Math 673/4

Multigrid Methods: A Mostly Matrix-Based Approach

Chapter 05: Multigrid: Algorithms and Abstract Convergence Theory

Abner J. Salgado and Steven M. Wise

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

F24/S25

Chapter 05, Part 3 of 3

Multigrid: Algorithms and Abstract Convergence Theory

A Failure of the Galerkin Condition

A Failure of the Galerkin Condition

In this section, we will describe a particular setting for which the Galerkin conditions fail. We will show that, with a little extra effort, multigrid convergence may still be attained. This general theory, which is based on ideas from the books by Bramble (1993) and Bramble and Zhang (2000), will apply to the analysis of multigrid algorithms for the cell-centered finite difference method discussed in Chapter 07.

Definition (Failure of the Galerkin Condition)

We say that **Assumption** (G3) holds iff

1. For all $0 \leq \ell \leq L$, A_ℓ is a symmetric positive definite matrix.
2. For all $1 \leq \ell \leq L$, $\mathsf{P}_{\ell-1} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ and $\mathsf{R}_{\ell-1} \in \mathbb{R}^{n_{\ell-1} \times n_\ell}$ are full rank matrices satisfying, for some $r \in (0,1)$,
$$\mathsf{R}_{\ell-1} = r\mathsf{P}_{\ell-1}^\top. \tag{G3}$$
3. For each $1 \leq \ell \leq L$,
$$\mathsf{R}_{\ell-1}\mathsf{A}_\ell\mathsf{P}_{\ell-1} = 2\mathsf{A}_{\ell-1}.$$

Remark (Failure of the Galerkin Condition)

If Assumption (G3) *holds, clearly the strong Galerkin condition (G0) fails. We have the equivalent relation*

$$\mathsf{A}_{\ell-1} = \frac{r}{2}\mathsf{P}_{\ell-1}^{\top}\mathsf{A}_{\ell}\mathsf{P}_{\ell-1},$$

for all $1 \leq \ell \leq L$.

Let us look at some clear implications of Assumption (G3).

Proposition (Not A Projection)

Let $1 \leq \ell \leq L$. *Recall that*

$$\tilde{\Pi}_\ell := \mathsf{P}_{\ell-1}\mathsf{A}_{\ell-1}^{-1}\mathsf{R}_{\ell-1}\mathsf{A}_\ell.$$

If assumption (G3) *holds, then*

$$\tilde{\Pi}_\ell^2 = 2\tilde{\Pi}_\ell.$$

In other words, the coarse-grid Ritz projection matrix, $\tilde{\Pi}_\ell$, *is not a projection. However, it is still true that*

$$\tilde{\Pi}_\ell^* = \tilde{\Pi}_\ell,$$

that is, $\tilde{\Pi}_\ell$ *is still self-adjoint in the* A_ℓ *inner product.*

Proof.

Exercise. □

Since the coarse-grid Ritz projection is not really a projection, we introduce a surrogate. For $1 \leq \ell \leq L$, define

$$\hat{\Pi}_\ell := \frac{1}{2}\tilde{\Pi}_\ell. \tag{1}$$

It turns out that this rescaling has suitable properties.

Proposition (Projection)

For $1 \leq \ell \leq L$ we have

$$\hat{\Pi}_\ell^* = \hat{\Pi}_\ell.$$

In addition, if Assumption (G3) *holds,*

$$\hat{\Pi}_\ell^2 = \hat{\Pi}_\ell.$$

In other words, if Assumption (G3) *holds,* $\hat{\Pi}_\ell$ *is a projection matrix.*

Proof.

Exercise. □

From the last result we immediately have the following.

Corollary (Projection)

Let, for all $1 \leq \ell \leq L$, $\hat{\Pi}_\ell$ *be defined in* (1). *Then,*

$$(\mathsf{I}_\ell - \hat{\Pi}_\ell)^* = \mathsf{I}_\ell - \hat{\Pi}_\ell.$$

Furthermore, if Assumption (G3) *holds, then*

$$(\mathsf{I}_\ell - \hat{\Pi}_\ell)^2 = \mathsf{I}_\ell - \hat{\Pi}_\ell.$$

Proof.

