

Randolph E. Bank · Harry Yserentant

# Multigrid Convergence: A Brief Trip Down Memory Lane

*Dedicated to Professor Wolfgang Hackbusch on the occasion of his 60<sup>th</sup> birthday*

Received: March 25, 2010 / Accepted: date

**Abstract** We make a brief algebraic survey of the highlights of the classical convergence theory for multigrid methods, in particular, the multigrid V-cycle.

**Keywords** Multigrid Convergence, V-Cycle

**Mathematics Subject Classification (2000)** 65M55, 65F10

## 1 Introduction

In this work we take a brief look at the classical convergence theory for multigrid methods developed in the late 1970's and 1980's. Much of this theory is due to the pioneering work of Wolfgang Hackbusch. Here we present the theory from a strictly algebraic point of view. After years of study and refinement, some aspects of the theory have been greatly simplified, but the central ideas in our analysis remain inspired by the original work. We attempt to be brief and concise in our presentation, and to avoid temptations to pursue extensions and generalizations. We direct our focus to the most essential aspects of the classical V-Cycle convergence theory, although some discussion of the W-Cycle is included, since historically convergence proofs for the W-Cycle predated those for the V-Cycle.

In Section 2, we present definitions and notation. In Section 3, we present our main assumption and prove some lemmas that connect our assumption to previous work, in particular the seminal work of Braess and Hackbusch [2]. More precisely, we show our assumption (2) is equivalent to the

---

Bank: The work of this author was supported by the National Science Foundation under contract DMS-xxxxxxx, and the Alexander von Humboldt Foundation through a Humboldt Research Award.

Yserentant: The work of this author was supported by the DFG-Research Center MATHEON.

---

Bank: Department of Mathematics, University of California, San Diego, La Jolla, California 92093-0112. Email: rbank@ucsd.edu.  
Yserentant: Institut für Mathematik, Technische Universität Berlin, StraÙ des 17.Juni 136, 10623 Berlin. Email: yserentant@math.tu-berlin.de.

well known weak approximation property, and dual norm estimates that often played a central role in earlier work.

In Section 4 we present the main convergence proof for the 2-level method, the W-Cycle and the V-Cycle. Our main results may be summarized as:

$$\frac{\kappa_2 - 1}{\kappa_2} = \|T\| \leq \|W\| \leq \|V\| \leq \frac{\kappa}{\kappa + 1},$$

where  $T$ ,  $W$ , and  $V$  are the error propagators for the two-level iteration, the W-Cycle, and the V-Cycle, respectively, and  $\kappa$  is the constant in our central assumption (2). The constant  $\kappa_2 \leq \kappa$  gives an exact characterization of two-level convergence. In exceptional circumstances  $\kappa_2 = \kappa$ .

Although we make few specific citations in our presentation, certainly many researchers made important contributions to multigrid convergence theory over many years. Indeed a complete list of references would be longer than our presentation. Thus we limit the references to the work of Hackbusch and Braess [2], two manuscripts that preceded this work but set the stage [1,4], and some important surveys and books [3,5–7] that give more complete coverage of multigrid theory, and contain many further references to the available literature.

## 2 Notation and Definitions

Let  $A$  be an  $N \times N$ , symmetric, positive definite matrix. We consider the solution of

$$Ax = b.$$

Typically  $A$  corresponds to the discretization of a self-adjoint elliptic partial differential equation by finite elements, finite differences or finite volumes.

Let  $N \equiv N_J > N_{J-1} > \dots > N_1$  denote the subspace dimensions of a  $J$  level method. Often the  $N_k$  correspond to different levels of refinement in the discretization process. Let  $\hat{R}_k$  be the  $N_k \times N_{k+1}$  matrix with rank  $N_k$ .  $\hat{R}_k$  is a so-called restriction, locally mapping the level  $k+1$  space to

the level  $k$  space. Let  $A \equiv A_J$  and define

$$A_k = \hat{R}_k A_{k+1} \hat{R}_k^t.$$

Note each  $A_k$  is symmetric, positive definite, and often corresponds to the usual discretization matrix for level  $k$ . Let the matrices  $R_k : \mathbb{R}^{N_J} \rightarrow \mathbb{R}^{N_k}$  be given by

