REMARKS ON MULTIGRID CONVERGENCE THEOREMS

by

Seymour V. Parter

Computer Sciences Technical Report #634

July 1986

REMARKS ON MULTIGRID CONVERGENCE THEOREMS[1]

by

Seymour V. Parter[2]

Technical Report #634

[1] Sponsored by the Air Force Office of Scientific Research under Contracts No. AFOSR-82-0275 and 86-0163.

[2] Computer Sciences Department, University of Wisconsin Madison, WI 53706.

REMARKS ON MULTIGRID CONVERGENCE THEOREMS

by

Seymour V. Parter

ABSTRACT

Multigrid has become an important iterative method for the solution of discrete elliptic equations. However, there is much to be done in the theory of convergence proofs. At the the present time there are two general two-level methods for general convergence proofs: an algebraic approach and a duality approach. While these theories do not give sharp estimates they provide good, general, rigorous convergence theorems. In this note we study the relationship between these theories. While the approach and thought-process leading to these theories are different, the results are essentially the same. Indeed, the basic estimated required by these theories are the same.

1. Introduction

Multigrid has become an important iterative method for the solution of discrete elliptic equations. However, there is much to be done in the theory of convergence proofs. The early proofs (e.g. [1], [13], [3], [6]) are limited to the W-cycle and yield statements of the form "If the number of smoothing steps is sufficiently large, then the method is convergent". In general, there is no estimate of (nor a recipe to estimate) "sufficiently large" nor has there been an estimate of the "rate of convergence". More recently, there have been many papers [2], [3], [5], [8], [9] [10], [11], [12], [14], which have given convergence proofs for the V-cycle and "any number of smoothings". Most of these results take the form "There is a constant C, depending on the regularity properties of the elliptic problem and the approximation properties of the finite-element subspaces, such that

$$\|\varepsilon^1\|_A \leq \frac{C}{C+k}\|\varepsilon^0\|_A \,, \tag{1}$$

where k is the number of smoothing steps, ε^0 is the error before the multigrid cycle, ε^1 is the error after that cycle and $\|\ \|_A$ denotes the energy norm". In [4], [16] the authors have studied very specific multigrid schemes and actually given numerical estimates of the rates of convergence for both V-cycle schemes and W-cycle schemes.

It is interesting and worthwhile to study the methodology of these convergence proofs. They all depend on a "two-level" analysis. That is, one obtains certain estimates of the process involving only two successive grids (or subspaces, $S_{j-1} \subset S_j$). Assuming that these estimates hold uniformly on all such pairs of successive grids, one then obtains a convergence rate for the entire multigrid process. These two-level analyses are of two kinds. One set is "algebraic" and seems to deal with certain spectral radii (e.g. [9], [10], [11], [12], [7]). The others use a "duality" argument (e.g. [2], [4], [5], [16], [17]). The estimates essential to these duality arguments have come from eigenfunction expansions and the constant C described above or from energy estimates.

Experimental results indicate that these estimates are not sharp - For example, in [7] the authors study a simple one dimensional example. The appropriate two-level constants were computed and the resulting estimate compared to experimental results. Similarly, the experimental results of [15] can be compared to the numerical estimates of [4]. In both cases the experiments indicate that the method is much better than the estimates obtained via these two-level theories. Nevertheless, the two-level theories are very useful and at this time provide us with an excellent approach to convergence theorems and bounds for the rates of convergence. In fact, we have no better general, rigorous, approach.

In this note we study the relationship between these theories. As we will show in section 2, while the approach and the thought-process leading to these estimates may be different, the results are essentially the same. Indeed, the basic required estimates are also the same.

I am indebted to Naomi Decker and David Kamowitz for their many useful discussions on this work.

2. Convergence Theories

We consider a finite dimensional linear vector space S_M with inner product $\langle\ ,\ \rangle$. Consider the problem

$$A_M U^{(M)} = f^{(M)} , \tag{2.1}$$

where A_M is a symmetric positive definite operator.