Exercise. □

Two-Grid Error Transfer Matrix

Let us now prove that a version of the Two-Grid algorithm converges. To do so, we recall that the two-grid error transfer matrix, E_1, is defined as

$$\mathsf{E}_1 = (\mathsf{K}_1^*)^{m_2} \left(\mathsf{I}_1 - \tilde{\Pi}_1\right) \mathsf{K}_1^{m_1}.$$

Two Definitions for Recall

Definition (α-Weak Approximation Property)

We say that the multigrid algorithm satisfies the α-**weak approximation property** or, equivalently, that **Assumption** (A2) holds, iff there are constants $\alpha \in (0, 1]$ and $C_{\mathrm{A2}} > 0$, such that, for every $1 \leq \ell \leq L$ and all $\boldsymbol{u}_\ell \in \mathbb{R}^{n_\ell}$,

$$\left| \left(\left(\mathsf{I}_\ell - \tilde{\Pi}_\ell \right) \boldsymbol{u}_\ell, \boldsymbol{u}_\ell \right)_{\mathsf{A}_\ell} \right| \leq \frac{C_{\mathrm{A2}}^{2\alpha}}{\rho_\ell^\alpha} \left\| \mathsf{A}_\ell \boldsymbol{u}_\ell \right\|_\ell^{2\alpha} \left\| \boldsymbol{u}_\ell \right\|_{\mathsf{A}_\ell}^{2(1-\alpha)} . \tag{A2}$$

Definition (First Smoothing Property)

We say that the multigrid algorithm satisfies the **first smoothing property** or, equivalently, that **Assumption** (S1), iff there is $C_{\mathrm{S1}} > 0$ such that for every $1 \leq \ell \leq L$ and all $\boldsymbol{u}_\ell \in \mathbb{R}^{n_\ell}$,

$$\left\| \mathsf{K}_\ell^m \boldsymbol{u}_\ell \right\|_{\mathsf{A}_\ell^2} \leq C_{\mathrm{S1}} \rho_\ell^{1/2} m^{-1/2} \left\| \boldsymbol{u}_\ell \right\|_{\mathsf{A}_\ell} . \tag{S1}$$

We also need a stability result from earlier in Chapter 05.

Lemma (Stability)

For every $0 \leq \ell \leq L$ *and all* $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$ *Richardson's smoother satisfies*

$$\left\| \mathsf{K}_\ell \boldsymbol{v}_\ell \right\|_{\mathsf{A}_\ell} \leq \left\| \boldsymbol{v}_\ell \right\|_{\mathsf{A}_\ell}, \tag{2}$$

$$\left(\mathsf{K}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_\ell \leq \left(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_\ell. \tag{3}$$

Theorem (α–Convergence of the Two-Grid Method)

Let $m_1 = m_2 = m \in \mathbb{N}$. Assume that (G3) *and* (A2) *hold, and that the smoothing is performed using Richardson's smoother. We have*

$$\left|(\mathsf{E}_1 \boldsymbol{v}_1, \boldsymbol{v}_1)_{\mathsf{A}_1}\right| \leq \left(\frac{C_{\mathrm{A2}} C_{\mathrm{S1}}}{m^{1/2}}\right)^{2\alpha} (\boldsymbol{v}_1, \boldsymbol{v}_1)_{\mathsf{A}_1},$$

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

Proof.

Let $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$ be arbitrary. The α–weak approximation property (A2) implies that

$$\begin{aligned}\left|(\mathsf{E}_1 \boldsymbol{v}_1, \boldsymbol{v}_1)_{\mathsf{A}_1}\right| &= \left|\left(\left(\mathsf{I}_1 - \tilde{\Pi}_1\right) \mathsf{K}_1^m \boldsymbol{v}_1, \mathsf{K}_1^m \boldsymbol{v}_1\right)_{\mathsf{A}_1}\right| \\ &\leq \frac{C_{\mathrm{A2}}^{2\alpha}}{\rho_1^{\alpha}} \left\|\mathsf{A}_1 \mathsf{K}_1^m \boldsymbol{v}_1\right\|_1^{2\alpha} \left\|\mathsf{K}_1^m \boldsymbol{v}_1\right\|_{\mathsf{A}_1}^{2(1-\alpha)},\end{aligned}$$

for some $C_{\mathrm{A2}} > 0$.

Proof (Cont.)