$$R_k = \hat{R}_k \hat{R}_{k+1} \dots \hat{R}_J.$$

By convention  $\hat{R}_J = I_J$ . Then

$$A_k = R_k A R_k^t.$$

The energy inner product and corresponding norm on the finest level  $J$  are defined by

$$(x, y)_J \equiv (x, y) = x^t A_J y,$$

$$\|x\|_J^2 \equiv \|x\|^2 = x^t A_J x.$$

For  $\hat{x}, \hat{y} \in \mathbb{R}^{N_k}$ , we have

$$(\hat{x}, \hat{y})_k \equiv \hat{x}^t A_k \hat{y} \equiv (x, y)_J,$$

$$\|\hat{x}\|_k^2 \equiv \hat{x}^t A_k \hat{x} \equiv \|x\|_J^2,$$

where  $x = R_k^t \hat{x}$ , and  $y = R_k^t \hat{y}$ . These usually correspond to the natural inner product and norm associated with the underlying boundary value problem.

The coarse grid projection from the finest space  $N_J$  to  $N_k$  with respect to the energy inner product,  $P_k : \mathbb{R}^{N_J} \rightarrow \mathbb{R}^{N_k}$ , is given by

$$P_k = A_k^{-1} R_k A_J$$

Here we observe  $P_k R_k^t x = x$  for all  $x \in \mathbb{R}^{N_k}$ . Thus we have the well known identities

$$A_k P_k = R_k A,$$

$$P_k R_k^t = I_k.$$

The error propagator for the coarse grid correction is

$$C_k = I_J - R_k^t A_k^{-1} R_k A_J = I_J - R_k^t P_k.$$

Note that  $C_k : \mathbb{R}^{N_J} \rightarrow \mathbb{R}^{N_J}$  is a self adjoint projection matrix in the energy inner product and that, as a projection,  $C_k^2 = C_k$ . The projection  $C_k$  induces the  $A$ -orthogonal decomposition of  $\mathbb{R}^{N_J}$  as

$$\mathbb{R}^{N_J} = \mathcal{V}_k \oplus \mathcal{W}_k,$$

where  $\mathcal{V}_k$  denotes the nullspace of  $C_k$  and  $\mathcal{W}_k$  its  $A$ -orthogonal complement.  $\mathcal{V}_k$  in some sense corresponds to the coarse discretization space at level  $k$ . Finally, we have the identity

$$C_{k-1} R_k^t = R_k^t \hat{C}_{k-1}$$

$$\hat{C}_{k-1} = I_k - \hat{R}_{k-1}^t A_{k-1}^{-1} \hat{R}_{k-1} A_k$$

where  $\hat{C}_{k-1}$  is  $N_k \times N_k$ ; i.e., it is the ‘‘local’’ two-level coarse grid error propagator between levels  $k-1$  and  $k$ .

Let  $B_k$  be the  $N_k \times N_k$  matrix used to be used as a preconditioner for  $A_k$ . We assume  $B_k$  is symmetric and positive definite. We note the eigenvalues  $\mu$  of  $B_k^{-1} A_k$  are real and positive since they satisfy the generalized eigenvalue problem

$$A_k x = \mu B_k x.$$

We assume the  $B_k$  are scaled such that

$$\frac{x^t A_k x}{x^t B_k x} \leq 1. \quad (1)$$

To approximately solve  $A_k x = b$ , we take  $x_0$  as an initial guess, solve  $B_k e_0 = b - A_k x_0$ , and set  $x_f = x_0 + e_0$ . The ‘‘local’’  $N_k \times N_k$  error propagator for the smoother at level  $k$  is given by

$$\hat{S}_k = I_k - B_k^{-1} A_k.$$

We note that often more than one smoothing step is used; it is straightforward to incorporate this into the definition of  $B_k$ . If  $m > 1$  smoothing iterations are used with a preconditioner  $\hat{B}_k$  then our analysis remains valid for  $B_k$  defined implicitly by  $I_k - B_k^{-1} A_k = (I_k - \hat{B}_k^{-1} A_k)^m$ . If  $\hat{B}_k$  satisfies (1) then a simple calculation shows that  $B_k$  does as well.