Consider a sequence of finite-dimensional spaces

$$\{S_j, j = 0, 1, \ldots, M\} \tag{2.2a}$$

with

$$\dim\ S_{j-1} < \dim\ S_j, \quad j = 1, 2, \ldots, M . \tag{2.2b}$$

Consider linear operators I_{j-1}^{j}, I_j^{j-1} which enable us to communicate between these spaces where

$$I_j^{j-1} : S_j \to S_{j-1} \quad \text{(projection)} , \tag{2.3a}$$

$$I_{j-1}^{j} : S_{j-1} \to S_j \quad \text{(interpolation)} . \tag{2.3b}$$

In this note we require that

$$I_j^{j-1} = (I_{j-1}^{j})^* . \tag{2.3c}$$

For each space S_j we define

$$\hat{A}_j = I_{j+1}^{j} \hat{A}_{j+1} I_j^{j+1}, \quad j = 0, 1, \ldots, M-1 \tag{2.4}$$

with $\hat{A}_M := A_M$. Finally, we have "smoothing" operators $G_j(u, f)$, $E_j(u, f)$. These are affine operators of the form

$$G_j(v, f) = \bar{G}_j v + \bar{K}_j f, \quad E_j(v, f) = \bar{E}_j v + \bar{F}_j f \tag{2.5}$$

where $\bar{G}_j, \bar{E}_j, \bar{K}_j$, and $\bar{F}_j$ are linear operators which satisfy the "consistency conditions"

$$v = G_j(v, f) \Longleftrightarrow \hat{A}_j v = f, \quad v = E_j(v, f) \Longleftrightarrow \hat{A}_j v = f , \tag{2.6a}$$

and

$$\|\bar{G}_j\|_A \le 1, \quad \|\bar{E}_j\|_A \le 1 \tag{2.6b}$$

where $\| \, \|_A$ denotes the "A norm" arising from the "A inner product",

$$\langle u, v \rangle_A = \langle \hat{A}_j u, v \rangle, \quad u, v \in S_j \ . \tag{2.7a}$$

With each $\bar{G}_j, \bar{E}_j$ we also define the A-adjoint operators $\bar{G}_j^*$, $\bar{E}_j^*$ which satisfy

$$\langle \bar{G}_j u, v \rangle_A = \langle u, \bar{G}_j^* v \rangle_A \ ,$$

$$\langle \bar{E}_j u, v \rangle_A = \langle u, \bar{E}_j^* v \rangle_A \ .$$

That is

$$\bar{G}_j^T \hat{A}_j = \hat{A}_j \bar{G}_j^*, \quad \bar{E}_j^T \hat{A}_j = \hat{A}_j \bar{E}_j^* \ . \tag{2.7b}$$

We are now able to define the multigrid iterative schemes, $MG(j, u, f)$, for the solution of (2.1). These schemes are defined recursively as follows. Let μ and m be positive integers. Assume u^j is known.

If $j = 0$ then

$$MG(0, u^0, f^0) = U^0 \ , \tag{2.8a}$$

where U^0 is the solution of

$$\hat{A}_0 U^0 = f^0 \ . \tag{2.8b}$$

If $1 \le j \le M$ perform the following five steps:

(i) Do m times:

$$u^j \leftarrow G_j(u^j, f^j) \ .$$

(ii) Set $r_j = f^j - \hat{A}_j u^j, \quad f^{j-1} = I_j^{j-1} r_j, \quad u^{j-1} = 0 \ .$

(iii) Do μ times:

$$u^{j-1} \leftarrow MG(j-1, u^{j-1}, f^{j-1}) \ .$$

(iv) Do $\qquad u^j \leftarrow u^j + I_{j-1}^j u^{j-1} \ .$

(v) Do m times

$$u^j \leftarrow E_j(u^j, f^j) \ .$$

Notes: When $\mu = 1$ we call these multigrid schemes - V-cycle schemes. When $\mu \geq 2$ we call these multigrid schemes - W-cycle schemes.