Next, recall that Richardson's smoother satisfies Assumption (S1). Using this fact, the stability of Richardson's method, and the α–weak approximation property (A2), we have

$$\begin{aligned}
\left|(\mathsf{E}_1 \boldsymbol{v}_1, \boldsymbol{v}_1)_{\mathsf{A}_1}\right| &\leq \frac{C_{\mathrm{A2}}^{2\alpha}}{\rho_1^{\alpha}} \left\|\mathsf{A}_1 \mathsf{K}_1^m \boldsymbol{v}_1\right\|_1^{2\alpha} \left\|\mathsf{K}_1^m \boldsymbol{v}_1\right\|_{\mathsf{A}_1}^{2(1-\alpha)} \\
&\overset{(2)}{\leq} \frac{C_{\mathrm{A2}}^{2\alpha}}{\rho_1^{\alpha}} \left\|\mathsf{A}_1 \mathsf{K}_1^m \boldsymbol{v}_1\right\|_1^{2\alpha} \left\|\boldsymbol{v}_1\right\|_{\mathsf{A}_1}^{2(1-\alpha)} \\
&\overset{(\mathrm{S1})}{\leq} \frac{C_{\mathrm{A2}}^{2\alpha}}{\rho_1^{\alpha}} C_{\mathrm{S1}}^{2\alpha} \rho_1^{\alpha} m^{-\alpha} \left\|\boldsymbol{v}_1\right\|_{\mathsf{A}_1}^{2} \\
&= \left(\frac{C_{\mathrm{A2}} C_{\mathrm{S1}}}{m^{1/2}}\right)^{2\alpha} (\boldsymbol{v}_1, \boldsymbol{v}_1)_{\mathsf{A}_1} .
\end{aligned}$$

□

Corollary (α-Convergence of Two-Grid Method)

In the setting of the previous theorem, we have

$$\left\| \mathsf{E}_1 \boldsymbol{v}_1 \right\|_{\mathsf{A}_1} \leq \left(\frac{C_{\mathrm{A}2} C_{\mathrm{S}1}}{m^{1/2}} \right)^{2\alpha} \left\| \boldsymbol{v}_1 \right\|_{\mathsf{A}_1},$$

for all $\boldsymbol{v}_1 \in \mathbb{R}^{n_1}$. *In other words, the two-grid method converges, provided* $m > 0$ *is sufficiently large.*

Proof.

Exercise. □

To conclude this section and, in fact, this chapter, we will prove that a version of the W-Cycle algorithm converges. To do so, we need a series of technical lemmas.

Lemma (Error Transfer)

Let $m_1 = m_2 = m \in \mathbb{N}$ and $p \in \mathbb{N}$, with $p > 1$. Assume that (G3) *holds. Then, for $0 \leq \ell \leq L$, the error transfer matrix for the W-Cycle algorithm, E_ℓ, is self-adjoint with respect to the A_ℓ-inner product. Furthermore, if there is a number $\gamma > 0$ such that, for all $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$,*

$$\left|(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\right| \leq \gamma \, (\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} \,,$$

then, for all $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$,

$$\left|(\mathsf{E}_\ell^p \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\right| \leq \gamma^p \, (\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} \,.$$

Finally, if p is even, then E_ℓ is, in addition, positive semi-definite with respect to the A_ℓ-inner product, that is,

$$0 \leq (\mathsf{E}_\ell^p \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} \,,$$

for all $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$.

Proof.

We leave it as an exercise for the reader to confirm that that E_ℓ is self-adjoint with respect to the A_ℓ-inner product.

If E_ℓ is self-adjoint with respect to the A_ℓ inner product, there is an orthonormal basis of eigenvectors (with respect to the A_ℓ-inner product) and the eigenvalues are all real. Now, expand $\boldsymbol{v}_\ell$ in this basis and estimate.

The details of the proof are left as an exercise. $\square$

Because, $\hat{\Pi}_\ell$, introduced earlier in this section, is a projection, we have the following ultimately useful result, which will be used in the convergence proof for the W-Cycle.

Lemma (Projection Estimate)

Assume that (G3) *holds. Then, for any* $1 \leq \ell \leq L$ *and all* $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$, *we have*

$$-\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} \leq (\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\,. \tag{4}$$

Proof.

