= \vec{u}_\ell - \tilde{\Pi}_\ell\vec{u}_\ell.\end{aligned}$$

It follows that

$$\begin{aligned}\left|\vec{u}_\ell - R_{\ell-1}\vec{u}_\ell\right|_{H^1}^2 &= (\vec{w}_\ell, A_\ell \vec{w}_\ell)_\ell \\ &= (\vec{w}_\ell, \vec{w}_\ell)_{A_\ell} \\ &= \left\|\vec{w}_\ell\right\|_{A_\ell}^2 \\ &= \left\|\vec{u}_\ell - \tilde{\Pi}_\ell \vec{u}_\ell\right\|_{A_\ell}^2.\end{aligned}$$

Finnally, using the norm equivalence in (5.4)

$$\begin{aligned}C_1 h_\ell^d \left\|\vec{u}_\ell - \tilde{\Pi}_\ell \vec{u}_\ell\right\|_\ell^2 &\overset{5.4}{\leq} \left\|\vec{u}_\ell - R_{\ell-1}\vec{u}_\ell\right\|_{L^2}^2 \\ &\overset{5.4}{\leq} \left|\vec{u}_\ell - R_{\ell-1}\vec{u}_\ell\right|_{H^1}^2 \\ &= Ch_\ell^2 \left\|\vec{u}_\ell - \tilde{\Pi}_\ell \vec{u}_\ell\right\|_{A_\ell}^2.\end{aligned}$$

The proof of Theorem.(5.19) showed that

$$C_6^{n_\ell} h_\ell^{d-2} \leq \rho_\ell \leq C_7^{n_\ell} h_\ell^{d-2},$$

combing this with the last estimate gives (5.17). □

Corollary 5.27. *Let $\mathcal{T}_\ell$ and V_ℓ as usual with A_ℓ the standard stiffness matrix for the model problem. Then, the weak approximation property, Assumption (A4) holds: $\exists\ C_{A4}$ such that*

$$\left(\left(I_\ell - \tilde{\Pi}_\ell\right)\vec{u}_\ell, \vec{u}_\ell\right)_{A_\ell} \leq C_{A4}^2 \rho_\ell^{-1} \left\|A_\ell \vec{u}_\ell\right\|_\ell^2 \tag{5.18}$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$.

Proof. This follows immediately from Theorem.(4.18). Since the Galerkin condition (A0) and the strong approximate property hold. □

Remark 5.28. *Therefore, using Richardson's smoother, the W and V-Cycle algorithms defined in Chapter.(4) converge. See Theorem(4.24), (4.28), and (4.32).*

5.3 Full Multigrid for Finite Element Methods

Definition 5.29. *Let MG be as in Definition.(4.2). Define*

$$\hat{\vec{u}}_0 := A_0^{-1}\vec{f}_0.$$

For $\ell = 1$ to $\ell = L$

$$\vec{u}_\ell^{\,(0)} = P_{\ell-1}\hat{\vec{u}}_{\ell-1}$$

$$\vec{u}_\ell^{\,(\sigma+1)} = \mathrm{MG}\left(\vec{f}_\ell, \ell, \vec{u}_\ell^{\,(\sigma)}\right), \quad 0 \leq \sigma \leq r-1$$

$$\hat{\vec{u}}_\ell = \vec{u}_\ell^{\,(r)}$$

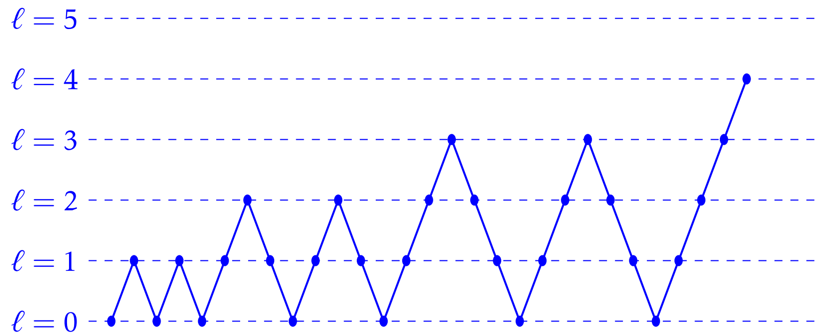


Figure 28: One full cycle with $r = 2$.

Hence

$$\vec{f}_ell := \begin{bmatrix} (f, \phi_{\ell,1})_{L^2} \\ (f, \phi_{\ell,2})_{L^2} \\ \vdots \\ \left(f, \phi_{\ell,n_\ell}\right)_{L^2} \end{bmatrix} \in \mathbb{R}^{n_\ell} (0 \le \ell \le L).$$

Theorem 5.30. *Suppose that, in general, for all* $\vec{u}_\ell^{\,(0)}$

$$\left\| \vec{u}_\ell^{\,E} - \mathrm{MG}\left(\vec{g}_\ell, \ell, \vec{u}_\ell^{\,(0)}\right) \right\|_{A_\ell} \le \gamma \left\| \vec{u}_\ell^{\,E} - \vec{u}_\ell^{\,(0)} \right\|_{A_\ell}, \tag{5.19}$$

where $0 < \gamma < 1$ *is independent of* ℓ *and*

$$\vec{u}_\ell^{\,E} := A_\ell^{-1} \vec{g}_\ell.$$

Suppose that r in the Full Multigrid algorithm in Definition. (5.29) is sufficiently large. Then there exists a constant $C > 0$ *such that*

$$\left| u_\ell^* - \hat{u}_\ell \right|_{H^1} \le C h_\ell |u|_{H^2}, \tag{5.20}$$

where

$$\vec{v}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}} \overset{\mathcal{B}_\ell}{\leftrightarrow} \vec{v}_{\ell-1} \in V_{\ell-1}. \quad (\textit{Full Multigrid})$$

$$a\left(u_\ell^*, v_\ell\right) = (f, v_\ell), \ \forall\, v_\ell \in V_\ell,$$

and

$$a(u, v) = (f, v), \ \forall\, v_\ell \in H_0^1(\Omega).$$

Proof. Define

$$\hat{e}_\ell = u_\ell^* - \hat{u}_\ell \in V_\ell.$$

This is the algebraic error not the FE error. Clearly $\hat{e}_0 \equiv 0$. Now

$$|\hat{e}_\ell|_{H^1}^2 = a(\hat{e}_\ell, \hat{e}_\ell) = \left\| \hat{\vec{e}}_\ell \right\|_{A_\ell}^2,$$

where

$$\hat{\vec{e}}_\ell = \vec{u}_\ell^{\,*} - \hat{\vec{u}}_\ell \in \mathbb{R}^{n_\ell} \overset{\mathcal{B}_\ell}{\leftrightarrow} \hat{e}_\ell = u_\ell^* - \hat{u}_\ell \in V_\ell.$$

Then

$$\begin{aligned}
|\hat{e}_\ell|_{H^1} &= \left\|\vec{u}^{\,*}_\ell - \hat{\vec{u}}_\ell\right\|_{A_\ell} \\
&\overset{(5.19)}{\leq} \gamma^r \left\|\vec{u}^{\,*}_\ell - P_{\ell-1}\hat{\vec{u}}_{\ell-1}\right\|_{A_\ell} \\
&= \gamma^r \left|u^*_\ell - \hat{u}_{\ell-1}\right|_{H^1} \\
&\leq \gamma^r \left\{\left|u^*_\ell - u\right|_{H^1} + \left|u - u^*_{\ell-1}\right|_{H^1} + \left|u^*_{\ell-1} - \hat{u}_{\ell-1}\right|_{H^1}\right\} \\
&\leq \gamma^r \left\{ch_\ell |u|_{H^2} + C\cdot 2h_\ell |u|_{H^2} + |\hat{e}_{\ell-1}|_{H^1}\right\} \\
&= \tilde{C}\gamma^r h_\ell |u|_{H^2} + \gamma^r |\hat{e}_{\ell-1}|_{H^1}.
\end{aligned}$$

By induction

$$\begin{aligned}
|\hat{e}_\ell|_{H^1} &\leq \left\{\tilde{C}h_\ell\gamma^r + \tilde{C}h_{\ell-1}\gamma^{2r} + \cdots + \tilde{C}h_0\gamma^{(\ell+1)r}\right\}|u|_{H^2} \\
&\leq h_\ell |u|_{H^2}\frac{\tilde{C}\gamma^r}{1-2\gamma^r},
\end{aligned}$$

provided

$$2\gamma^r < 1.$$

Setting

$$C := \frac{\tilde{C}\gamma^r}{1-2\gamma^r},$$

the theorem is proven. □

Remark 5.31. *The operation count for the full multigrid algorithm is $\mathcal{O}(n_L)$. In this sense, multigrid is optimal.*

5.4 Multigrid and Subspace Corrections

Definition 5.32. *Suppose $0 \leq j < \ell$. Define*

$$P_j^\ell := P_{\ell-1}P_{\ell-2}\cdots P_j \in \mathbb{R}^{n_\ell\times n_j}$$

In particular,

$$P_{\ell-1}^\ell = P_\ell \in \mathbb{R}^{n_\ell\times n_{\ell-1}}.$$

Lemma 5.33. *Suppose $v_j \in V_j$ for some $0 \leq j < \ell$, where $V_0 \subset V_1 \subset \cdots \subset V_\ell \subset \cdots$ are the usual nested FE spaces. Let $\vec{v}_j \in \mathbb{R}^{n_j}$ be the coordinate vector of v_j is basis $\mathcal{B}_j$. Then, the unique coordinate vector of v_j in the basis $\mathcal{B}_j$ is*

$$P_j^\ell \vec{v}_j \in \mathbb{R}^{n_\ell}.$$

Proof. Simple exercise. □

Definition 5.34. *Define $R_j^\ell \in \mathbb{R}^{n_j\times n_\ell}$, for $0 \leq j < \ell$ via*

$$R_j^\ell = \left(P_j^\ell\right)^T.$$

R_j^ℓ is called the multilevel restriction matrix and P_j^ℓ is called the multilevel prolongation matrix.

Lemma 5.35. *With the usual construction for the conforming FE method, we have, for any $0 \le j < \ell$,*

$$A_j = R_j^\ell A_\ell P_j^\ell \in \mathbb{R}^{n_j \times n_j}.$$

Proof. This follows because the Galerkin condition holds:

$$\begin{aligned} A_j &= R_j A_{j+1} P_j \\ &= R_j R_{j+1} A_{j+2} P_{j+1} P_j \\ &\vdots \\ &= R_j \cdots R_{\ell-1} A_\ell P_{\ell-1} \cdots P_j \\ &= R_j^\ell A_\ell P_j^\ell. \end{aligned}$$

□

Definition 5.36. *For any $0 \le j < \ell$ define the multilevel Ritz projection matrix via*

$$\Pi_j^\ell := A_j^{-1} R_j^\ell A_\ell \in \mathbb{R}^{n_j \times n_\ell}.$$

R_j^ℓ is called the multilevel restriction matrix and P_j^ℓ is called the multilevel prolongation matrix.