We view  $\hat{S}_k$  as an  $N_J \times N_J$  matrix operating on vectors in  $\mathbb{R}^{N_J}$  by setting

$$S_k = I_J - R_k^t B_k^{-1} R_k A_J = I_J - R_k^t B_k^{-1} A_k P_k.$$

Note that  $S_k$  is self adjoint in the energy inner product. Similar to  $C_k$  we have the identity

$$S_k R_k^t = R_k^t \hat{S}_k.$$

We now consider multilevel iterations. The error propagator for the symmetric V-Cycle  $V_k$  is defined recursively as

$$V_2 = S_2^{1/2} C_1 S_2^{1/2}$$

$$V_k = S_k^{1/2} S_{k-1}^{1/2} V_{k-1} S_{k-1}^{1/2} S_k^{1/2} \quad 3 \leq k \leq J.$$

(Here and in future, the  $S_k^{1/2}$  is computed with respect to the energy inner product.) It is easy to express  $V_k$  directly as

$$V_k = S_k^{1/2} S_{k-1} \dots S_2 C_1 S_2 \dots S_{k-1} S_k^{1/2}.$$

The operators  $S_J^{1/2}$  formally appear so that smoothers can be applied just once on each level. Algorithmically, one starts with a smoothing step and ends with a smoothing step. That is, one applies the operator

$$\tilde{V}_k = S_k^{1/2} V_k S_k^{1/2} = S_k S_{k-1} \dots S_2 C_1 S_2 \dots S_{k-1} S_k.$$

whose norm can be estimated by  $\|\tilde{V}_k\| \leq \|V_k\|$ .

The error propagator for the symmetric W-Cycle  $W_k$  is defined recursively as

$$W_2 = S_2^{1/2} C_1 S_2^{1/2}$$

$$W_k = S_k^{1/2} S_{k-1}^{1/2} W_{k-1} S_{k-1} W_{k-1} S_{k-1}^{1/2} S_k^{1/2} \quad 3 \leq k \leq J.$$

Unlike the V-Cycle, a non-recursive definition for the W-Cycle is very complicated.

### 3 Central Assumption

Our proof for the convergence of the V-Cycle represents in its essence an algebraic version of the proof of Braess and Hackbusch. We begin in this section with the discussion of the basic estimate on which the proof of Braess and Hackbusch relies and state this estimate first in the form of an abstract assumption. Let  $x \in \mathbb{R}^{N_J}$  and define, for  $x \neq 0$ , the functional

$$\rho(x) = \frac{\|S_J^{1/2}x\|^2}{\|x\|^2}.$$

The central assumption is then as follows: for all  $x \in \mathbb{R}^{N_J}$ , there exists  $\kappa \geq 1$ , independent of  $J$  and  $N_J$ , such that

$$\|C_{J-1}x\| \leq \sqrt{\kappa(1-\rho(x))}\|x\|. \quad (2)$$

It relates the action of the smoother and of the coarse grid correction, that is basic for the fast convergence of multigrid methods. We see below how this assumption is related to several assumptions commonly made in multigrid analysis and with that indirectly to the regularity properties of the underlying continuous problem.

**Lemma 1** *The estimate*

$$\|C_{J-1}x\|^2 \leq \kappa \|(B_J^{-1}A_J)^{1/2}x\|^2 \quad (3)$$

*holds if and only if (2) holds*

**Proof** The proof trivially follows from the identity

$$\begin{aligned} \|(B_J^{-1}A_J)^{1/2}x\|^2 &= (x, B_J^{-1}A_Jx) \\ &= (x, (I_J - S_J)x) \\ &= (1 - \rho(x))\|x\|^2. \end{aligned}$$

□

In light of (3), the constant  $\kappa$  is easily seen to be the largest eigenvalue of the generalized eigenvalue problem

$$(C_{J-1}x, C_{J-1}\chi) = \mu (B_J^{-1}A_Jx, \chi). \quad (4)$$

Here we present two additional lemmas relating our analysis to previous work.