Using a previous corollary,

$$
\begin{aligned}
0 \le 2\left\|(\mathsf{I}_\ell - \hat{\Pi}_\ell)\boldsymbol{v}_\ell\right\|_{\mathsf{A}_\ell}^2 &= 2\left((\mathsf{I}_\ell - \hat{\Pi}_\ell)\boldsymbol{v}_\ell, (\mathsf{I}_\ell - \hat{\Pi}_\ell)\boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} \\
&= 2\left((\mathsf{I}_\ell - \hat{\Pi}_\ell)\boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} \\
&\overset{(1)}{=} \left((2\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} \\
&= \left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} + (\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}.
\end{aligned}
$$

□

Lemma (Some Projection Identities)

Assume that (G3) *holds. Then, for any* $1 \leq \ell \leq L$ *and all* $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$*, we have*

$$\left(\Pi_{\ell-1}\boldsymbol{v}_\ell, \Pi_{\ell-1}\boldsymbol{v}_\ell\right)_{\mathsf{A}_{\ell-1}} = r\left(\boldsymbol{v}_\ell, \tilde{\Pi}_\ell \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} = r\left(\tilde{\Pi}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell}. \tag{5}$$

Similarly, for all $1 \leq \ell \leq L$*, any vector* $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$*, and every matrix* $\mathsf{B}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}\times n_{\ell-1}}$*,*

$$\left(\mathsf{B}_{\ell-1}\Pi_{\ell-1}\boldsymbol{v}_\ell, \Pi_{\ell-1}\boldsymbol{v}_\ell\right)_{\mathsf{A}_{\ell-1}} = r\left(\mathsf{P}_{\ell-1}\mathsf{B}_{\ell-1}\Pi_{\ell-1}\boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell}. \tag{6}$$

Proof.

Suppose that $1 \leq \ell \leq L$ and recall that

$$\mathsf{A}_{\ell-1}\Pi_{\ell-1} = \mathsf{R}_{\ell-1}\mathsf{A}_\ell.$$

Proof (Cont.)

Let now $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$ be arbitrary. We compute

$$\begin{aligned}
(\Pi_{\ell-1}\boldsymbol{v}_\ell, \Pi_{\ell-1}\boldsymbol{v}_\ell)_{\mathsf{A}_{\ell-1}} &= (\mathsf{A}_{\ell-1}\Pi_{\ell-1}\boldsymbol{v}_\ell, \Pi_{\ell-1}\boldsymbol{v}_\ell)_{\ell-1} \\
&= (\mathsf{R}_{\ell-1}\mathsf{A}_\ell\boldsymbol{v}_\ell, \Pi_{\ell-1}\boldsymbol{v}_\ell)_{\ell-1} \\
&= \left(\mathsf{A}_\ell\boldsymbol{v}_\ell, \mathsf{R}_{\ell-1}^\top\Pi_{\ell-1}\boldsymbol{v}_\ell\right)_\ell \\
&= r\,(\mathsf{A}_\ell\boldsymbol{v}_\ell, \mathsf{P}_{\ell-1}\Pi_{\ell-1}\boldsymbol{v}_\ell)_\ell \\
&= r\left(\mathsf{A}_\ell\boldsymbol{v}_\ell, \tilde{\Pi}_\ell\boldsymbol{v}_\ell\right)_\ell \\
&= r\left(\boldsymbol{v}_\ell, \tilde{\Pi}_\ell\boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} .
\end{aligned}$$

The second identity is similarly derived. □

We now have available all of the necessary tools to prove convergence of the symmetric W-cycle algorithm.

Theorem (α–Weak Convergence of the W-Cycle)

Let $m_1 = m_2 = m \in \mathbb{N}$, and $p \in \mathbb{N}$ be even. Assume that (G3) *and* (A2) *hold, and that smoothing is performed using Richardson's smoothing. Then, for any* $0 < \gamma < 1$, *there is an $m \in \mathbb{N}$, such that, for all $1 \leq \ell \leq L$, the W-Cycle error matrix, E_ℓ, satisfies*

$$\left| (\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} \right| \leq \gamma \left(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_{\mathsf{A}_\ell},$$

for all $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$. In particular, it suffices to choose m that satisfies

$$m \geq \left(\frac{1 - \gamma^p}{\gamma - \gamma^p} \right)^{1/\alpha} C_{\mathrm{A2}}^2 C_{\mathrm{S1}}^2. \tag{7}$$

Proof.