Lemma 5.37. *We have, for $0 \le j < \ell$,*

$$\Pi_j^\ell := \Pi_j \times \Pi_{j+1} \times \cdots \times \Pi_{\ell-1}.$$

Proof. The matrix product on the RHS is

$$\Pi_j \cdots \Pi_{\ell-1} = A_j^{-1} R_j A_{j+1} \times A_{j+1}^{-1} R_{j+1} A_{j+2} \times \cdots \times A_{\ell-1}^{-1} R_{\ell-1} A_\ell.$$

□

Definition 5.38. *Define, for any $0 \le j < \ell$*

$$\tilde{\Pi}_{\ell,j} := P_j^\ell \Pi_j^\ell \in \mathbb{R}^{n_\ell \times n_\ell}.$$

Observe that, for $j = \ell - 1$, we have

$$\tilde{\Pi}_{\ell,j} = \tilde{\Pi}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}.$$

Theorem 5.39. *Let $a(\cdot,\cdot)$, V_ℓ, et cetera , be defined as usual for the conforming FE method described in Chapter.5. Let $0 \le j < \ell$ and $\vec{u}_\ell \in V_\ell$ be arbitrary. Set*

$$u_j' = \underset{\in V_j \subset V_\ell}{R_j \vec{u}_\ell} \overset{\mathcal{B}_j}{\leftrightarrow} \vec{u}_j' \in \mathbb{R}^{n_j}.$$

Then, if $\vec{u}_\ell$ is the coordinate vector of $\vec{u}_\ell \in V_\ell$ in the basis $\mathcal{B}_\ell$, it follows that the unique representation of $R_j u_\ell \in V_j$ in the basis $\mathcal{B}_j$ is precisely

$$\vec{u}_j' = \Pi_j^\ell \vec{u}_\ell \in \mathbb{R}^{n_j}.$$

And the unique representation of $R_j \vec{u}_\ell \in V_\ell$ in the basis $\mathcal{B}_\ell$ is precisely

$$\tilde{\Pi}_{\ell,j} \vec{u}_\ell \in \mathbb{R}^{n_\ell}.$$

Proof. Let $\vec{u}_\ell \in V_\ell$ be given. $R_j \vec{u}_\ell$ is defined as the unique solution to

$$a\left(\mathcal{R}_j \vec{u}_\ell, v_j\right) = a\left(\vec{u}_\ell, v_j\right), \quad \forall\, v_j \in V_j.$$

Then

$$a\left(\mathcal{R}_j \vec{u}_\ell, v_j\right) = \vec{u}_j', \vec{v}_{j_{A_j}}.$$

On the other hand

$$a\left(\vec{u}_\ell, v_j\right) = \left(\vec{u}_\ell, P_j^\ell \vec{v}_j\right)_{A_\ell},$$

where

$$\vec{v}_j \in \mathbb{R}^{n_j} \overset{\mathcal{B}_\ell}{\leftrightarrow} v_j \in V_j,$$

as usual. Going further, we have

$$\begin{aligned} a\left(\vec{u}_\ell, v_j\right) &= \left(A_\ell \vec{u}_\ell, P_j^\ell \vec{v}_j\right)_\ell \\ &= \left(R_j^\ell A_\ell \vec{u}_\ell, \vec{v}_j\right)_j, \end{aligned}$$

and

$$a\left(\mathcal{R}_j \vec{u}_\ell, v_j\right) = A_j u_j', v_{j_j}.$$

Therefore,

$$A_j u_j' = R_j^\ell A_\ell \vec{u}_\ell,$$

or

$$u_j' = A_j^{-1} R_j^\ell A_\ell \vec{u}_\ell = \Pi_j^\ell \vec{u}_\ell.$$

The second part follows from lemma.(5.33). □

Definition 5.40. *Define, for any $0 \le j < \ell$*

$$T_j^\ell := \Pi_j^\ell - K_j^m \Pi_j^\ell \in \mathbb{R}^{n_j \times n_\ell},$$

where m is a given integer exponent. Define

$$\tilde{T}_{\ell,j} = P_j^\ell T_j^\ell \in \mathbb{R}^{n_\ell \times n_\ell},$$

$\tilde{T}_{\ell,j}$ is called a subspace "projection" matrix.

5.5 Properties of the "Projections"

Lemma 5.41. *Let $0 \le j < \ell$. Then*

$$\Pi_j^\ell P_j^\ell = I_j, \tag{5.21}$$

and

$$\tilde{\Pi}_{\ell,j}^2 = \tilde{\Pi}_{\ell,j}.$$

Proof. The Galerkin condition holds in the sense that

$$A_j = R_j^\ell A_\ell P_j^\ell. \tag{5.22}$$

By definition

$$\Pi_j^\ell = A_j^{-1} R_j^\ell A_\ell,$$

so that

$$\begin{aligned} \Pi_j^\ell P_j^\ell &= A_j^{-1} R_j^\ell A_\ell P_j^\ell \\ &= A_j^{-1} A_j = I_j. \end{aligned}$$

Now,

$$\tilde{\Pi}_{\ell,j}^2 = P_j^\ell \Pi_j^\ell P_j^\ell \Pi_j^\ell = P_j^\ell \Pi_j^\ell = \tilde{\Pi}_{\ell,j}.$$

□

Definition 5.42. *Let $0 \le j < \ell$. Define*

$$T_j^{'\ell} := \Pi_j^\ell - (K_j^*)^m \Pi_j^\ell, \tilde{T}_{\ell,j}^{'} := P_j^\ell T_j^{'\ell},$$

where

$$K_j^* = I_j - S_j^T A_j.$$

Lemma 5.43. *Let $0 \le j < \ell$. Then*

$$\tilde{T}_{\ell,j}^* = \tilde{T}_{\ell,j}. \tag{5.23}$$

And

$$\tilde{T}_{\ell,j}^* = \tilde{T}_{\ell,j}^{'}.$$

Proof. Recall

$$\left(\tilde{\Pi}_{\ell,j} \vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell} = \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}^* \vec{v}_\ell\right)_{A_\ell},$$

for all $\vec{u}_\ell, \vec{v}_\ell \in \mathbb{R}^{n_\ell}$. Then

$$\begin{aligned} \left(\tilde{\Pi}_{\ell,j} \vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell} &= \left(P_j^\ell \Pi_j^\ell \vec{u}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\ &= \left(\Pi_j^\ell \vec{u}_\ell, R_j^\ell A_\ell \vec{v}_\ell\right)_j \\ &= \left(A_j^{-1} R_j^\ell A_\ell \vec{u}_\ell, R_j^\ell A_\ell \vec{v}_\ell\right)_j \\ &= \left(R_j^\ell A_\ell \vec{u}_\ell, A_j^{-1} R_j^\ell A_\ell \vec{v}_\ell\right)_j \end{aligned}$$

$$
\begin{aligned}
&= \left(A_\ell \vec{u}_\ell, \tilde{\Pi}_{\ell,j} \vec{v}_\ell\right)_\ell \\
&= \left(\vec{u}_\ell, \tilde{\Pi}^*_{\ell,j} \vec{v}_\ell\right)_{A_\ell}.
\end{aligned}
$$

Now

$$\tilde{T}_{\ell,j} = \tilde{\Pi}_{\ell,j} - P_j^\ell K_j^m \Pi_j^\ell.$$

Therefore,

$$
\begin{aligned}
\left(\tilde{T}_{\ell,j}\vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell} &= \left(\tilde{\Pi}_{\ell,j}\vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell} - \left(P_j^\ell K_j^m \Pi_j^\ell \vec{u}_\ell, \vec{v}_\ell\right)_{A_\ell} \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(P_j^\ell K_j^m \Pi_j^\ell \vec{u}_\ell, A_\ell \vec{v}_\ell\right)_\ell \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(K_j^m \Pi_j^\ell \vec{u}_\ell, R_j^\ell A_\ell \vec{v}_\ell\right)_j \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(K_j^m \Pi_j^\ell \vec{u}_\ell, A_j A_j^{-1} R_j^\ell A_\ell \vec{v}_\ell\right)_j \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(K_j^m \Pi_j^\ell \vec{u}_\ell, A_j^{-1} R_j^\ell A_\ell \vec{v}_\ell\right)_{A_j} \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(\Pi_j^\ell \vec{u}_\ell, \left(K_j^m\right)^* \Pi_j^\ell \vec{v}_\ell\right)_{A_j} \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(\Pi_j^\ell \vec{u}_\ell, \left(K_j^*\right)^m \Pi_j^\ell \vec{v}_\ell\right)_{A_j} \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(\Pi_j^\ell \vec{u}_\ell, \left(K_j^*\right)^m \Pi_j^\ell \vec{v}_\ell\right)_{A_j} \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(A_j^{-1} R_j^\ell A_\ell \vec{u}_\ell, \left(K_j^*\right)^m \Pi_j^\ell \vec{v}_\ell\right)_{A_j} \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(R_j^\ell A_\ell \vec{u}_\ell, \left(K_j^*\right)^m \Pi_j^\ell \vec{v}_\ell\right)_j \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(A_\ell \vec{u}_\ell, P_j^\ell \left(K_j^*\right)^m \Pi_j^\ell \vec{v}_\ell\right)_\ell \\
&= \left(\vec{u}_\ell, \tilde{\Pi}_{\ell,j}\vec{v}_\ell\right)_{A_\ell} - \left(\vec{u}_\ell, P_j^\ell \left(K_j^*\right)^m \Pi_j^\ell \vec{v}_\ell\right)_{A_\ell} \\
&= \left(\vec{u}_\ell, \tilde{T}'_{\ell,j}\vec{v}_\ell\right)_{A_\ell}.
\end{aligned}
$$

So,

$$\tilde{\Pi}^*_{\ell,j} = \tilde{\Pi}_{\ell,j},$$

and

$$\tilde{T}^*_{\ell,j} = \tilde{T}'_{\ell,j}.$$

□

Remark 5.44. *We note that, in general*

$$\tilde{T}^2_{\ell,j} \neq \tilde{T}_{\ell,j}.$$

$\tilde{T}_{\ell,j}$ is not a projection matrix, in this sense. We call $\tilde{T}_{\ell,j}$ (and $\tilde{T}'_{\ell,j}$) projection-like operators.

Theorem 5.45. *Let $0 \le j < \ell$. Then*

$$\left(I_\ell - \tilde{\Pi}_{\ell,j}\right)\left(I_\ell - \tilde{T}_{\ell,j}\right) = I_\ell - \tilde{\Pi}_{\ell,j}, \tag{5.24}$$

and

$$\left(I_\ell - \tilde{T}^*_{\ell,j}\right)\left(I_\ell - \tilde{\Pi}_{\ell,j}\right) = I_\ell - \tilde{\Pi}_{\ell,j}. \tag{5.25}$$

Proof. The LHS of (5.24) is

$$M_\ell := I_\ell - \tilde{\Pi}_{\ell,j} - \tilde{T}_{\ell,j} + \tilde{\Pi}_{\ell,j}\tilde{T}_{\ell,j}.$$

By definition,

$$\begin{aligned}\tilde{\Pi}_{\ell,j}\tilde{T}_{\ell,j} &= P_j^\ell \Pi_j^\ell P_j^\ell T_j^\ell \\ &\overset{5.21}{=} P_j^\ell I_j T_j^\ell = \tilde{T}_{\ell,j}.\end{aligned}$$

So

$$M_\ell := I_\ell - \tilde{\Pi}_{\ell,j} - \tilde{T}_{\ell,j} + \tilde{T}_{\ell,j} = I_\ell - \tilde{\Pi}_{\ell,j}.$$

For the second identity, we denote the LHS of (5.25) as

$$M_\ell' = I_\ell - \tilde{T}_{\ell,j}^* - \tilde{\Pi}_{\ell,j} + \tilde{T}_{\ell,j}^*\tilde{\Pi}_{\ell,j}.$$

Then

$$\begin{aligned}\tilde{T}_{\ell,j}^*\tilde{\Pi}_{\ell,j} &= P_j^\ell T_j^{'\ell} P_j^\ell \Pi_j^\ell \\ &= P_j^\ell \left(\Pi_j^\ell - (K_j^*)^m \Pi_j^\ell\right) P_j^\ell \Pi_j^\ell \\ &\overset{5.21}{=} P_j^\ell \left(I_j^\ell - (K_j^*)^m\right)\Pi_j^\ell = \tilde{T}_{\ell,j}^*.\end{aligned}$$

So

$$M_\ell' = I_\ell - \tilde{\Pi}_{\ell,j}^*,$$

as desired. □

Lemma 5.46. *Let $0 \le j < \ell$. Then*

$$I_\ell - \tilde{\Pi}_{\ell,j} = \left(I_\ell - \tilde{T}_{\ell,j}^*\right)\left(I_\ell - \tilde{\Pi}_{\ell,j}\right)\left(I_\ell - \tilde{T}_{\ell,j}\right). \tag{5.26}$$