**Lemma 2** *Assume there exists a  $\kappa \geq 1$ , such that for any  $x \in \mathbb{R}^{N_J}$ , there exists  $\chi \in \mathbb{R}^{N_{J-1}}$  such that*

$$\|x - R_{J-1}^t\chi\| \leq \sqrt{\kappa} \|(B_J^{-1}A_J)^{1/2}x\|. \quad (5)$$

*Then (5) holds if and only if (3) holds.*

**Proof** Estimate (5) is called the *weak approximation property*. Assume (5). Then

$$\begin{aligned} \|C_{J-1}x\|^2 &= (C_{J-1}x, C_{J-1}x) \\ &= (C_{J-1}x, x) \\ &= (C_{J-1}x, x - R_{J-1}^t\chi) \\ &\leq \sqrt{\kappa} \|C_{J-1}x\| \|(B_J^{-1}A_J)^{1/2}x\|. \end{aligned}$$

On the other hand, assume (3) and choose  $\chi = P_{J-1}x$ . Then  $\|x - R_{J-1}^t\chi\| = \|C_{J-1}x\|$ . □

The weak approximation property can be used to validate our main assumption (2). Many analyses, like the adaptation of the proof of Braess and Hackbusch in [7], also make use a pair of (discrete) dual norms, and the induced intermediate scale of norms.

**Lemma 3** *Estimate (3) holds if and only if*

$$\max_{\|w\|=1, w \in \mathcal{W}_{J-1}} \frac{w^t B_J w}{w^t A_J w} \leq \kappa. \quad (6)$$

**Proof** Observe that  $\|(B_J A_J^{-1})^{1/2}w\|^2 = w^t B_J w$ , and  $\|(B_J^{-1}A_J)^{1/2}w\|^2 = w^t A_J B_J^{-1}A_J w$  form a pair of dual norms with respect to the energy inner product. Assume (3),  $w \in \mathcal{W}_{J-1}$ , and let  $z$  be defined by  $A_J z = B_J w$ ; then

$$\begin{aligned} \|(B_J^{-1}A_J)^{1/2}z\|^2 &= z^t A_J B_J^{-1}A_J z \\ &= w^t B_J w \\ &= \|(B_J A_J^{-1})^{1/2}w\|^2. \end{aligned}$$

Then using (3) and the fact that  $w \in \mathcal{W}_{J-1}$

$$\begin{aligned} w^t B_J w &= w^t A_J z \\ &= w^t A (z - (I_{J-1} - C_{J-1})z) \\ &\leq \|w\| \|C_{J-1}z\| \\ &\leq \sqrt{\kappa} \|w\| \|(B_J^{-1}A_J)^{1/2}z\| \\ &= \sqrt{\kappa} \|w\| \|(B_J A_J^{-1})^{1/2}w\|. \end{aligned}$$

This argument is a discrete form of the well known Nitsche trick. Conversely, assume (6). Then

$$\begin{aligned} \|C_{J-1}w\|^2 &= (C_{J-1}w, w) \\ &\leq \|(B_J A_J^{-1})^{1/2}C_{J-1}w\| \|(B_J^{-1}A_J)^{1/2}w\| \\ &\leq \sqrt{\kappa} \|C_{J-1}w\| \|(B_J^{-1}A_J)^{1/2}w\|. \end{aligned}$$

□

Thus we see that Lemma 3 provides another equivalent alternative for validating assumption (2).

The nature of the initially stated basic assumption becomes particularly apparent from its equivalence to condition 6. In the finite element context, for a second-order Laplace-type boundary value problem, the norm induced by the smoother is in many cases equivalent to a correspondingly weighted  $L_2$ -norm. Condition 6 then essentially means that

$$\|\mathcal{P}_k u - \mathcal{P}_{k-1} u\|_{L_2} \leq c 2^{-k} \|\mathcal{P}_k u - \mathcal{P}_{k-1} u\|_{H^1}$$

holds for all functions  $u$  in the continuous solution space  $H^1$  and all levels  $k$ , where  $\mathcal{P}_k$  here denotes the projection that maps the continuous solution to its finite element approximation of level  $k$ . It is not very difficult to deduce from this condition the estimate

$$\|u - \mathcal{P}_k u\|_{L_2} \leq \frac{1}{3} c 2^{-k} \|u - \mathcal{P}_k u\|_{H^1}$$

for the functions  $u \in H^1$ . That means, that one gains in the  $L_2$ -norm one order of convergence compared to the  $H^1$ -norm. This property is usually shown with help of the Nitsche trick that is based on  $H^2$ -regularity and more or less equivalent to it. In this scenario, the initially stated condition means that we study a problem with full elliptic regularity. This is the main drawback of the Braess and Hackbusch approach that was later overcome by other techniques (see [3, 6, 7]) that have their own disadvantages.