Fix $1 \leq \ell \leq L$. Recall that

$$\mathsf{E}_\ell = (\mathsf{K}_\ell^*)^m \left(\mathsf{I}_\ell - \mathsf{P}_{\ell-1}\left(\mathsf{I}_{\ell-1} - \mathsf{E}_{\ell-1}^p\right)\Pi_{\ell-1}\right)\mathsf{K}_\ell^m \in \mathbb{R}^{n_\ell \times n_\ell}.$$

We first observe that E_ℓ is not necessarily positive semi-definite. It is, however, self-adjoint with respect to the A_ℓ-inner product, as it is guaranteed by a previous lemma. Set

$$\boldsymbol{w}_\ell = \mathsf{K}_\ell^m \boldsymbol{v}_\ell.$$

Using the definition of E_ℓ and (6), it follows that

$$(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} = \left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} + \frac{1}{r}\left(\mathsf{E}_{\ell-1}^p \Pi_{\ell-1}\boldsymbol{w}_\ell, \Pi_{\ell-1}\boldsymbol{w}_\ell\right)_{\mathsf{A}_{\ell-1}}.$$

Proof (Cont.)

We now begin the proof of convergence. We fix $\gamma \in (0,1)$, and proceed by induction on ℓ. The base case, $\ell = 0$, is trivial. The case $\ell = 1$ is true by virtue of the related Two-Grid convergence result. For the induction hypothesis, let us assume that

$$\left| \left(\mathsf{E}_{\ell-1} \boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1} \right)_{\mathsf{A}_{\ell-1}} \right| \leq \gamma \left(\boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1} \right)_{\mathsf{A}_{\ell-1}},$$

for all $\boldsymbol{v}_{\ell-1} \in \mathbb{R}^{n_{\ell-1}}$. Using properties of the error matrix, it follows that

$$0 \leq \left(\mathsf{E}_{\ell-1}^{p} \boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1} \right)_{\mathsf{A}_{\ell-1}} = \left| \left(\mathsf{E}_{\ell-1}^{p} \boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1} \right)_{\mathsf{A}_{\ell-1}} \right| \leq \gamma^{p} \left(\boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1} \right)_{\mathsf{A}_{\ell-1}}.$$

Proof (Cont.)

Using this last estimate and identity (5), it follows that

$$
\begin{aligned}
(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} &= \left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} + \frac{1}{r}\left(\mathsf{E}^p_{\ell-1}\Pi_{\ell-1}\boldsymbol{w}_\ell, \Pi_{\ell-1}\boldsymbol{w}_\ell\right)_{\mathsf{A}_{\ell-1}} \\
&\leq \left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} + \frac{\gamma^p}{r}\left(\Pi_{\ell-1}\boldsymbol{w}_\ell, \Pi_{\ell-1}\boldsymbol{w}_\ell\right)_{\mathsf{A}_{\ell-1}} \\
&\overset{(5)}{=} \left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} + \gamma^p\left(\tilde{\Pi}_\ell\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} \qquad (8) \\
&= (1-\gamma^p)\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} + \gamma^p\left(\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} \\
&\leq (1-\gamma^p)\left|\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell}\right| + \gamma^p\left(\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell}.
\end{aligned}
$$

Proof (Cont.)

On the other hand, since E_ℓ^p is self-adjoint positive semi-definite and using (4),

$$
\begin{aligned}
-\left(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell\right)_{\mathsf{A}_\ell} &= -\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} - \frac{1}{r}\left(\mathsf{E}_{\ell-1}^p \Pi_{\ell-1}\boldsymbol{w}_\ell, \Pi_{\ell-1}\boldsymbol{w}_\ell\right)_{\mathsf{A}_{\ell-1}} \\
&\leq -\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} \\
&= -(1-\gamma^p)\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} - \gamma^p\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} \\
&\overset{(4)}{\leq} -(1-\gamma^p)\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} + \gamma^p\left(\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell} \\
&\leq (1-\gamma^p)\left|\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell}\right| + \gamma^p\left(\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell}.
\end{aligned}
\tag{9}
$$

Proof (Cont.)