Proof. Since the Galerkin condition holds, $\tilde{\Pi}_{\ell,j}$ is a bona fide projection matrix (Lemma (5.41))

$$\tilde{\Pi}_{\ell,j}^2 = \tilde{\Pi}_{\ell,j},$$

and

$$\left(I_\ell - \tilde{\Pi}_{\ell,j}\right)^2 = I_\ell - \tilde{\Pi}_{\ell,j},$$

is a direct consequence. By the last result

$$\begin{aligned}I_\ell - \tilde{\Pi}_{\ell,j} &= \left(I_\ell - \tilde{\Pi}_{\ell,j}\right)\left(I_\ell - \tilde{\Pi}_{\ell,j}\right) \\ &\overset{(5.24)(5.25)}{=} \left(I_\ell - \tilde{T}_{\ell,j}^*\right)\left(I_\ell - \tilde{\Pi}_{\ell,j}\right)\left(I_\ell - \tilde{\Pi}_{\ell,j}\right)\left(I_\ell - \tilde{T}_{\ell,j}\right) \\ &= \left(I_\ell - \tilde{T}_{\ell,j}^*\right)\left(I_\ell - \tilde{\Pi}_{\ell,j}\right)\left(I_\ell - \tilde{T}_{\ell,j}\right).\end{aligned}$$

□

Lemma 5.47. *Let $0 \le i < j < \ell$. Then*

$$P_j^\ell\left(I_j - \tilde{T}_{j,i}\right) = \left(I_\ell - \tilde{T}_{\ell,i}\right)P_j^\ell. \tag{5.27}$$

Proof.

$$\begin{aligned}
P_j^\ell\left(I_j - \tilde{T}_{j,i}\right) &\overset{5.21}{=} P_j^\ell\left(I_j - \tilde{T}_{j,i}\right)\Pi_j^\ell P_j^\ell \\
&= \left\{P_j^\ell \Pi_j^\ell - P_j^\ell P_i^j T_i^j \Pi_j^\ell\right\} P_j^\ell \\
&= \left\{\tilde{\Pi}_{\ell,j} - P_i^\ell\left(\Pi_i^j - K_i^m \Pi_i^j\right)\Pi_j^\ell\right\} P_j^\ell \\
&= \left\{\tilde{\Pi}_{\ell,j} - P_i^\ell \Pi_i^j + P_i^\ell K_i^m \Pi_i^\ell\right\} P_j^\ell \\
&= \left\{\tilde{\Pi}_{\ell,j} - \tilde{T}_{\ell,i}\right\} P_j^\ell \\
&= P_j^\ell \Pi_j^\ell P_j^\ell - \tilde{T}_{\ell,i} P_j^\ell \\
&\overset{5.21}{=} \left(I_\ell - \tilde{T}_{\ell,i}\right) P_j^\ell .
\end{aligned}$$

□

Corollary 5.48. *Let $0 \le i < j < \ell$. Then*

$$P_j^\ell\left(I_j - \tilde{\Pi}_{j,i}\right) = \left(I_\ell - \tilde{\Pi}_{\ell,i}\right) P_j^\ell . \tag{5.28}$$

Theorem 5.49. *Let $V_\ell, \mathcal{T}_\ell$ and $a(\cdot,\cdot)$ be definded as usual. Consider the symmetric V-cycle algorithm: $m = m_1 = m_2$ and $p = 1$. The error propagation matrix can be expressed as*

$$\begin{aligned}
E_\ell &= (K_\ell^*)^m\left(I_\ell - \tilde{T}_{\ell,\ell-1}^*\right)\times\cdots\times\left(I_\ell - \tilde{T}_{\ell,1}^*\right)\left(I_\ell - \tilde{\Pi}_{\ell,0}^*\right) \\
&\quad \times\left(I_\ell - \tilde{T}_{\ell,1}\right)\times\cdots\times\left(I_\ell - \tilde{T}_{\ell,\ell-1}\right)(K_\ell)^m,
\end{aligned} \tag{5.29}$$

for all $\ell \ge 1$.

Proof. Define the quantity

$$M_{\ell,j} := I_\ell - \tilde{\Pi}_{\ell,j} + P_j^\ell E_j \Pi_j^\ell ,$$

for any $0 \le j < \ell$. Observe that when $j = 0$

$$M_{\ell,0} = I_\ell - \tilde{\Pi}_{\ell,0},$$

since $E_0 = 0$. Now by Theorem.(4.8),

$$M_{\ell,j} = I_\ell - \tilde{\Pi}_{\ell,j} + P_j^\ell\left(K_j^*\right)^m\left(I_j - \tilde{\Pi}_{j,j-1} + P_{j-1}^j E_{j-1}\Pi_{j-1}^j\right)K_j^m \Pi_{j-1}^j . \tag{5.30}$$

In other words,

$$M_{\ell,j} = I_\ell - \tilde{\Pi}_{\ell,j} + P_j^\ell M_{j,j-1} K_j^m \Pi_{j-1}^j .$$

Now, observe that

$$\begin{aligned}
P_j^\ell\left(K_j^*\right)^m &\overset{5.21}{=} P_j^\ell\left(K_j^*\right)^m \Pi_j^\ell P_j^\ell \\
&= \left(P_j^\ell \Pi_j^\ell - P_j^\ell \Pi_j^\ell + P_j^\ell\left(K_j^*\right)^m \Pi_j^\ell\right) P_j^\ell \\
&= \left(P_j^\ell \Pi_j^\ell - \tilde{T}_{\ell,j}^*\right) P_j^\ell \qquad (5.31) \\
&= P_j^\ell \Pi_j^\ell P_j^\ell - \tilde{T}_{\ell,j}^* P_j^\ell
\end{aligned}$$

$$\stackrel{5.21}{=} \left(I_\ell - \tilde{T}^*_{\ell,j}\right) P^\ell_j.$$

Similarly,

$$\begin{aligned}
K^m_j \Pi^\ell_j &= \Pi^\ell_j - \Pi^\ell_j + K^m_j \Pi^\ell_j \\
&= \Pi^\ell_j - T^\ell_j \\
&\stackrel{5.21}{=} \Pi^\ell_j P^\ell_j \left(\Pi^\ell_j - T_{\ell,j}\right) \qquad (5.32) \\
&= \Pi^\ell_j \left(P^\ell_j \Pi^\ell_j - \tilde{T}_{\ell,j}\right) \\
&= \Pi^\ell_j P^\ell_j \Pi^\ell_j - \Pi^\ell_j \tilde{T}_{\ell,j} \\
&\stackrel{5.21}{=} \Pi^\ell_j \left(I_\ell - \tilde{T}_{\ell,j}\right).
\end{aligned}$$

Putting (5.30)- (5.32) together, we have

$$\begin{aligned}
M_{\ell,j} &= I_\ell - \tilde{\Pi}_{\ell,j} + \left(I_\ell - \tilde{T}^*_{\ell,j}\right) P^\ell_j \left\{ I_j - \tilde{\Pi}_{j,j-1} + P^j_{j-1} E_{j-1} \Pi^j_{j-1} \right\} \Pi^\ell_j \left(I_\ell - \tilde{T}_{\ell,j}\right) \\
&\stackrel{5.26}{=} \left(I_\ell - \tilde{T}^*_{\ell,j}\right) \left\{ I_\ell - \tilde{\Pi}_{\ell,j} + P^\ell_j \left(I_j - \tilde{\Pi}_{j,j-1} + P^j_{j-1} E_{j-1} \Pi^j_{j-1} \right) \Pi^\ell_j \right\} \left(I_\ell - \tilde{T}_{\ell,j}\right) \\
&= \left(I_\ell - \tilde{T}^*_{\ell,j}\right) \left\{ I_\ell - \tilde{\Pi}_{\ell,j} + \tilde{\Pi}_{\ell,j} - \tilde{\Pi}_{\ell,j-1} + P^\ell_{j-1} E_{j-1} \Pi^\ell_{j-1} \right\} \left(I_\ell - \tilde{T}_{\ell,j}\right).
\end{aligned}$$

Or

$$M_{\ell,j} = \left(I_\ell - \tilde{T}^*_{\ell,j}\right) M_{\ell,j-1} \left(I_\ell - \tilde{T}_{\ell,j}\right).$$

Therefore

$$\begin{aligned}
M_{\ell,\ell-1} &= \left(I_\ell - \tilde{T}^*_{\ell,\ell-1}\right) M_{\ell,\ell-2} \left(I_\ell - \tilde{T}_{\ell,\ell-1}\right) \\
&= \left(I_\ell - \tilde{T}^*_{\ell,\ell-1}\right)\left(I_\ell - \tilde{T}^*_{\ell,\ell-2}\right) M_{\ell,\ell-3} \left(I_\ell - \tilde{T}_{\ell,\ell-2}\right)\left(I_\ell - \tilde{T}_{\ell,\ell-1}\right) \\
&= \left(I_\ell - \tilde{T}^*_{\ell,\ell-1}\right) \times \cdots \times \left(I_\ell - \tilde{T}^*_{\ell,1}\right)\left(I_\ell - \tilde{\Pi}_{\ell,0}\right) \times \left(I_\ell - \tilde{T}_{\ell,1}\right)\left(I_\ell - \tilde{T}_{\ell,2}\right) \times \cdots \times \left(I_\ell - \tilde{T}_{\ell,\ell-1}\right)
\end{aligned}$$

But recall that

$$E_\ell = \left(K^*_\ell\right)^m M_{\ell,\ell-1} K^m_\ell,$$

from Theorem.(4.8). The theorem is proven. □

Corollary 5.50. *For the one-sided V-cycle with only pre-smoothing ($p = 1, m := m_1 > 0$ and $m_2 = 0$), we have*

$$E^{pre}_\ell = \left(I_\ell - \tilde{\Pi}_{\ell,0}\right)\left(I_\ell - \tilde{T}_{\ell,1}\right)\left(I_\ell - \tilde{T}_{\ell,2}\right) \times \cdots \times \left(I_\ell - \tilde{T}_{\ell,\ell-1}\right) K^m_\ell.$$

And for the algorithm with only post-smoothing ($p = 1, m := m_2 > 0$ and $m_1 = 0$), we have

$$E^{post}_\ell = \left(K^*_\ell\right)^m \left(I_\ell - \tilde{T}^*_{\ell,\ell-1}\right) \times \cdots \times \left(I_\ell - \tilde{T}^*_{\ell,1}\right)\left(I_\ell - \tilde{\Pi}_{\ell,0}\right).$$

Therefore

$$E_\ell = E^{post}_\ell \times E^{pre}_\ell.$$

Moreover

$$\left(E^{post}_\ell\right)^* = E^{pre}_\ell.$$

Clearly

$$E^*_\ell = E_\ell = \left(E^{pre}_\ell\right)^* E^{pre}_\ell$$

is SPSD.

Theorem 5.51. *Both of the one-sided method converge for any $m > 0$ if Richardson's method is used for smoothing.*

Proof. We have shown, in Theorem.(4.31) that there is some C_0 such that

$$\left\|E_\ell \vec{u}_\ell\right\|_{A_\ell} \leq \frac{C_0}{m + C_0}\left\|\vec{u}_\ell\right\|_{A_\ell},$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$. We wish the prove that

$$\left\|E_\ell^{\text{pre}} \vec{u}_\ell\right\|_{A_\ell} \leq \gamma,$$

for some $0 \leq \gamma < 1$. Observe that

$$\begin{aligned}
\left\|E_\ell^{\text{pre}} \vec{u}_\ell\right\|_{A_\ell}^2 &= \left(E_\ell^{\text{pre}} \vec{u}_\ell, E_\ell^{\text{pre}} \vec{u}_\ell\right)_{A_\ell} \\
&= \left(\vec{u}_\ell, \left(E_\ell^{\text{pre}}\right)^* E_\ell^{\text{pre}} \vec{u}_\ell\right)_{A_\ell} \\
&= \left(\vec{u}_\ell, E_\ell \vec{u}_\ell\right)_{A_\ell} \\
&\overset{C.S.}{\leq} \left\|\vec{u}_\ell\right\|_{A_\ell}\left\|E_\ell \vec{u}_\ell\right\|_{A_\ell} \\
&\leq \left\|\vec{u}_\ell\right\|_{A_\ell} \frac{C_0}{m + C_0}\left\|\vec{u}_\ell\right\|_{A_\ell} \\
&\leq \frac{C_0}{m + C_0}\left\|\vec{u}_\ell\right\|_{A_\ell}^2.
\end{aligned}$$

Thus

$$\left\|E_\ell^{\text{pre}} \vec{u}_\ell\right\|_{A_\ell} \leq \sqrt{\frac{C_0}{m + C_0}}\left\|\vec{u}_\ell\right\|_{A_\ell}.$$

□

Theorem 5.52. *For the W-cycle algorithm with only pre-smoothing :$m := m_1 > 0, m_2 = 0, p = 2$. The error propagation matrix is*

$$E_\ell^{w,pre} = F_\ell E_\ell^{pre},$$

where E_ℓ^{pre} is defined above, and $F_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ is a matrix with

$$\|F_\ell\|_{A_\ell} \leq 1.$$

consequently the one-sided W-cycle method with pre-smoothing converges for any $m > 0$.