#### 4 Convergence Rate Estimates

We begin by estimating the rate of convergence of the two-level scheme, given by  $\|T_J\|$  where  $T_J = S_J^{1/2}C_{J-1}S_J^{1/2}$ . First note that since  $C_{J-1} = C_{J-1}^2$ ,

$$\begin{aligned} \|T_J\| &= \|S_J^{1/2}C_{J-1}S_J^{1/2}\| \\ &= \|C_{J-1}S_J^{1/2}\|^2 \\ &= \|S_J^{1/2}C_{J-1}\|^2 \\ &= \|C_{J-1}S_J C_{J-1}\| \\ &= \|C_{J-1}(I_J - B_J^{-1}A_J)C_{J-1}\|. \end{aligned}$$

Thus we can frame our analysis in terms of the generalized eigenvalue problem

$$(B_J^{-1}A_J C_{J-1}x, C_{J-1}\chi) = \mu(C_{J-1}x, C_{J-1}\chi)$$

on  $\mathcal{W}_{J-1}$ . Since  $B_J$  is positive definite, the eigenvalues  $\mu$  of  $B_J^{-1}A_J$  on  $\mathcal{W}_{J-1}$  lie in the positive interval  $0 < \lambda \leq \mu \leq 1$ . It is important to observe that this generalized eigenvalue problem is similar to, but more restrictive than, the eigenvalue problem for  $\kappa$  given in (4). (Replace  $x$  and  $\chi$  in (4) with  $C_{J-1}x$  and  $C_{J-1}\chi$  and use  $C_{J-1}^2 = C_{J-1}$ .) Thus  $\kappa \geq 1/\lambda \equiv \kappa_2$ ,  $1 - \lambda = (\kappa_2 - 1)/\kappa_2$ , and we have proved:

#### Theorem 1

$$\frac{\kappa_2 - 1}{\kappa_2} = \|T_J\| \leq \frac{\kappa - 1}{\kappa},$$

where  $\kappa_2 \leq \kappa$  and  $\kappa$  is given in (2).

We remark that our proof shows that if smoothing preserves the invariant subspace  $\mathcal{V}_{J-1}$ , then  $\kappa_2 = \kappa$ . It is this small ‘‘gap’’ between the two generalized eigenvalue problems that precludes this approach from showing in general that two-level convergence implies V-Cycle convergence.

We also remark that the above proof shows the best rate of convergence for the two-level iteration is found by minimizing over all subspaces  $\mathcal{V}_{J-1}$ . Clearly this is achieved when  $\mathcal{V}_{J-1}$  is the span of eigenvectors associated with the  $N_{J-1}$  smallest eigenvalues of  $B_J^{-1}A_J$ . This subspace is not necessarily unique, and as a practical matter, might be difficult to compute.

In the following theorem, we compare the V-Cycle and W-Cycle with the two-level iteration of Theorem 1.

#### Theorem 2 For $J \geq 2$ ,

$$\|T_J\| \leq \|W_J\| \leq \|V_J\|. \quad (7)$$

**Proof** We begin with the left-hand inequality in (7). Let  $\tilde{W}_{J-1} = S_{J-1}^{1/2}W_{J-1}S_{J-1}^{1/2}$ . Then

$$\begin{aligned} \|W_J\| &= \|S_J^{1/2}\tilde{W}_{J-1}S_J^{1/2}\| \\ &= \|\tilde{W}_{J-1}S_J^{1/2}\|^2 \\ &= \|(C_{J-1} + \tilde{W}_{J-1} - C_{J-1})S_J^{1/2}\|^2 \\ &= \|(C_{J-1} + \Delta)S_J^{1/2}\|^2. \end{aligned}$$

Since  $\Delta$  represents the error in the coarse grid correction  $C_{J-1}\Delta = 0$ . Thus

$$\begin{aligned} \|W_J\| &= \max_{\|x\|=1} (\{C_{J-1} + \Delta\}S_J^{1/2}x, \{C_{J-1} + \Delta\}S_J^{1/2}x) \\ &= \max_{\|x\|=1} \|C_{J-1}S_J^{1/2}x\|^2 + \|\Delta S_J^{1/2}x\|^2 \\ &\geq \|C_{J-1}S_J^{1/2}\|^2 \\ &= \|T_J\|. \end{aligned}$$