The cases $\bar{G}_j = I$ and $\bar{K}_j = 0$, $\bar{E}_j = I$ and $\bar{F}_j = 0$ fail to satisfy consistency condition (2.6a). Nevertheless, we may consider them as (extreme) special cases of "smoothing" operators. When $\bar{G}_j = I$ and $\bar{K}_j = 0$ the multigrid scheme $MG(j, u, f)$ becomes the coarse-to-fine cycle $M/_j(u, f)$ [see [12], [14]]. When $\bar{E}_j = I$ and $\bar{F}_j = 0$ the multigrid scheme becomes the fine-to-coarse cycle $M\backslash_j(u, f)$.

Remark 1: When the spaces $\{S_j\}$ are "nested", i.e., $S_{j-1} \subset S_j$ and one chooses I^j_{j-1} as the natural injection map (the identity restricted to S_{j-1}) we are dealing with a "finite element" multigrid scheme. As long as one chooses the coarse grid operators $\hat{A}_j$ by the recipe (2.4) we will say we have a "Galerkin" multigrid scheme. Such schemes are essentially finite element schemes. All general proofs for finite-element schemes apply to Galerkin schemes. The current convergence theory for V-cycles and even much of the convergence theory for W-cycles is for such Galerkin schemes. However, in practice one may choose another recipe for the coarse grid operator $\hat{A}_j$, $j < M$.

Let

$$R_j := \text{Range } \ I^j_{j-1} \ , \tag{2.9a}$$

$$N_j := \text{Nullspace } \ I^{j-1}_j \hat{A}_j \ . \tag{2.9b}$$

Let

$$T_j := \hat{A}\text{-orthogonal projection onto } \ N_j \ , \tag{2.10a}$$

$$\$_j := \hat{A}\text{-orthogonal projection onto } \ R_j \ . \tag{2.10b}$$

We now collect some basic results of [9], [12], [7].

Theorem 2.1: Consider the V-cycle multigrid algorithms. Let ε_j be the error before the start of the j^{th} multigrid $(MG(j, u^j, f^j))$, and let $\bar{\varepsilon}_j$ be the error after one iteration.

Suppose there are fixed numbers α_k, $\bar{\alpha}_k$, $0 < \alpha_k$, $\bar{\alpha}_k \leq 1$ which satisfy

$$\|\bar{E}^k_j u\|^2_A \leq \bar{\alpha}_k \|T_j u\|^2_A + \|\$_j u\|^2_A \qquad j = 1, 2, \ldots, M \ , \tag{2.11a}$$

$$\|T_j \bar{G}^k_j u\|^2_A + \alpha_k \|\$_j \bar{G}^k_j u\|^2_A \leq \alpha_k \|u\|^2_A \qquad j = 1, 2, \ldots, M \ . \tag{2.11b}$$

Then

$$\|\bar{\varepsilon}_j\|_A^2 \leq \alpha_k \bar{\alpha}_k \|\varepsilon_j\|_A^2 . \tag{2.12}$$

In the special cases noted above, we have

$$\bar{G}_j = I \quad \Rightarrow \alpha = 1 , \tag{2.12a}$$

$$\bar{E}_j = I \quad \Rightarrow \bar{\alpha} = 1 . \tag{2.12b}$$

Proof: The proof for the case $\bar{E}_j = I$, i.e., $M\backslash_j(u,f)$, first appeared in [12]. The case $\bar{G}_j = I$, i.e., $M/_j(u,f)$ appeared in [7]. The proof for the general case is given in [9]. In fact, [9] provides exact formulae for the optimal α_k and $\bar{\alpha}_k$. These are

$$\alpha_k = \rho[(I - (\bar{G}_j^*)^k \$_j (\bar{G}_j)^k)^{-1} (\bar{G}_j^*)^k T_j (\bar{G}_j)^k] , \tag{2.13a}$$