Proof. Exercise. □

Theorem 5.53. *For the 2-sided W-cycle algorithm with $p = 2, m := m_1 = m_2$, the error propagation matrix is*

$$E_\ell^W = \left(E_\ell^{pre}\right)^* D_\ell E_\ell^{pre}$$

where

$$E_\ell^{pre} = \left(I_\ell - \tilde{\Pi}_{\ell,0}\right)\left(I_\ell - \tilde{T}_{\ell,1}\right)\left(I_\ell - \tilde{T}_{\ell,2}\right) \times \cdots \times \left(I_\ell - \tilde{T}_{\ell,\ell-1}\right) K_\ell^m.$$

and

$$\|D_\ell\|_{A_\ell} \le 1, \;\; \forall\, \ell \ge 1.$$

The algorithm converges if the symmetric V-cycle algorithm converges with the uniform contraction $0 < \gamma < 1$, i.e.

$$\left\|E_\ell \vec{u}_\ell\right\|_{A_\ell} \le \gamma \left\|\vec{u}_\ell\right\|_{A_\ell},$$

for all $\vec{u}_\ell \in \mathbb{R}^{n_\ell}$, with $E_\ell = \left(E_\ell^{pre}\right)^* E_\ell^{pre}$. Here γ may (and usually does) depend upon m.

Proof. Exercise. □

6 Subspace Decompositions

Definition 6.1.

$$0 < m_0 < m_1 < \cdots < m_\ell < \cdots < m_L \le n_L.$$

Let the matrices

$$Q_j^L \in \mathbb{R}^{n_L \times m_j}$$

be defined for $0 \le j \le L$. We say that Assumption(S1) holds, or , equivalently, that the Q_j^L prolongation matrices gives us a subspace decomposition of $\mathbb{R}^{n_L}$, iff for every $\vec{u}_L \in \mathbb{R}^{n_L}$, there exist vector

$$w_j \in \mathbb{R}^{m_j}, \ \ 0 \le j \le L,$$

such that

$$\vec{u}_L = \sum_{j=0}^{L} Q_j^L \vec{w}_j \tag{6.1}$$

Herein

$$(\vec{u}_L, \vec{u}_L) := \sum_{j=1}^{n_L} (\vec{u}_L)_j (\vec{v}_L)_j \quad \forall \vec{u}_L, \ \vec{v}_L \in \mathbb{R}^{n_L}.$$

And

$$(\vec{u}_\ell, \vec{u}_\ell)_\ell := \sum_{j=1}^{m_\ell} (\vec{u}_\ell)_j (\vec{v}_\ell)_j \quad \forall \vec{u}_\ell, \ \vec{v}_\ell \in \mathbb{R}^{m_\ell}.$$

Definition 6.2. *The matrix $C \in \mathbb{R}^{n_L \times n_L}$ is called an additive subspace preconditioner iff there are SPD matrices $C_\ell \in \mathbb{R}^{m_\ell \times m_\ell}$, for each $0 \le \ell \le L$, such that*

$$C = \sum_{\ell=0}^{L} Q_\ell^L C_\ell^{-1} Z_\ell^L, \tag{6.2}$$

where

$$Z_\ell^L = \left(Q_\ell^L\right)^T \in \mathbb{R}^{m_\ell \times n_L},$$

i.e.

$$\left(\vec{u}_\ell, Q_\ell^L \vec{v}_L\right)_\ell = \left(\left(Q_\ell^L\right)^T \vec{u}_\ell, \vec{v}_L\right) = \left(Z_\ell^L \vec{u}_\ell, \vec{v}_L\right)$$

for all $\vec{u}_\ell \in \mathbb{R}^{m_\ell}$ and $\vec{v}_L \in \mathbb{R}^{n_L}$.

Lemma 6.3. *Suppose that assumption (S1) holds. Then C is SPD with respect to $(\cdot,\cdot)$, and, consequently if A is also SPD w.r.t $(\cdot,\cdot)$, then CA is SPD w.r.t. $(\cdot,\cdot)_{C^{-1}}$ and $(\cdot,\cdot)_A$.*

Proof. C is clearly symmetric, since each C_ℓ^{-1} is symmetric. Now, let $\vec{u}_L \in \mathbb{R}^{n_L}$ be arbitrary. Then

$$\begin{aligned} (\vec{u}_L, C\vec{u}_L) &\overset{(6.2)}{=} \left(\vec{u}_L, \sum_{\ell=0}^{L} Q_\ell^L C_\ell^{-1} Z_\ell^L \vec{u}_L\right) \\ &= \sum_{\ell=0}^{L} \left(\vec{u}_L, Q_\ell^L C_\ell^{-1} Z_\ell^L \vec{u}_L\right) \end{aligned}$$

$$= \sum_{\ell=0}^{L}\left(Z_\ell^L\vec{u}_L, C_\ell^{-1}Z_\ell^L\vec{u}_L\right)_\ell \geq 0,$$

since C_ℓ^{-1} is SPD w.r.t. $(\cdot,\cdot)_\ell$, $0 \leq \ell \leq L$. Suppose

$$\sum_{\ell=0}^{L}\left(Z_\ell^L\vec{u}_L, C_\ell^{-1}Z_\ell^L\vec{u}_L\right)_\ell = 0.$$

Since, again, C_ℓ^{-1} is SPD, it must be that

$$Z_\ell^L\vec{u}_L = \vec{0}, \quad 0 \leq \ell \leq L. \tag{6.3}$$

In this case, since (S1) holds, we have

$$\begin{aligned}
\left\|\vec{u}_L\right\|^2 &= (\vec{u}_L, \vec{u}_L) \\
&\overset{S1}{=} \left(\vec{u}_L, \sum_{\ell=1}^{L} Q_\ell^L\vec{w}_\ell\right) \\
&= \sum_{\ell=1}^{L}\left(\vec{u}_L, Q_\ell^L\vec{w}_\ell\right) \\
&= \sum_{\ell=1}^{L}\left(Z_\ell^L\vec{u}_L, \vec{w}_\ell\right)_\ell \\
&\overset{(6.3)}{=} \sum_{\ell=1}^{L}\left(\vec{0}, \vec{w}_\ell\right)_\ell = 0.
\end{aligned}$$

Thus

$$\left\|\vec{u}_L\right\| = 0.$$

Hence $\vec{u}_L = \vec{0}$, which shows that C is SPD w.r.t. $(\cdot,\cdot)$. The results concerning CA are easy to show. □

Theorem 6.4. *Suppose that (S1) holds. Then, for any $\vec{u}_L \in \mathbb{R}^{n_L}$,*

$$\begin{aligned}
(\vec{u}_L, \vec{u}_L)_{C^{-1}} :&= \left(\vec{u}_L, C^{-1}\vec{u}_L\right) \\
&= \min_{\vec{u}=\sum_{\ell=0}^{L} Q_\ell^L\vec{w}_\ell} = \sum_{\ell=0}^{L}(C_\ell\vec{w}_\ell, \vec{w}_\ell)_\ell.
\end{aligned} \tag{6.4}$$

Proof. Since each $C_\ell \in \mathbb{R}^{m_\ell\times m_\ell}$ is SPD w.r.t. $(\cdot,\cdot)_\ell$, $(\cdot,\cdot)_{C_\ell^{-1}}$ is a bona fide inner product. Therefore

$$\begin{aligned}
\left(\vec{u}_\ell, C^{-1}\vec{v}_\ell\right) &=: (\vec{u}_\ell, \vec{v}_\ell)_{C_\ell^{-1}} \\
&\overset{C.S.}{\leq} \left\|\vec{u}_\ell\right\|_{C_\ell^{-1}}\left\|\vec{v}_\ell\right\|_{C_\ell^{-1}} \\
&= \sqrt{(\vec{u}_\ell, \vec{u}_\ell)_{C_\ell^{-1}}}\sqrt{(\vec{v}_\ell, \vec{v}_\ell)_{C_\ell^{-1}}}.
\end{aligned}$$

Let $\vec{u}_L \in \mathbb{R}^{n_L}$ be arbitrary. Then

$$\vec{u}_L \overset{(S1)}{=} \sum_{\ell=0}^{L} Q_\ell^L\vec{w}_\ell, \;\; \exists\vec{w}_\ell \in \mathbb{R}^{m_\ell}.$$

We have

$$
\begin{aligned}
(\vec{u}_L, \vec{u}_L)_{C^{-1}} &= \sum_{\ell=0}^{L} \left(\vec{u}_L, Q_\ell^L \vec{w}_\ell\right)_{C^{-1}} \\
&= \sum_{\ell=0}^{L} \left(Z_\ell^L C^{-1} \vec{u}_L, \vec{w}_\ell\right)_\ell \\
&= \sum_{\ell=0}^{L} \left(Z_\ell^L C^{-1} \vec{u}_L, C_\ell \vec{w}_\ell\right)_{C_\ell^{-1}} \\
&\overset{C.S.}{\leq} \sum_{\ell=0}^{L} \left\|Z_\ell^L C^{-1} \vec{u}_L\right\|_{C_\ell^{-1}} \left\|C_\ell \vec{w}_\ell\right\|_{C_\ell^{-1}} \\
&\overset{C.S.}{\leq} \left(\sum_{\ell=0}^{L} \left\|Z_\ell^L C^{-1} \vec{u}_L\right\|_{C_\ell^{-1}}^2\right)^{1/2} \left(\sum_{\ell=0}^{L} \left\|C_\ell \vec{w}_\ell\right\|_{C_\ell^{-1}}^2\right)^{1/2} \\
&\overset{C.S.}{\leq} \left(\sum_{\ell=0}^{L} \left(Z_\ell^L C^{-1} \vec{u}_L, Z_\ell^L C^{-1} \vec{u}_L\right)_{C_\ell^{-1}}\right)^{1/2} \left(\sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, C_\ell \vec{w}_\ell)_{C_\ell^{-1}}\right)^{1/2} \\
&\leq \left(\sum_{\ell=0}^{L} \left(Z_\ell^L C^{-1} \vec{u}_L, C_\ell^{-1} Z_\ell^L C^{-1} \vec{u}_L\right)_\ell\right)^{1/2} \left(\sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell\right)^{1/2} \\
&\overset{(6.2)}{=} \left(\left(C^{-1} \vec{u}_L, C C^{-1} \vec{u}_L\right)\right)^{1/2} \left(\sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell\right)^{1/2} \\
&= \left\|\vec{u}_L\right\|_{C^{-1}} \left(\sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell\right)^{1/2}.
\end{aligned}
$$

So,

$$
(\vec{u}_L, \vec{u}_L)_{C^{-1}} \leq \sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell .
$$

For the particular choice

$$
\vec{w}_\ell = C_\ell^{-1} Z_\ell^L C^{-1} \vec{u}_L \in \mathbb{R}^{m_\ell}, \quad 0 \leq \ell \leq L,
$$

we have

$$
\vec{u}_L = \sum_{\ell=0}^{L} Q_\ell^L \vec{w}_\ell,
$$

and

$$
\begin{aligned}
\sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell &= \sum_{\ell=0}^{L} \left(C_\ell C_\ell^{-1} Z_\ell^L C^{-1} \vec{u}_L, C_\ell^{-1} Z_\ell^L C^{-1} \vec{u}_L\right)_\ell \\
&= \sum_{\ell=0}^{L} \left(C^{-1} \vec{u}_L, Q_\ell^L C_\ell^{-1} Z_\ell^L C^{-1} \vec{u}_L\right) \\
&\overset{(6.2)}{=} \left(C^{-1} \vec{u}_L, C C^{-1} \vec{u}_L\right) \\
&= (\vec{u}_L, \vec{u}_L)_{C^{-1}} .
\end{aligned}
$$