Combining estimates (8) and (9), we have

$$\left|(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\right| \leq (1-\gamma^p)\left|\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell}\right| + \gamma^p\,(\boldsymbol{w}_\ell, \boldsymbol{w}_\ell)_{\mathsf{A}_\ell}\,.$$

To conclude we use the error estimate for the Two-Grid method and the stability of Richardson's smoother to obtain, provided (7),

$$\begin{aligned}
\left|(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\right| &\leq (1-\gamma^p)\left(\frac{C_{\mathrm{A}2}C_{\mathrm{S}1}}{m^{1/2}}\right)^{2\alpha}(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} + \gamma^p\,(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell} \\
&= \left[(1-\gamma^p)\left(\frac{C_{\mathrm{A}2}C_{\mathrm{S}1}}{m^{1/2}}\right)^{2\alpha} + \gamma^p\right](\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\,, \\
&\leq \gamma\,(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\,.
\end{aligned}$$

The proof is complete. □

Corollary (Convergence of the W-Cycle)

With the same hypotheses as in the last theorem, it follows that, for every $0 < \gamma < 1$, *there is an* $m > 0$, *such that*

$$\left\|\mathsf{E}_\ell \boldsymbol{v}_\ell\right\|_{\mathsf{A}_\ell} \leq \gamma \left\|\boldsymbol{v}_\ell\right\|_{\mathsf{A}_\ell}, \tag{10}$$

for all $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$. *In other words, the W-Cycle method converges, provided* $m > 0$ *is sufficiently large.*

Proof.

Exercise. □

Room for Improvement

While the previous results already show convergence of the W-cycle, following Bramble and Zhang (2000), we can do better and prove that the W-cycle converges even when $m_1 = m_2 = m = 1$. Doing so requires some of auxiliary results. First, we state a special form of Young's inequality.

Lemma (Young's Inequality)

Let $p \in (1, \infty)$ *and define*

$$q := \frac{p}{p-1}.$$

Then, for every $a, b \in \mathbb{R}$*, any* $\tau > 0$*, and all* $\alpha \in (0,1)$*, it follows that*

$$|ab| \leq \tau^{\alpha p} \frac{|a|^p}{p} + \frac{1}{\tau^{\alpha q}} \frac{|b|^q}{q}.$$

In particular, upon choosing

$$p = \frac{1}{\alpha}, \qquad q = \frac{1}{1-\alpha},$$

we have, for any $\tau > 0$ *and any* $\alpha \in (0,1)$*,*

$$|ab| \leq \tau\alpha|a|^{1/\alpha} + \tau^{-\alpha/(1-\alpha)}(1-\alpha)|b|^{1/(1-\alpha)}. \tag{11}$$

Proof.

Exercise. □

Lemma (Technical Estimate)

Let $0 < \alpha < 1$, $K > 0$, *and set*

$$K_\alpha := (1-\alpha)^{1-\alpha}\alpha^\alpha K^\alpha, \tag{12}$$

Define, for $0 \le x \le 1$, $0 < \beta < 1$, *and* $m \in \mathbb{N}$,

$$\Gamma(\beta; x; m) := (1-\beta)\frac{K^\alpha}{m^\alpha}(1-x)^\alpha x^{1-\alpha} + \beta x. \tag{13}$$

Assume that $M > 0$ *is chosen so that*

$$M \ge \max\left\{\frac{1}{2}\left(4K_\alpha\right)^{1/\alpha}, m^\alpha\right\}. \tag{14}$$

Then

$$\Gamma(\delta^2; x; m) \le \delta,$$

where

$$\delta = \frac{M}{M + m^\alpha}, \tag{15}$$

Proof.

Using (11), it follows that

$$\Gamma(\beta; x; m) \leq L(\beta; x; m), \tag{16}$$

where

$$\begin{aligned} L(\beta; x; m) &:= (1-\beta)\left[\alpha K \tau (1-x) + (1-\alpha)(\tau m)^{-\alpha/(1-\alpha)} x\right] \\ &\quad + \beta x, \end{aligned} \tag{17}$$

for any $\tau > 0$. Observe that $x \mapsto L(\beta; x; m)$ is linear, with

$$L(\beta; 0; m) = (1-\beta)\alpha K \tau > 0,$$

and

$$L(\beta; 1; m) = \beta + (1-\beta)(1-\alpha)(\tau m)^{-\alpha/(1-\alpha)} > 0.$$

Thus, for any $\beta \in (0,1)$, any $x \in [0,1]$, and any $m \in \mathbb{N}$,

$$0 \leq \Gamma(\beta; x; m) \leq \max\left\{L(\beta; 0; m), L(\beta; 1; m)\right\}.$$

Proof (Cont.)