$$\bar{\alpha}_k = \rho[(I - (\bar{E}_j)^k \$_j (\bar{E}_j^*)^k)^{-1} (\bar{E}_j)^k T_j (\bar{E}_j^*)^k] , \tag{2.13b}$$

where ρ denotes the spectral radius of an operator. These formulae say

$$\bar{\alpha}_k(\bar{E}_j) = \alpha_k(\bar{E}_j^*) . \tag{2.14}$$

Theorem 2.2: Let $\bar{G}_j = \bar{E}_j^*$. Then $\alpha_k = \bar{\alpha}_k$. Let $\alpha_1 < 1$ be any number satisfying (2.11a) with k = 1. Let $c = \alpha_1/(1-\alpha_1)$, so that

$$\alpha_1 = \frac{c}{c+1} . \tag{2.15a}$$

Then (2.11a) holds for $k \geq 1$ with α_k given by

$$\alpha_k = \frac{c}{c+k} . \tag{2.15b}$$

Proof: See [12] and [9].

Theorem 2.3: Consider the fine-to-coarse W-cycle multigrid algorithm, $M\backslash_j(u^i, f^j)$ with $\mu \geq 2$. Let $0 < \eta < \dfrac{1}{2}$ and,

$$\|T_j \bar{G}_j^k u\|_A^2 \leq \eta \|u\|_A^2 .$$

Consider the polynomial

$$f(x) = (1-\eta)x^{\mu} - x + \eta \,. \tag{2.16}$$

This polynomial has a root $x = 1$ and another positive root $\bar{x}$, $0 < \bar{x} < 1$. If, as before, ε_j is the error before the start of the j^{th} multigrid and $\bar{\varepsilon}_j$ the error after that iteration then

$$\|\bar{\varepsilon}_j\|_A^2 \le \bar{x}\|\varepsilon_j\|_A^2 \,. \tag{2.17}$$

In particular, when $\mu = 2$

$$\bar{x} = \frac{\eta}{1-\eta} \,. \tag{2.18}$$

Proof: See Theorem 1 of [10].

There is another approach to this problem. This approach, based on a duality argument, has been used by Bank and C. Douglas [2], Braess [4] and Verfürth [16]. The following is our abstraction of their arguments.

With each $u \in S_j$ we associate real functions $g(u)$, $e(u)$, $t(u)$, $t^*(u)$ which satisfy

$$0 \le g(u),\ e(u),\ t(u),\ t^*(u) \le 1 \tag{2.19}$$

and

$$\|\bar{G}_j u\|_A^2 \le g(u)\|u\|_A^2 \,, \quad \|\bar{E}_j^* u\|_A^2 \le e(u)\|u\|_A^2 \,, \tag{2.20a}$$

$$\|T_j\bar{G}_j u\|_A^2 \le t(u)g(u)\|u\|_A^2 \,, \quad \|T_j\bar{E}_j^* u\|_A^2 \le t^*(u)e(u)\|u\|_A^2 \,. \tag{2.20b}$$

Remark: In many cases we may associate a real variable $\sigma \in [0,1]$ with $u \in S_j$ so that

$$\sigma = \sigma(u) \,, \tag{2.21a}$$

and we write

$$g(u) = g_0(\sigma), \ \ e(u) = e_0(\sigma) \,, \tag{2.21b}$$

$$t(u) = t_0(\sigma), \ \ t^*(u) = t_0^*(\sigma) \,. \tag{2.21c}$$

The basic convergence estimates obtained via this duality approach seem to be limited to the symmetric case where $\bar{E}_j = \bar{G}_j^*$ or the one sided cases, either $\bar{G}_j = I$ or $\bar{E}_j = I$. Let us develop the argument and then collect the results.