□

Theorem 6.5. *The eigenvalues of CA are positive, provided A is SPD w.r.t. $(\cdot,\cdot)$ and assumption (S1) holds. Moreover*

$$\lambda_{\max}(CA) = \max_{\vec{u}_L \in \mathbb{R}_*^{n_L}} \frac{(A\vec{u}_L, \vec{u}_L)}{\displaystyle\min_{\vec{u}_L = \sum_{\ell=0}^{L} Q_\ell^L \vec{w}_\ell} \sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell}, \tag{6.5}$$

$$\lambda_{\min}(CA) = \min_{\vec{u}_L \in \mathbb{R}_*^{n_L}} \frac{(A\vec{u}_L, \vec{u}_L)}{\displaystyle\min_{\vec{u}_L = \sum_{\ell=0}^{L} Q_\ell^L \vec{w}_\ell} \sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell}. \tag{6.6}$$

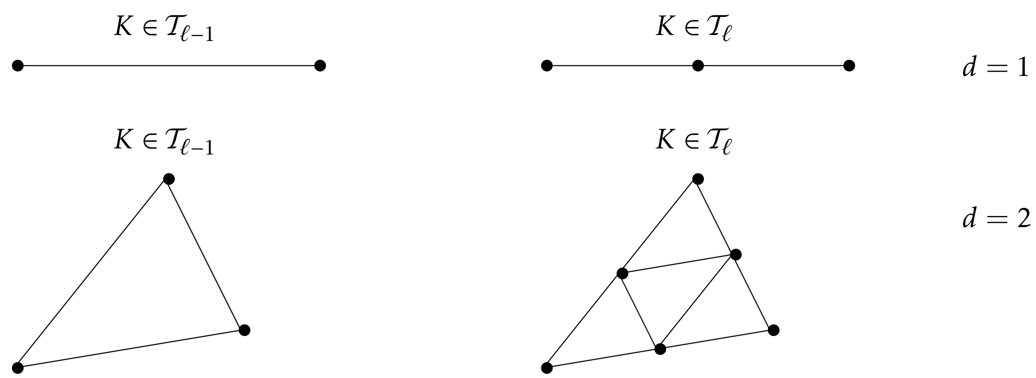
Proof. Recall that CA is SPD w.r.t. $(\cdot,\cdot)_{C^{-1}}$. Thus the eigenvalues are positive real and the corresponding eigenvectors may be chosen so that they form an orthonormal basis for $\mathbb{R}^{n_L}$ w.r.t. $(\cdot,\cdot)_{C^{-1}}$. Moreover the Rayleigh quotient formula holds

$$\begin{aligned}\lambda_{\max}(CA) &= \max_{\vec{u}_L \in \mathbb{R}_*^{n_L}} \frac{(CA\vec{u}_L, \vec{u}_L)_{C^{-1}}}{(\vec{u}_L, \vec{u}_L)_{C^{-1}}} \\ &\overset{(6.4)}{=} \max_{\vec{u}_L \in \mathbb{R}_*^{n_L}} \frac{(A\vec{u}_L, \vec{u}_L)}{\displaystyle\min_{\vec{u}_L = \sum_{\ell=0}^{L} Q_\ell^L \vec{w}_\ell} \sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell}.\end{aligned}$$

The formula for $\lambda_{\min}(CA)$ is proven in a similar way. □

6.1 Hierarchical Basis

Definition 6.6. *Suppose $d = 1$ or $d = 2$. Suppose Ω is an open interval ($d = 1$) or an open convex polygonal domain ($d = 2$). Suppose $\mathcal{T}_0$is an initial conforming triangulation ($d = 2$) or partition ($d = 1$) of Ω into triangles ($d = 1$) or subintervals ($d = 1$). Let $\mathcal{T}_\ell$ be the family of triangulations obtained by subdividing each triangular of $\mathcal{T}_{\ell-1}$ into 4 similar triangles by joining the edge midpoints ($d = 2$) or subdividing each interval of $\mathcal{T}_{\ell-1}$ into 2 equal subintervals.*



Set

$$V_\ell := \left\{ v \in C^0(\overline{\Omega}) | v|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_\ell, v|_{\partial\Omega} \equiv 0 \right\} \tag{6.7}$$

for all $0 \le \ell \le L$. Define

$$n_\ell := \dim(V_\ell).$$

Set $W_0 := V_0$, and , for $1 \le \ell \le L$, define

$$W_\ell := \left\{ v \in V_\ell | v(\vec{N}_{\ell-1,j}) = 0, \ \forall \ 1 \le j \le n_{\ell-1} \right\}. \tag{6.8}$$

Recall that $\left\{ \vec{N}_{\ell,j} \right\}_{j=1}^{n_\ell} \subset \Omega$ is set of interior vertices of $\mathcal{T}_\ell$. Set

$$m_\ell := \dim(W_\ell).$$

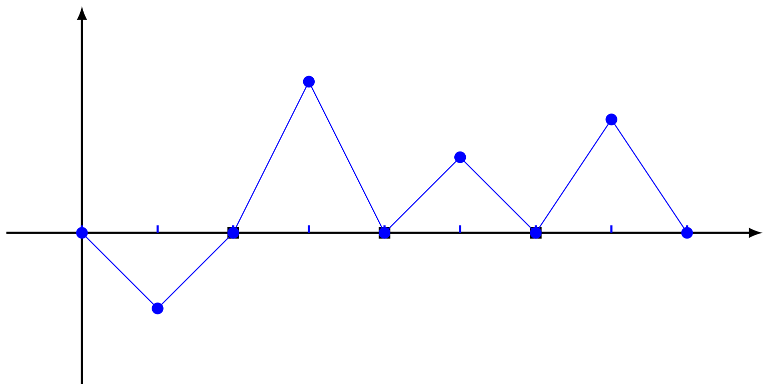
By B_ℓ^V? $0 \le \ell \le L$, we denote the family of Lagrange nodal bases of V_ℓ

$$B_\ell^V := \left\{ \psi_{\ell,j} \right\}_{j=1}^{n_\ell},$$

with the property that

$$\psi_{\ell,j}\left(\vec{N}_{\ell,i} \right) = \delta_{ij}, \quad 1 \le i, j \le n_\ell.$$

Example. *d=1, $\Omega = (0,1)$, $\ell = 2$,*



Lemma 6.7. *For the spaces $W_\ell \subseteq V_\ell$ as in the last definition, we have*

$$V_\ell = V_{\ell-1} \oplus W_\ell \quad 1 \le \ell \le L, \tag{6.9}$$

and

$$V_L = V_0 \oplus V_1 \oplus \cdots \oplus W_L. \tag{6.10}$$

Proof. Let $\mathcal{I}_\ell : C^0(\overline{\Omega}) \to V_\ell$ be the standard Lagrange linear nodal interpolation operator. It has the property that

$$\mathcal{I}_\ell(v)\left(\vec{N}_{\ell,i}\right) = v\left(\vec{N}_{\ell,i}\right), \quad 1 \le i \le n_\ell. \tag{6.11}$$

Let $v_\ell \in V_\ell$ be a given arbitrary function. Write

$$v_\ell = \mathcal{I}_{\ell-1}(v_\ell) + \{v_\ell - \mathcal{I}_{\ell-1}(v_\ell)\}.$$

Clearly

$$\mathcal{I}_{\ell-1}(v_\ell) \in V_{\ell-1},$$

and

$$v_\ell - \mathcal{I}_{\ell-1}(v_\ell) \in W_\ell.$$

Indeed

$$v_\ell\left(\vec{N}_{\ell-1,j}\right) - \mathcal{I}_{\ell-1} v_\ell\left(\vec{N}_{\ell-1,j}\right) = 0.$$

for every $1 \le j \le n_{\ell-1}$. And, this decomposition must be unique. Suppose not. Then

$$v_\ell = v_{\ell-1}^{(i)} + w_\ell^{(i)}, \quad i = 1, 2.$$

So,

$$\begin{aligned} 0 &= \left(v_{\ell-1}^{(1)} - v_{\ell-1}^{(2)}\right) + \left(w_\ell^{(1)} - w_\ell^{(2)}\right) \\ &=: v_{\ell-1} - w_\ell. \end{aligned}$$

Therefore

$$V_{\ell-1} \ni v_\ell = w_\ell \in W_\ell.$$

Clearly, both functions must be identically zero. This proves (6.9). Identity (6.10) follows from (6.9). □

Definition 6.8. *For $1 \le \ell \le L$, define $B_\ell^W := \left\{\phi_{\ell,i}\right\}_{i=1}^{m_\ell} \subset W_\ell$ such that*

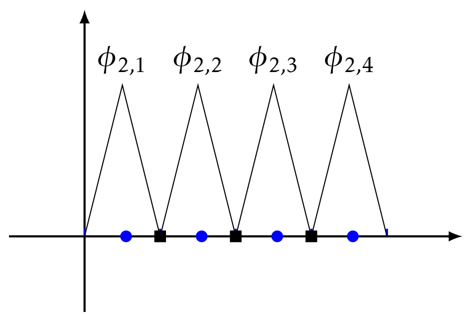
$$W_\ell \ni \phi_{\ell,i}\left(\vec{N}_{\ell,j}^W\right) = \delta_{ij}, \quad 1 \le i,j \le m_\ell,$$

where

$$\left\{\vec{N}_{\ell,j}^W\right\}_{j=1}^{m_\ell} := \left\{\vec{N}_{\ell,j}\right\}_{j=1}^{n_\ell} \setminus \left\{\vec{N}_{\ell,j}\right\}_{j=1}^{n_{\ell-1}}.$$

$B_0^W = B_0^V$.

Example. *d=1*, $\Omega = (0,1)$, $\ell = 2$, $n_1 = 3$, $m_2 = 4, n_2 = 7$



Lemma 6.9. *Defintions (6.6) and (6.8) are consisted in the definition of m_ℓ, and moreover, B_ℓ^W is a basis for W_ℓ, for each $1 \le \ell \le L$.*

Proof. Exercise. □

Lemma 6.10. *Suppose $W_\ell \subseteq V_\ell, 0 \le \ell \le L$, are as defined in Definition .(6.6). Then*

$$H_\ell := \cup_{j=0}^{\ell} B_j^W$$

is a basis for V_ℓ, for any $1 \le \ell \le L$.

Proof. The result follows if we can show that

$$span(H_\ell) = V_\ell,$$

and H_ℓ is linearly independent.

Suppose $v_\ell \in V_\ell$ is arbitrary. Then, there exist unique $w_j \in W_j$, $0 \le j \le \ell$, such that

$$v_\ell = w_0 + w_1 + \cdots + w_\ell.$$

In fact, we can write down this decomposition explicitly:

$$v_\ell = \mathcal{I}_0 v_\ell + (\mathcal{I}_1 v_\ell - \mathcal{I}_0 v_\ell) + (\mathcal{I}_2 v_\ell - \mathcal{I}_1 v_\ell) + \cdots + (v_\ell - \mathcal{I}_{\ell-1} v_\ell)$$