To prove the right-hand inequality in (7), we begin by considering the *non-symmetric* V-Cycle. Let

$$\begin{aligned} U_2 &= S_2^{1/2}C_1 \\ U_k &= S_k^{1/2}S_{k-1}^{1/2}U_{k-1} \quad 3 \leq k \leq J. \end{aligned}$$

Notice that  $V_k = U_k U_k^t$  and  $\|V_k\| = \|U_k\|^2$ . We will prove by induction that

$$W_k = U_k Z_k U_k^t, \quad (8)$$

$$\|Z_k\| \leq 1. \quad (9)$$

Since  $W_2 = U_2 U_2^t$ ,  $Z_2 = I_2$ . Now assume (8)-(9) hold for  $k$  to show for  $k+1$ . Now

$$W_{k+1} = S_{k+1}^{1/2}\tilde{W}_k S_{k+1}^{1/2} = U_{k+1}Z_{k+1}U_{k+1}^t,$$

where

$$Z_{k+1} = Z_k U_k^t S_k U_k Z_k.$$

But

$$\|Z_{k+1}\| \leq \|Z_k\|^2 \|U_k\|^2 \|S_k\| \leq 1,$$

and the induction is closed.  $\square$

We begin our analysis of the V-Cycle with two technical lemmas.

#### Lemma 4 Let $\kappa \geq 1$ , and let

$$R(\gamma) = \max_{0 \leq \rho \leq 1} ((1 - \gamma) \min\{1, \kappa(1 - \rho)\} + \gamma)\rho.$$

Then

$$R(\gamma) \leq \frac{\kappa}{\kappa + 1}. \quad (10)$$

**Proof** If  $0 \leq \rho \leq (\kappa - 1)/\kappa$ , then  $\min\{1, \kappa(1 - \rho)\} = 1$  and (10) is immediate. If  $\rho \geq (\kappa - 1)/\kappa$ , then  $\min\{1, \kappa(1 - \rho)\} = \kappa(1 - \rho)$ , and

$$R(\gamma) = \max_{(\kappa-1)/\kappa \leq \rho \leq 1} ((1 - \gamma)\kappa(1 - \rho) + \gamma)\rho.$$

Now  $R(\gamma)$  is an increasing function of  $\gamma$  for any  $(\kappa - 1)/\kappa \leq \rho \leq 1$ . Thus

$$\begin{aligned} R(\gamma) &\leq R\left(\frac{\kappa}{\kappa+1}\right) \\ &= \frac{\kappa}{\kappa+1} \max_{(\kappa-1)/\kappa \leq \rho \leq 1} (2 - \rho)\rho \\ &= \frac{\kappa}{\kappa+1}. \end{aligned}$$

□

The next lemma analyzes the error between an approximate coarse grid correction, given by  $\tilde{V}_{J-1} = S_{J-1}^{1/2} V_{J-1} S_{J-1}^{1/2}$ , and the exact coarse grid correction  $C_{J-1}$ .

**Lemma 5** Suppose that, for  $v \in \mathcal{V}_{J-1}$ ,

$$\|\tilde{V}_{J-1}v\| \leq \gamma\|v\|.$$

Then

$$\|\tilde{V}_{J-1} - (1 - \gamma)C_{J-1}\| \leq \gamma.$$

**Proof** Let  $x = v + w$ ,  $v \in \mathcal{V}_{J-1}$  and  $w \in \mathcal{W}_{J-1}$ . Then

$$\begin{aligned} \tilde{V}_{J-1}x &= \tilde{V}_{J-1}v + w, \\ C_{J-1}x &= w. \end{aligned}$$

Thus

$$\begin{aligned} \|(\tilde{V}_{J-1} - (1 - \gamma)C_{J-1})x\|^2 &= \|(\tilde{V}_{J-1}v + \gamma w)\|^2 \\ &\leq \gamma^2(\|v\|^2 + \|w\|^2) \\ &= \gamma^2\|x\|^2. \end{aligned}$$

□

We now turn to convergence of the V-Cycle. The classic approach is essentially an induction proof, which states that if the  $J - 1$  level V-Cycle converges at a given rate, then so does the  $J$  level V-Cycle. The base case for the induction, the two-level estimate, is already given in Theorem 1.