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

$$0 \leq \Gamma(\delta^2; x; m) \leq \max\left\{L(\delta^2; 0; m), L(\delta^2; 1; m)\right\}. \tag{18}$$

Thus, if we can show that

$$L(\delta^2; 0; m) \leq \delta \tag{19}$$

and

$$L(\delta^2; 1; m) \leq \delta \tag{20}$$

the proof will be complete.

Proof (Cont.)

Let us then show (20). We have

$$
\begin{aligned}
(1-\delta)^\alpha \left(\frac{1+\delta}{\delta}\right) &\overset{(15)}{=} \left(\frac{m^\alpha}{M+m^\alpha}\right)^\alpha \frac{2M+m^\alpha}{M} \\
&\leq \left(\frac{m^\alpha}{M+m^\alpha}\right)^\alpha \frac{2(M+m^\alpha)}{M} \\
&= m^{\alpha^2} \frac{2(M+m^\alpha)^{1-\alpha}}{M} \\
&\overset{(14)}{\leq} m^{\alpha^2} \frac{2(2M)^{1-\alpha}}{M} \\
&= m^{\alpha^2} \frac{4}{(2M)^\alpha} \\
&\overset{(14)}{\leq} \frac{m^{\alpha^2}}{K_\alpha} \\
&\leq \frac{m^\alpha}{K_\alpha}.
\end{aligned}
$$

Proof (Cont.)

Using (12) and some algebra, the last estimate is equivalent to

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

Now, we can find a number $\tau > 0$, such that,

$$\frac{(1-\alpha)^{1-\alpha}(1+\delta)^{1-\alpha}}{m^{\alpha}\delta^{1-\alpha}} \leq \tau^{\alpha} \leq \left(\frac{\delta}{\alpha(1-\delta^2)K}\right)^{\alpha}. \tag{21}$$

This will fix the value of τ in the definition of L.

Proof (Cont.)

Finally, the two inequalities in (21) are clearly equivalent to

$$\tau^{\alpha} \leq \left(\frac{\delta}{\alpha(1-\delta^2)K} \right)^{\alpha}$$

and

$$\frac{(1-\alpha)^{1-\alpha}(1+\delta)^{1-\alpha}}{m^{\alpha}\delta^{1-\alpha}} \leq \tau^{\alpha},$$

which are, respectively, equivalent to (19) and (20), after some algebra. The proof is complete. □

Lastly, we need an estimate that we established earlier for the proof of the Braess-Hackbusch Result.

Lemma (Richardson's Smoother)

Suppose that smoothing is done with Richardson's smoother, that is,

$$\mathsf{S}_\ell = \Lambda_\ell^{-1}\mathsf{I}_\ell,$$

where

$$\rho_\ell \le \Lambda_\ell \le C_{\mathrm{R}}\rho_\ell,$$

for some $C_{\mathrm{R}} \ge 1$ *that is independent of* ℓ*. Then, for any* $m \ge 1$, $\ell \ge 1$, *and all* $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$, *we have*

$$\left((\mathsf{I}_\ell - \mathsf{K}_\ell)\, \mathsf{K}_\ell^{2m}\boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_{\mathsf{A}_\ell} \le \frac{1}{2m}\left(\left(\mathsf{I}_\ell - \mathsf{K}_\ell^{2m}\right)\boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_{\mathsf{A}_\ell}. \tag{22}$$

Consequently,

$$\rho_\ell^{-1}\left\|\mathsf{A}_\ell\mathsf{K}_\ell^m\boldsymbol{v}_\ell\right\|_\ell^2 \le \frac{C_{\mathrm{R}}}{2m}\left(\|\boldsymbol{v}_\ell\|_{\mathsf{A}_\ell}^2 - \|\mathsf{K}_\ell^m\boldsymbol{v}_\ell\|_{\mathsf{A}_\ell}^2 \right). \tag{23}$$

Theorem (α–Weak Convergence of the W-Cycle)