Setting

$$w_0 := \mathcal{I}_0 v_\ell,$$
$$w_j := \mathcal{I}_j v_\ell - \mathcal{I}_{j-1} v_\ell, \quad 1 \le j \le \ell$$

and noticing that

$$\mathcal{I}_\ell v_\ell = v_\ell,$$

gives the result. Now, since $w_j \in W_j$, $0 \le j \le \ell$, there are unique coefficients $C_{j1}, \cdots, C_{jm_j} \in \mathbb{R}$ such that

$$w_j = \sum_{k=1}^{m_j} C_{jk}\phi_{jk}.$$

Hence

$$v_\ell = \sum_{j=0}^{\ell} w_j = \sum_{j=0}^{\ell}\sum_{k=1}^{m_j} C_{jk}\phi_{jk}$$

and

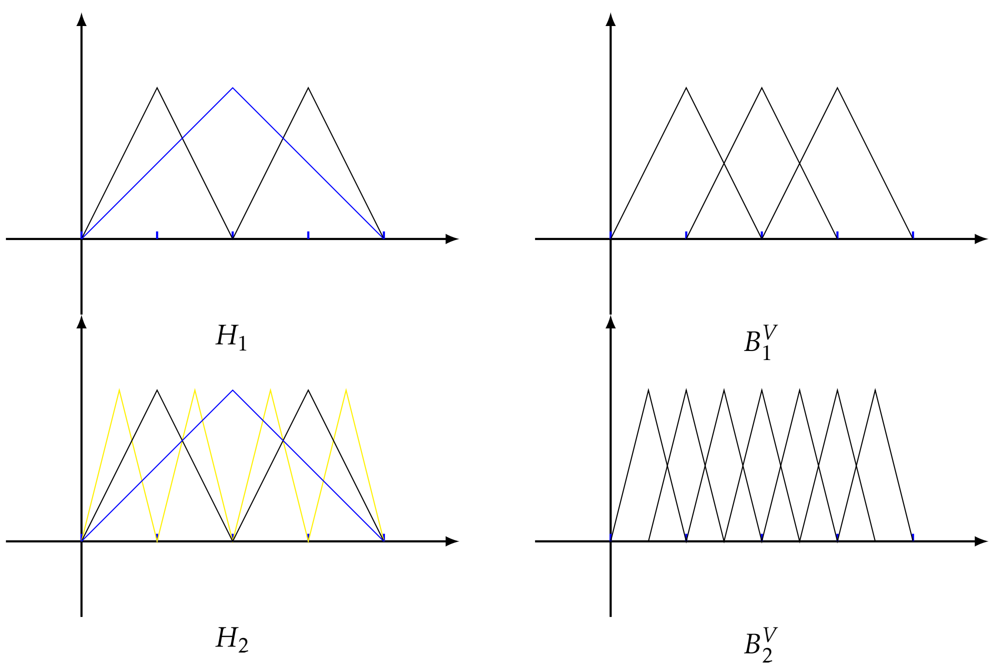
$$V_\ell \subseteq span(H_\ell).$$

On the other hand, it should be clear that

$$span(H_\ell) \subseteq V_\ell.$$

Since $\#(H_\ell) = \#B_\ell^V = n_\ell$, it follows that H_ℓ is linearly independent, since B_ℓ^V is a basis. $\square$

Example. *d=1*



Now, we need to connect the spaces W_j to V_ℓ where $0 \le j \le \ell$.

Remark 6.11. *Clearly*

$$W_j \subset V_j \subset V_\ell, \quad 0 \le j \le \ell.$$

Moreover precisely, for each $0 \le j \le \ell$, and each $1 \le i \le m_j$

$$\phi_{j,i} \in V_\ell.$$

Therefore, there are unique numbers

$$q_{j,k,i}^{\ell} \in \mathbb{R}, \quad 1 \le k \le n_\ell, \ 1 \le i \le m_j,$$

such that

$$\phi_{j,i} = \sum_{k=1}^{n_\ell} q_{j,k,i}^{\ell} \psi_{\ell,k}.$$

Definition 6.12. *Define the matrix $Q_j^\ell \in \mathbb{R}^{n_\ell \times m_j}$ via*

$$\left[Q_j^\ell\right]_{i,k} := q_{j,i,k}^{\ell}.$$

Q_j^ℓ is called an auxiliary prolongation matrix.

Lemma 6.13. *Suppose that $\vec{w}_j \in \mathbb{R}^{m_j}$ is the coordinate vector of the function $w_j \in W_j$ w.r.t. the basis B_j^W. The coordinate vector of $w_j \in V_\ell$ in the basis B_ℓ^V is simply*

$$Q_j^\ell \vec{w}_j \in \mathbb{R}^{n_\ell}.$$

Proof. Exercise. □

Remark 6.14. *Note that the family of spaces W_j are hierarchical, but are not nested*

$$W_0 \not\subset W_1 \not\subset W_2 \cdots.$$

It makes no sense to stack the prolongation matrices

$$Q_j^\ell \neq Q_k^\ell Q_j^k,$$

for $j < k < \ell$. In fact, the product is not defined.

Definition 6.15. *Define the operator $B_j : W_j \to W_j'$ via*

$$B_j[w_1](w_2) = \sum_{i=1}^{m_j} w_1\left(\vec{N}_{j,i}^W\right) w_2\left(\vec{N}_{j,i}^W\right)$$

Define the matrix $B_j \in \mathbb{R}^{m_j \times m_j}$ via

$$\begin{aligned}
\left[B_j\right]_{i,k} : &= B_j[\phi_{j,i}](\phi_{j,k}) \\
&= \sum_{r=1}^{m_j} \phi_{ji}\left(\vec{N}_{j,r}^W\right) \phi_{jk}\left(\vec{N}_{j,r}^W\right) \\
&= \sum_{r=1}^{m_j} \delta_{ir}\delta_{rk} = \delta_{ik}.
\end{aligned}$$

Definition 6.16. *Let $A_L \in \mathbb{R}^{n_L \times n_L}$ be the SPD matrix defined via*

$$[A_L]_{ij} = a(\phi_{L,j}, \phi_{L,i}), \ \ 1 \le i, j \le n_L,$$

where

$$a(u,v) = (\nabla u, \nabla v)_{L^2}, \ \ \forall u, v \in H_0^1(\Omega).$$

The hierarchical basis preconditioner for A_L is defined as

$$C_H = \sum_{\ell=0}^{L} Q_\ell^L C_\ell^{-1} Z_\ell^L \overset{??}{=} \sum_{\ell=0}^{L} Q_\ell^L Z_\ell^L, \tag{6.12}$$

where $Q_L \in \mathbb{R}^{n_L \times m_\ell}$ is the auxiliary prolongation matrix from the Definition.6.12 and

$$Z_\ell^L = \left(Q_\ell^L\right)^T.$$

Lemma 6.17. *Assumption (S1) holds for the hierarchical basis decomposition. In particular, for any $u_L \in V_L$, there exist unique $w_\ell \in W_\ell, 0 \le \ell \le L$, such that*

$$u_L = \sum_{\ell=0}^{L} w_\ell.$$

equivalently, for any $\vec{u}_L \in \mathbb{R}^{n_L}$, there are unique vectors $w_\ell \in \mathbb{R}^{m_\ell}$, such that

$$\vec{u}_L = \sum_{\ell=0}^{L} Q_\ell^L w_\ell.$$

Consequently, C_H defined in (6.12) is SPD.

Proof. This follows from lemma.6.7 and lemma.6.3. □

Remark 6.18. *Our goal is now to show that*

$$\lambda_{\min}(C_H A_L) \ge C_1 \left(1 + |\ln h_L|^2\right)^{-1},$$

and

$$\lambda_{\max}(C_H A_L) \le C_2,$$

where $C_1, C_2 > 0$ are independent of h, using Theorem.6.5. In this case

$$\frac{\lambda_{\max}}{\lambda_{\min}} =: \kappa(C_H A_L) \le \frac{C_2}{C_1}\left(1 + |\ln h_L|^2\right),$$

suppose that

$$h_L = (1/2)^L,$$

then

$$|\ln h_L|^2 = L^2 \left|\ln(1/2)\right|^2.$$

This analysis will only work for $d = 2$.

Now, we need some technical results, for more details see [2].

Lemma 6.19. *For any $0 \le \ell \le L$, with $\mathcal{I}_{-1} \equiv 0$,*

$$\|\mathcal{I}_\ell u_L - \mathcal{I}_{\ell-1} u_L\|_{L^2(\Omega)} \le C h_\ell \left(1 + \sqrt{L-\ell}\right) |u_L|_{H^1(\Omega)}. \tag{6.13}$$

for all $u_L \in V_L$, where $\Omega \subset \mathbb{R}^2$ (i.e. d=2).

Proof. Define the piecewise constant function $\bar{u}_L^\ell$ such that

$$\bar{u}_L^\ell|_K := \frac{1}{|K|}\int_K u_L(\vec{x})d\vec{x}.$$

Then

$$\begin{aligned}
\|\mathcal{I}_\ell u_L - \mathcal{I}_{\ell-1} u_L\|_{L^2(\Omega)}^2 &= \left\|\mathcal{I}_\ell u_L - \mathcal{I}_{\ell-1}(\mathcal{I}_\ell u_L)\right\|_{L^2(\Omega)}^2 \\
&\le c h_\ell^2 \sum_{K \in \mathcal{T}_h} |\mathcal{I}_\ell u_L|_{H^1(K)}^2 \quad \text{(interpolation error)} \\
&= c h_\ell^2 \sum_{K \in \mathcal{T}_h} \left|\mathcal{I}_\ell u_L - \bar{u}_L^\ell\right|_{H^1(K)}^2 \quad (\bar{u}_L^\ell = \text{constant}) \\
&\le c h_\ell^2 \sum_{K \in \mathcal{T}_h} \left\|\mathcal{I}_\ell u_L - \bar{u}_L^\ell\right\|_{L^\infty(K)}^2 \quad \text{(inverse inequality)} \\
&\le c h_\ell^2 \sum_{K \in \mathcal{T}_h} \left\|u_L - \bar{u}_L^\ell\right\|_{L^\infty(K)}^2 \quad (u_L \in C^0(\bar{\Omega})).
\end{aligned}$$

Next, we use the estimate (for d=2):

$$\left\|u_L - \bar{u}_L^\ell\right\|_{L^\infty(K)}^2 \le C(1 + ln(h_\ell)/h_L)\,|u_L|_{H^1(K)}^2.$$

This estimate requires that

$$\int_K (u_L - \bar{u}_L^\ell) d\vec{x} = 0.$$

See chapter. 7 of [2] for details. Now notice that

$$h_\ell = h_0 2^{-\ell} \quad 1 \le \ell \le L.$$

So,

$$\ln(h_\ell/h_L) = \ln(2^{L-\ell}) = (L-\ell)\ln 2.$$

The result follows. ☐

Lemma 6.20. *There is some constant $C_1 > 0$ such that*

$$\lambda_{\min}(C_H A_L) \ge C_1(1 + |\ln h_L|^2)^{-1}. \tag{6.14}$$

Proof. By definition for any $w_{\ell,1}, w_{\ell,2} \in W_\ell$

$$C_\ell[w_{\ell,1}](w_{\ell,2}) = \sum_{i=1}^{m_\ell} w_{\ell,1}(\vec{N}_{\ell,i}^W) w_{\ell,2}(\vec{N}_{\ell,i}^W).$$

Let

$$\vec{w}_{\ell,\alpha} \in \mathbb{R}^{m_\ell} \overset{\mathcal{B}_\ell^W}{\leftrightarrow} w_{\ell,\alpha}, \quad \alpha = 1,2.$$

Then

$$\begin{aligned}
(C_\ell \vec{w}_{\ell,1}, \vec{w}_{\ell,2})_\ell &= \sum_{i=1}^{m_\ell} [\vec{w}_{\ell,1}]_i [\vec{w}_{\ell,2}]_i \\
&= \sum_{i=1}^{m_\ell} w_{\ell,1}(\vec{N}_{\ell,i}^W) w_{\ell,2}(\vec{N}_{\ell,i}^W) \\
&= C_\ell [w_{\ell,1}](w_{\ell,2}) \\
&= C_\ell [w_{\ell,2}](w_{\ell,1}) \\
&=: \left\langle w_{\ell,1}, w_{\ell,2} \right\rangle_{C_\ell}.
\end{aligned}$$

This is like a mass-lumping inner product. All that is missing is a factor of h_ℓ^2. As in Lemma.5.14, there are constant $\tilde{C}_1, \tilde{C}_2 > 0$ such that, for all $0 \le \ell \le L$,

$$\tilde{C}_1 h_\ell^2 \left\langle w_{\ell,\alpha}, w_{\ell,\alpha} \right\rangle_{C_\ell} \le \left\| w_{\ell,\alpha} \right\|_{L^2(\Omega)}^2 \le \tilde{C}_2 h_\ell^2 \left\langle w_{\ell,\alpha}, w_{\ell,\alpha} \right\rangle_{C_\ell} \tag{6.15}$$