**Theorem 3** For the general case of a V-Cycle, assume that

$$\|\tilde{V}_{J-1}v\| \leq \gamma\|v\| \leq \frac{\kappa}{\kappa+1}\|v\| \quad (11)$$

for  $v \in \mathcal{V}_{J-1}$ . Then

$$\|\tilde{V}_J\| \leq \|V_J\| = \|S_J^{1/2}\tilde{V}_{J-1}S_J^{1/2}\| \leq \frac{\kappa}{\kappa+1}. \quad (12)$$

**Proof** Let  $e \in \mathbb{R}_J^N$  and  $0 \leq \gamma \leq \kappa/(\kappa + 1)$  be given. Then

$$\begin{aligned} (e, S_J^{1/2}\tilde{V}_{J-1}S_J^{1/2}e) &= (S_J^{1/2}e, \tilde{V}_{J-1}S_J^{1/2}e) \\ &= \{1 - \gamma\}(C_{J-1}S_J^{1/2}e, C_{J-1}S_J^{1/2}e) \\ &\quad + (S_J^{1/2}e, \{V_{J-1} - (1 - \gamma)C_{J-1}\}S_J^{1/2}e). \end{aligned} \quad (13)$$

For the first term on the right hand side of (13), we have

$$(C_{J-1}S_J^{1/2}e, C_{J-1}S_J^{1/2}e) \leq \min\{1, \kappa(1 - \rho(e))\}\|S_J^{1/2}e\|^2,$$

due to assumption (2), the trivial estimate  $\|C_{J-1}\| \leq 1$ , and the estimate  $\|S_J^{1/2}e\|^2 = (e, S_J e) \leq \|e\|\|S_J e\|$ , from which follows  $1 - \rho(S_J^{1/2}e) \leq 1 - \rho(e)$ . For the second term, we use Lemma 5 to see that

$$(S_J^{1/2}e, \{V_{J-1} - (1 - \gamma)C_{J-1}\}S_J^{1/2}e) \leq \gamma\|S_J^{1/2}e\|^2.$$

Finally,  $\|S_J^{1/2}e\|^2 \leq \rho(e)\|e\|^2$ . Combining these estimates,

$$(e, S_J^{1/2}V_{J-1}S_J^{1/2}e) \leq R(\gamma)\|e\|^2,$$

and second inequality in (12) follows from Lemma 4. The first inequality in (12), necessary to validate the hypothesis (11) for the next level, and thus close the induction, follows from

$$\|\tilde{V}_J\| = \|S_J^{1/2}V_J S_J^{1/2}\| \leq \|S_J^{1/2}\|^2\|V_J\| \leq \|V_J\|.$$

□

To summarize, in this section we have proved the chain of estimates

$$\frac{\kappa_2 - 1}{\kappa_2} = \|T_J\| \leq \|W_J\| \leq \|V_J\| \leq \frac{\kappa}{\kappa + 1}.$$

The possibility that  $\kappa_2 = \kappa$  in some exceptional but unlikely circumstances indicates that this is probably the best estimate possible through this classical approach.

## References

1. R. E. BANK AND T. DUPONT, *An optimal order process for solving elliptic finite element equations*, Math. Comp., 36 (1981), pp. 35–51.
2. D. BRAESS AND W. HACKBUSCH, *A new convergence proof for the multigrid method including the V cycle*, SIAM J. Numer. Anal., 20 (1983), pp. 967–975.
3. J. H. BRAMBLE AND X. ZHANG, *The analysis of multigrid methods*, in Handbook of numerical analysis, Vol. VII, Handb. Numer. Anal., VII, North-Holland, Amsterdam, 2000, pp. 173–415.
4. W. HACKBUSCH, *Convergence of multi-grid iterations applied to difference equations*, Math. Comp., 34 (1980), pp. 425–440.
5. W. HACKBUSCH, *Multigrid methods and applications*, vol. 4 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1985.
6. J. XU, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.
7. H. YSERENTANT, *Old and new convergence proofs for multigrid methods*, in Acta numerica, 1993, Acta Numer., Cambridge Univ. Press, Cambridge, 1993, pp. 285–326.