Let $m_1 = m_2 = m \in \mathbb{N}$ and $p = 2$. Assume that (G3) and (A2) hold, and that the smoothing is performed using Richardson's smoother. Define K_α as in (12) with

$$K = \frac{C_{\mathrm{A2}}^2 C_{\mathrm{R}}}{2}.$$

Let $M > 0$ be chosen so that (14) holds. Then, for every $1 \leq \ell \leq L$, the W-Cycle error matrix, E_ℓ, satisfies

$$\left| \left(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_{\mathsf{A}_\ell} \right| \leq \frac{M}{M + m^\alpha} \left(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell \right)_{\mathsf{A}_\ell}, \tag{24}$$

for all $\boldsymbol{v}_\ell \in \mathbb{R}^{n_\ell}$.

Proof.

The proof is by induction. Suppose that

$$\left|(\mathsf{E}_{\ell-1}\boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1})_{\mathsf{A}_{\ell-1}}\right| \leq \gamma\,(\boldsymbol{v}_{\ell-1}, \boldsymbol{v}_{\ell-1})_{\mathsf{A}_{\ell-1}}\,, \qquad \gamma = \frac{M}{M+m^\alpha}.$$

In the proof of the previous theorem, we observed that, for $\boldsymbol{w}_\ell = \mathsf{K}_\ell^m \boldsymbol{v}_\ell$,

$$\left|(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\right| \leq (1-\gamma^2)\left|\left((\mathsf{I}_\ell - \tilde{\Pi}_\ell)\boldsymbol{w}_\ell, \boldsymbol{w}_\ell\right)_{\mathsf{A}_\ell}\right| + \gamma^2\,(\boldsymbol{w}_\ell, \boldsymbol{w}_\ell)_{\mathsf{A}_\ell}\,.$$

Proof (Cont.)

Using (A2) and an estimate used in the proof of the Braess-Hackbusch Result, it follows that

$$
\begin{aligned}
I_\ell &:= \left|(\mathsf{E}_\ell \boldsymbol{v}_\ell, \boldsymbol{v}_\ell)_{\mathsf{A}_\ell}\right| \\
&\overset{(A2)}{\leq} (1-\gamma^p)\frac{C_{\mathrm{A2}}^{2\alpha}}{\rho_\ell^\alpha} \|\mathsf{A}_\ell \boldsymbol{w}_\ell\|_\ell^{2\alpha} \|\boldsymbol{w}_\ell\|_{\mathsf{A}_\ell}^{2(1-\alpha)} + \gamma^p (\boldsymbol{w}_\ell, \boldsymbol{w}_\ell)_{\mathsf{A}_\ell} \\
&\overset{(23)}{\leq} (1-\gamma^2) C_{\mathrm{A2}}^{2\alpha} \left(\frac{C_{\mathrm{R}}}{2m}\left(\|\boldsymbol{v}_\ell\|_{\mathsf{A}_\ell}^2 - \|\boldsymbol{w}_\ell\|_{\mathsf{A}_\ell}^2\right)\right)^\alpha \|\boldsymbol{w}_\ell\|_{\mathsf{A}_\ell}^{2(1-\alpha)} + \gamma^2 \|\boldsymbol{w}_\ell\|_{\mathsf{A}_\ell}^2 \\
&= \Gamma\left(\gamma^2; x; m\right) \|\boldsymbol{v}_\ell\|_{\mathsf{A}_\ell}^2,
\end{aligned}
$$

where Γ is defined as in (13), with

$$
x := \frac{\|\boldsymbol{w}_\ell\|_{\mathsf{A}_\ell}^2}{\|\boldsymbol{v}_\ell\|_{\mathsf{A}_\ell}^2}.
$$

Proof (Cont.)

Observe that, owing to the stability of Richardson's smoother, it follows that

$$0 \leq x \leq 1.$$

Thus, with condition (14) imposed on M, we can apply Lemma 14 to conclude that

$$\Gamma\left(\gamma^2; x; m\right) \leq \gamma.$$

The proof is complete. □

Remark (Open Problem)

The avid reader may have noticed that all the convergence results presented in this section only pertain the W-cycle. To the best of our knowledge, and at the time of this writing, convergence of the standard V-cycle under condition (G3) *is an open question.*