Therefore, for any $w_\ell \in W_\ell \overset{\mathcal{B}_\ell^W}{\leftrightarrow} \vec{w}_\ell \in \mathbb{R}^{m_\ell}$,

$$\begin{aligned}
(C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell &= h_\ell^{-2} h_\ell^2 \langle w_\ell, w_\ell \rangle_{C_\ell} \\
&\overset{6.15}{\le} \tilde{\tilde{C}}_1 h_\ell^{-2} \| w_\ell \|_{L^2(\Omega)}^2 \\
&\overset{\mathcal{I}_{-1}=0}{=} \tilde{\tilde{C}}_1 h_\ell^{-2} \| w_\ell - \mathcal{I}_{\ell-1} w_\ell \|_{L^2(\Omega)}^2 \\
&\le \tilde{\tilde{C}}_2 | w_\ell |_{H^1(\Omega)}^2 \qquad \text{(interp. error)} \\
&\le \tilde{\tilde{C}}_3 h_\ell^{-2} \| w_\ell \|_{L^2(\Omega)}^2 \qquad \text{(inverse ineq.)} \\
&\overset{6.15}{\le} \tilde{\tilde{C}}_4 (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell .
\end{aligned} \tag{6.16}$$

Therefore, there are constants $\tilde{\tilde{C}}_5, \tilde{\tilde{C}}_6 > 0$ such that we have the equivalence

$$\tilde{\tilde{C}}_5 \sum_{\ell=0}^{L} | w_\ell |_{H^1(\Omega)}^2 \le \sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell \le \tilde{\tilde{C}}_6 \sum_{\ell=0}^{L} | w_\ell |_{H^1(\Omega)}^2, \tag{6.17}$$

for any $w_\ell \in W_\ell$, in general. Now, let $u_L \in V_L$ be given and

$$u_L = \sum_{\ell=0}^{L} w_\ell, \exists ! w_\ell \in W_\ell.$$

Recall that

$$w_\ell = \mathcal{I}_\ell u_L - \mathcal{I}_{\ell-1} u_L, \quad 1 \le \ell \le L,$$

and

$$w_0 = \mathcal{I}_0 u_L.$$

Then, from (6.16)

$$\sum_{\ell=0}^{L} (C_\ell \vec{w}_\ell, \vec{w}_\ell)_\ell \le \tilde{\tilde{C}}_1 \sum_{\ell=0}^{L} h_\ell^{-2} \| w_\ell \|_{L^2(\Omega)}^2$$

$$\begin{aligned}
&\overset{(6.13)}{\leq} && C\sum_{\ell=0}^{L}\left(1+\sqrt{L-\ell}\right)^2 |u_L|^2_{H^1(\Omega)} \\
&\leq && C\sum_{\ell=0}^{L}(1+L-\ell)\,|u_L|^2_{H^1(\Omega)} \\
&\leq && C\left(1+L+L^2\right)|u_L|^2_{H^1(\Omega)} \\
&\overset{L\geq 1}{\leq} && CL^2\,|u_L|^2_{H^1(\Omega)}.
\end{aligned}$$

But

$$\begin{aligned}
|u_L|^2_{H^1(\Omega)} &= (\nabla u_L, \nabla u_L) \\
&= a(u_L, u_L) \\
&= (A\vec{u}_L, \vec{u}_L).
\end{aligned}$$

And

$$\begin{aligned}
|\ln h_L|^2 &= \left|\ln(h_0 2^{-L})\right|^2 \\
&= \left|\ln(h_0) - L\ln(2)\right|^2 \\
&= \ln^2(h_0) - 2\ln(h_0)L\ln(2) + L^2\ln^2(2).
\end{aligned}$$

So

$$L^2 \leq C\left(1+\left|\ln(h_L)\right|^2\right), \quad \exists C>0.$$

Thus,

$$\sum_{\ell=0}^{L}(C_\ell\vec{w}_\ell, \vec{w}_\ell) \leq C\left(1+\left|\ln(h_L)\right|^2\right)(A\vec{u}_L, \vec{u}_L),$$

and it follows from Theorem.6.5 that

$$\lambda_{\min}(C_H A_L) \geq C_1\left(1+|\ln h_L|^2\right)^{-1}.$$

□

Lemma 6.21. *Let $a_j, b_j \geq 0, -\infty < j < \infty$, with*

$$s_1 := \sum_{j=-\infty}^{\infty} a_j \leq \infty,$$

and

$$s_2 := \sum_{j=-\infty}^{\infty} b_j \leq \infty.$$

Then

$$\sum_{j=-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty} a_{j-k}b_k\right) \leq s_1^2 s_2. \tag{6.18}$$

Proof. Exercise. □

Lemma 6.22. *For any $v_\ell \in V_\ell$ and $v_k \in V_k$, $0 \le \ell \le k \le L$, and d=2, $\exists C > 0$ such that*

$$\int_\Omega \nabla v_\ell \nabla v_k d\vec{x} \le C 2^{(\ell-k)/2} |v_\ell|_{H^1(\Omega)} \left(h_k^{-1} \|v_k\|_{L^2(\Omega)} \right). \tag{6.19}$$

Proof. For any $K \in \mathcal{T}_h$, since $\Delta v_\ell|_K \equiv 0$,

$$\begin{aligned}
\int_K \nabla v_\ell \nabla v_k d\vec{x} &= \int_{\partial K} \frac{\partial v_\ell}{\partial n} v_k ds \\
&\le C h_\ell^{-1} |v_\ell|_{H^1(K)} \int_{\partial K} v_k ds \\
&\le \left(C h_\ell^{-1} |v_\ell|_{H^1(K)} \right) \left(h_k \sum_{\vec{N}_k \in \partial K} v_k(\vec{N}_k) \right) \\
&\overset{C.S.}{\le} \left(C h_\ell^{-1} |v_\ell|_{H^1(K)} \right) \left(h_k \left(\frac{h_\ell}{h_k} \right)^{1/2} \left(\sum_{\vec{N}_k \in \partial K} \left| v_k(\vec{N}_k) \right|^2 \right)^{1/2} \right) \\
&\overset{(5.4)}{\le} C \left(\frac{h_\ell}{h_k} \right)^{1/2} |v_\ell|_{H^1(K)} h_k^{-1} \|v_k\|_{L^2(K)}.
\end{aligned}$$

Thus

$$\begin{aligned}
\int_\Omega \nabla v_\ell \nabla v_k d\vec{x} &= \sum_{K \in \mathcal{T}_h} \int_K \nabla v_\ell \nabla v_k d\vec{x} \\
&\le C 2^{(\ell-k)/2} \sum_{K \in \mathcal{T}_h} |v_\ell|_{H^1(K)} h_k^{-1} \|v_k\|_{L^2(K)} \\
&\overset{C.S.}{\le} C 2^{(\ell-k)/2} |v_\ell|_{H^1(\Omega)} h_k^{-1} \|v_k\|_{L^2(\Omega)}.
\end{aligned}$$

□

Lemma 6.23. *(Strengthened Cauchy-Schwarz Inequality) For any $w_\ell \in W_\ell$ and $w_k \in W_k$, $0 \le \ell \le k \le L$,*

$$\int_\Omega \nabla w_\ell \nabla w_k d\vec{x} \le C 2^{(\ell-k)/2} |w_\ell|_{H^1(\Omega)} |w_k|_{H^1(\Omega)}. \tag{6.20}$$

Proof. Observe that

$$w_k = w_k - \mathcal{I}_{\ell-1}(w_k).$$

We use the interpolation error estimate

$$\left\| w_k - \mathcal{I}_{k-1}(w_k) \right\|_{L^2(\Omega)} \le C h_k |w_k|_{H^1(\Omega)},$$

to conclude that

$$\|w_k\|_{L^2(\Omega)} \le C h_k |w_k|_{H^1(\Omega)}.$$

Now, we use the last result. Since $w_\ell \in V_\ell$ and $w_k \in V_k$,

$$\int_\Omega \nabla w_\ell \nabla w_k d\vec{x} \le C^{(\ell-k)/2} |w_\ell|_{H^1(\Omega)} h_k^{-1} \|w_k\|_{L^2(\Omega)}$$

$$\leq \quad C^{(\ell-k)/2}|w_\ell|_{H^1(\Omega)}|w_k|_{H^1(\Omega)}$$

□

Lemma 6.24. *There is a constant $C_2 > 0$ such that*

$$\lambda_{\max}(C_H A_L) \leq C_2,$$

independent of L.

Proof. Let $v_L \in V_L$ be arbitrary.

$$v_L \in V_L \overset{\mathcal{B}_L}{\leftrightarrow} \vec{v}_L \in \mathbb{R}^{n_L}.$$

There exist unique $w_\ell \in W_\ell \overset{\mathcal{B}_\ell^W}{\leftrightarrow} \vec{w}_\ell \in \mathbb{R}^{m_\ell}$ such that

$$v_L = \sum_{\ell=0}^{L} w_\ell \leftrightarrow \vec{v}_L = \sum_{\ell=0}^{L} Q_\ell^L \vec{w}_\ell.$$

Then

$$\begin{aligned}
(\vec{v}_L, \vec{v}_L)_{A_L} &= (\vec{v}_L, A_L \vec{v}_L) \\
&= a(\vec{v}_L, \vec{v}_L) \\
&= a\left(\sum_{\ell=0}^{L} w_\ell, \sum_{k=0}^{L} w_k\right) \\
&= \int_\Omega \left(\nabla \sum_{\ell=0}^{L} w_\ell\right)\left(\nabla \sum_{k=0}^{L} w_k\right) d\vec{x} \\
&= \sum_{\ell,k=0}^{L} \int_\Omega \nabla w_\ell \nabla w_k d\vec{x} \\
&\overset{(6.20)}{\leq} C \sum_{\ell,k=0}^{L} 2^{-|\ell-k|/2} |w_\ell|_{H^1(\Omega)} |w_k|_{H^1(\Omega)} \\
&\leq C \sum_{\ell=0}^{L} \left(\sum_{k=0}^{L} 2^{-|\ell-k|/2} |w_k|_{H^1(\Omega)}\right) |w_\ell|_{H^1(\Omega)} \qquad (6.21) \\
&\overset{C.S.}{\leq} C \left\{\sum_{\ell=0}^{L} \left(\sum_{k=0}^{L} 2^{-|\ell-k|/2} |w_k|_{H^1(\Omega)}\right)^2\right\}^{1/2} \left\{\sum_{\ell=0}^{L} |w_\ell|^2_{H^1(\Omega)}\right\}^{1/2} \\
&\overset{(6.18)}{\leq} C \left\{\sum_{\ell=0}^{L} |w_\ell|^2_{H^1(\Omega)}\right\}^{1/2} \left\{\sum_{\ell=0}^{L} |w_\ell|^2_{H^1(\Omega)}\right\}^{1/2} \\
&= c \sum_{\ell=0}^{L} |w_\ell|^2_{H^1(\Omega)} \\
&\overset{(6.17)}{\leq} C_2 \sum_{\ell=0}^{L} (\vec{w}_\ell, \vec{w}_\ell)_{C_\ell}.
\end{aligned}$$

Recall that, since decomposition are unique

$$\begin{aligned}
\lambda_{\max}(C_H A_L) &\overset{6.5}{=} \max_{\vec{u}_L \in \mathbb{R}_*^{n_L}} \frac{(\vec{u}_L, \vec{u}_L)_{A_L}}{\sum_{\ell=0}^{L} (\vec{w}_\ell, \vec{w}_\ell)_{C_\ell}} \\
&\overset{6.21}{=} \max_{\vec{u}_L \in \mathbb{R}_*^{n_L}} \frac{\mathbb{C}_2 \sum_{\ell=0}^{L} (\vec{w}_\ell, \vec{w}_\ell)_{C_\ell}}{\sum_{\ell=0}^{L} (\vec{w}_\ell, \vec{w}_\ell)_{C_\ell}} \\
&\leq C_2.
\end{aligned}$$

□

Theorem 6.25. *There is a constant $C > 0$ independent of L, such that*

$$\kappa(C_H A_L) = \frac{\lambda_{\max}(C_H A_L)}{\lambda_{\min}(C_H A_L)} \leq C\left(1 + |\ln h_L|^2\right).$$

independent of L.

Proof. Follows from Lemma.6.20 and 6.23 □

6.2 The BPX Preconditioner

For this method we choose

$$W_\ell := V_\ell, \quad 0 \leq \ell \leq L.$$

Thus

$$W_L = V_L$$

and

$$m_\ell = n_\ell, \quad 0 \leq \ell \leq L.$$

Definition 6.26. *Define the operator $C_\ell : V_\ell \to V_\ell'$ via*

$$C_\ell[v_{\ell,1}](v_{\ell,2}) = \sum_{i=1}^{n_\ell} w_{\ell,1}(\vec{N}_{\ell,i}^W) w_{\ell,2}(\vec{N}_{\ell,i}^W)$$

The matrix $C_\ell \in \mathbb{R}^{m_\ell \times m_\ell}$ is defined as

$$[C_\ell]_{jk} = C_\ell\left[\phi_{\ell,j}\right](\phi_{\ell,k}) = \delta_{jk}, \quad 1 \leq j,k \leq n_\ell,$$

where $\mathcal{B}_\ell = \left\{\phi_{\ell,j}\right\}_{j=1}^{n_\ell}$ is the Lagrange nodal basis for the piecewise linear FE space $V_\ell, 0 \leq \ell \leq L$. The BPX preconditioner is

$$C_{BPX} := \sum_{\ell=0}^{L} P_\ell^L C_\ell^{-1} \mathcal{R}_\ell^L = \sum_{\ell=0}^{L} P_\ell^L \mathcal{R}_\ell^L, \tag{6.22}$$

where $P_\ell^L \in \mathbb{R}^{n_L \times n_\ell}$ is the standard prolongation matrix from Chap.5 and $\mathcal{R}_\ell^L = \left(P_\ell^L\right)^T$.

Lemma 6.27. *Assumption (S1) holds for the BPX framework, i.e., for every $u_L \in V_L$, there exists $v_\ell \in V_\ell, 0 \le \ell \le L$, such that*

$$u_L = \sum_{\ell=0}^{L} v_\ell,$$

or, equivalently

$$\vec{u}_L = \sum_{\ell=0}^{L} P_\ell^l \vec{v}_\ell,$$

with

$$V_\ell \ni v_\ell \overset{\mathcal{B}_\ell}{\leftrightarrow} \vec{v}_\ell \in \mathbb{R}^{n_\ell},$$

and

$$V_L \ni u_L \overset{\mathcal{B}_\ell}{\leftrightarrow} \vec{u}_L \in \mathbb{R}^{n_L}.$$

Proof. This is trivial because of the nestedness of the the spaces

$$V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{L-1} \subset V_l.$$

□

Remark 6.28. *Note that the decomposition os no longer unique.*

Lemma 6.29. *For any $v_j \in V_j, v_\ell \in V_\ell$,*

$$\int_\Omega \nabla v_j \nabla v_\ell d\vec{x} \le C 2^{-|j-\ell|/2} \left(h_j^{-1} \left\| v_j \right\|_{L^2(\Omega)} \right) \left(h_\ell^{-1} \left\| v_\ell \right\|_{L^2(\Omega)} \right), \tag{6.23}$$

for some $C > 0$.

Proof. This is follows from (6.19) and the inverse inequality

$$\left| v_j \right|_{H^1(\Omega)} \le c h_j^{-1} \left\| v_j \right\|_{L^2(\Omega)}.$$

□

Lemma 6.30. *For some $C_2 > 0$ that is independent of L,*

$$\lambda_{\max}\left(B_{BPX} A_L\right) \le C_2.$$

for some $C > 0$.

Proof. Let $u_L \in V_L$ be arbitrary. There exists $v_\ell \in V_\ell, 0 \le \ell \le L$, such that

$$u_L = \sum_{\ell=0}^{L} v_\ell,$$

or

$$\vec{u}_L = \sum_{\ell=0}^{L} P_\ell^l \vec{v}_\ell.$$

The decomposition is not unique, however. Then

$$
\begin{aligned}
(\vec{u}_L,\vec{u}_L)_{A_L} &= (\vec{u}_L, A_L\vec{u}_L) \\
&= a(\vec{u}_L,\vec{u}_L) \\
&= a\left(\sum_{j=0}^{L} v_j, \sum_{\ell=0}^{L} v_\ell\right) \\
&= \sum_{\ell,j=0}^{L} a(v_j,v_\ell) \\
&\overset{(6.23)}{\leq} C\sum_{\ell,j=0}^{L} 2^{-|j-\ell|/2} h_j^{-1}\|v_\ell\|_{L^2(\Omega)} h_\ell \|v_k\|_{L^2(\Omega)} \\
&\overset{(6.18)}{\leq} C\sum_{j=0}^{L} h_j^{-2}\left\|v_j\right\|_{L^2(\Omega)} \\
&\overset{(5.4)}{\leq} C_2\sum_{j=0}^{L}\left(\vec{v}_j,\vec{v}_j\right)_{C_j} \\
&= C_2\sum_{j=0}^{L}\left(C_j\vec{v}_j,\vec{v}_j\right)_j
\end{aligned}
$$

Now,

$$
\begin{aligned}
\lambda_{\max}(C_{BPX}A_L) &\overset{6.5}{=} \max_{\vec{u}_L\in\mathbb{R}_*^{n_L}} \frac{(\vec{u}_L,\vec{u}_L)_{A_L}}{\displaystyle\min_{\vec{u}_L=\sum_{\ell=0}^{L}P_\ell^L\vec{v}_\ell'} \sum_{\ell=0}^{L}\left(\vec{u}_\ell',\vec{u}_\ell'\right)_{C_\ell}} \\
&\leq \max_{\vec{u}_L\in\mathbb{R}_*^{n_L}} \frac{C_2\sum_{\ell=0}^{L}(C_\ell\vec{w}_\ell,\vec{w}_\ell)_\ell}{\displaystyle\min_{\vec{v}_\ell'}\sum_{\ell=0}^{L}(C_\ell\vec{w}_\ell,\vec{w}_\ell)} \\
&\leq C_2.
\end{aligned}
$$

Recall that the minimum was achievable, so we could take $\vec{v}_\ell = \vec{v}_\ell'$. wee the proof of Theorem.6.4. □

Lemma 6.31. *There is a constant $C_1 > 0$ that is independent of L, such that*

$$\lambda_{\min}(B_{BPX}A_L) \geq C_1.$$

for some $C > 0$.

Proof. Let $u_L \in V_L$ be arbitrary. Set

$$v_\ell =: \mathcal{R}_\ell u_L - R_{\ell-1}u_L, \quad 0 \leq \ell \leq L,$$

where $\mathcal{R}_\ell : H_0^1(\Omega) \to V_\ell$ is the Ritz projection for $0 \leq \ell \leq L$ and $R_{-1} \equiv 0$. Since

$$\mathcal{R}_\ell u_L = u_L,$$

it follows that

$$u_L = \sum_{\ell=0}^{L} v_\ell \overset{\mathcal{B}_\ell}{\leftrightarrow} \vec{u}_\ell = \sum_{\ell=0}^{L} P_\ell^L v_\ell.$$

Moreover,

$$a(v_j, v_\ell) = 0, \quad o \le j \neq \ell \le L. \tag{6.24}$$

To see this, recall that, in general,

$$a(R_j u_L, v_j') = a(u_L, v_j'), \quad \forall v_j' \in V_j.$$

Suppose $j < \ell$, for definiteness. Then

$$a(R_j u_L, v_\ell') = a(u_L, v_\ell'), \quad \forall v_\ell' \in V_\ell.$$

In particular, since

$$v_j := R_j u_L - R_{j-1} u_L \in V_j \subset V_\ell,$$

and

$$a(\mathcal{R}_\ell u_L, v_j) = a(u_L, v_j),$$

likewise

$$a(R_{\ell-1} u_L, v_j) = a(u_L, v_j),$$

Subtracting, we have

$$a(\mathcal{R}_\ell u_L - R_{\ell-1} u_L, v_j) = 0$$

To make further progress, let us assume that Ω is convex. Then the standard regularity condition holds. And, for $1 \le \ell \le L$

$$\begin{aligned}
h_\ell^{-2} \|v_\ell\|_{L^2(\Omega)}^2 &= h_\ell^{-2} \|\mathcal{R}_\ell u_L - R_{\ell-1} u_L\|_{L^2(\Omega)}^2 \\
&= h_\ell^{-2} \|\mathcal{R}_\ell u_L - R_{\ell-1} \mathcal{R}_\ell u_L\|_{L^2(\Omega)}^2 \\
&\overset{5.13}{\le} C h_\ell^{-2} h_\ell^2 |\mathcal{R}_\ell u_L - R_{\ell-1} \mathcal{R}_\ell u_L|_{H^1(\Omega)}^2 \\
&= C |\mathcal{R}_\ell u_L - R_{\ell-1} \mathcal{R}_\ell u_L|_{H^1(\Omega)}^2 \\
&= C |v_\ell|_{H^1(\Omega)}^2 .
\end{aligned} \tag{6.25}$$

To see that $R_{\ell-1} = R_{\ell-1}\mathcal{R}_\ell$, let $u \in H_0^1(\Omega)$ be arbitrary. Then

$$a(R_{\ell-1}(\mathcal{R}_\ell u), v_{\ell-1}') = a(\mathcal{R}_\ell u, v_{\ell-1}'), \quad \forall v_{\ell-1}' \in V_{\ell-1}.$$

But,

$$a(\mathcal{R}_\ell u, v_{\ell-1}') = a(u, v_{\ell-1}'), \quad \forall v_{\ell-1}' \in V_{\ell-1}.$$

Since

$$a(\mathcal{R}_\ell u, v_\ell') = a(u, v_\ell'), \quad \forall v_\ell' \in V_\ell,$$

and

$$V_{\ell-1} \subset V_\ell.$$

But

$$a(R_{\ell-1} u, v_{\ell-1}') = a(u, v_{\ell-1}'), \quad \forall v_{\ell-1}' \in V_{\ell-1}.$$

Hence

$$a(R_{\ell-1}(\mathcal{R}_\ell u), v_{\ell-1}') = a(R_{\ell-1} u, v_{\ell-1}'), \quad \forall v_{\ell-1}' \in V_{\ell-1}.$$

And we conclude that $R_{\ell-1} = R_{\ell-1}\mathcal{R}_\ell$ since

$$R_{\ell-1}(\mathcal{R}_\ell u), R_{\ell-1}u \in V_{\ell-1}.$$

Estimate (6.24) holds trivially for $\ell = 0$. Finally,

$$\begin{aligned}\sum_{\ell=0}^{L}(C_\ell \vec{v}_\ell, \vec{v}_\ell)_\ell &\overset{(5.4)}{\le} C\sum_{\ell=0}^{L} h_\ell^{-2}\|v_\ell\|_{L^2(\Omega)}^2 \\ &\overset{(6.25)}{\le} C_1^{-1}\sum_{\ell=0}^{L}|v_\ell|_{H^1(\Omega)}^2 \\ &\overset{(6.24)}{=} C_1^{-1}|u_L|_{H^1(\Omega)}^2 .\end{aligned} \tag{6.26}$$

Finally,

$$\begin{aligned}\lambda_{\min}(C_{BPX}A_L) &= \min_{\vec{u}_L\in\mathbb{R}_*^{n_L}} \frac{(\vec{u}_L, \vec{u}_L)_{A_L}}{\min\limits_{\vec{u}_L=\sum_{\ell=0}^{L}P_\ell^L\vec{v}_\ell} \sum_{\ell=0}^{L}\left(\vec{u}_\ell', \vec{u}_\ell'\right)_{C_\ell}} \\ &\ge \min_{\vec{u}_L\in\mathbb{R}_*^{n_L}} \frac{(A_L\vec{u}_L, \vec{u}_L)_L}{\min\limits_{\vec{v}_\ell'}\sum_{\ell=0}^{L}(C_\ell\vec{v}_\ell, \vec{v}_\ell)} \\ &\ge \min_{\vec{u}_L\in\mathbb{R}_*^{n_L}} \frac{(A_L\vec{u}_L, \vec{u}_L)_L}{C_1^{-1}|u_L|_{H^1(\Omega)}} \\ &= C_1 .\end{aligned}$$

□

Theorem 6.32.

$$\kappa(B_{BPX}A_L) = \frac{\lambda_{\max}(B_{BPX}A_L)}{\lambda_{\min}(B_{BPX}A_L)} \le \frac{C_2}{C_1}.$$

Proof. Follows from Lemma.6.30 and 6.31. □

References

[1] D. Braess, *Finite elements: Theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, 2001. 92

[2] S. C. Brenner and R. Scott, *The mathematical theory of finite element methods*, vol. 15, Springer, 2008. 119

[3] Y. Saad, *Iterative methods for sparse linear systems*, Siam, 2003. 30