First of all, in steps (i) and (v) let m = 1. This involves no loss of generality as we are merely identifying $\bar{E}_j^m$ with $\bar{E}_j$, $\bar{G}_j^m$ with $\bar{G}_j$. As usual, let ε_j denote the error before the j multigrid cycle and $\bar{\varepsilon}_j$ the error after that cycle. Let $\varepsilon_j^{(1)}$ denote the error after step (i). Let $\varepsilon_j^{(2)}$ denote the error after step (iv). Then

$$\varepsilon_j^{(1)} = \bar{G}_j\varepsilon_j = T_j\bar{G}_j\varepsilon_j + \$_j\bar{G}_j\varepsilon_j \; . \tag{2.22a}$$

Since the next steps, (ii), (iii) and (iv) are an effort to solve for $\$_j\bar{G}_j\varepsilon_j$, we have

$$\varepsilon_j^{(2)} = T_j\bar{G}_j\varepsilon_j + \$_j(\bar{G}_j\varepsilon_j - \hat{\varepsilon}) \; , \tag{2.22b}$$

where, we assume that

$$\|\$_j(\bar{G}_j\varepsilon_j - \hat{\varepsilon})\|_A \le \delta_{j-1}^{\mu}\|\$_j\bar{G}\varepsilon_j\|_A \; , \tag{2.22c}$$

with the quantity δ_{j-1} a bound on the error for the $(j-1)$ multigrid $MG(j-1, \, . \, , \, .)$. In fact, we write

$$\$_j(\bar{G}_j\varepsilon_j - \hat{\varepsilon}) = \delta_{j-1}^{\mu}\$_j a \tag{2.23a}$$

with

$$\|\$_j a\|_A \le \|\$_j\bar{G}_j\varepsilon_j\|_A \; . \tag{2.23b}$$

Let w be an arbitrary element of S_j and consider the innerproduct

$$\langle w, \bar{\varepsilon}_j\rangle_A = \langle w, \bar{E}_j\varepsilon_j^{(2)}\rangle_A \; . \tag{2.24}$$

Using (2.22b) and (2.23) we have

$$\begin{aligned}\langle w, \bar{\varepsilon}_j\rangle_A &= \langle \bar{E}_j^* w, T_j\bar{G}_j\varepsilon_j + \delta_{j-1}^{u}\$_j a\rangle_A \\ &= (1-\delta_{j-1}^{\mu})\langle \bar{E}_j^* w, T_j\bar{G}_j\varepsilon_j\rangle_A \\ &+ \delta_{j-1}^{\mu}\langle \bar{E}_j^* w, T_j\bar{G}_j\varepsilon_j + \$_j a\rangle_A \; .\end{aligned}$$

Using the A-orthogonality of T_j and $\$_j$ and (2.23b) together with the basic estimates (2.20a), (2.20b) we have

$$\begin{aligned}\langle w, \bar{\varepsilon}_j\rangle_A \le (1-\delta_{j-1}^{\mu})&[t^*(w)e(w)t(\varepsilon_j)g(\varepsilon_j)]^{\frac{1}{2}}\|w\|_A \cdot \|\varepsilon_j\|_A \\ &+ \delta_{j-1}^{\mu}[e(w)g(\varepsilon_j)]^{\frac{1}{2}}\|w\|_A \cdot \|\varepsilon_j\|_A \; .\end{aligned} \tag{2.25}$$

Hence

$$\frac{\|\bar{\varepsilon}_j\|_A}{\|\varepsilon_j\|_A} \le \sup_{u,w} \left\{ (1-\delta_{j-1}^{\mu})[t^*(w)e(w)t(u)g(u)]^{\frac{1}{2}} + \delta_{j-1}^{\mu}[e(w)g(u)]^{\frac{1}{2}} \right\} \tag{2.26}$$

Now let us consider the cases of interest.

Theorem 2.4: Let $\delta = \delta(\bar{G}_j, \mu)$ and $\bar{\delta} = \bar{\delta}(\bar{E}_j, \mu)$ be two constants which are independent of j [i.e., depend on μ and the families of $\{\bar{G}_j\}$, $\{\bar{E}_j\}$] and satisfy

$$0 < \delta,\ \bar{\delta} \le 1 \tag{2.27}$$

and, for every $u \in S_j$

$$(1-\delta^{\mu})[t(u)g(u)]^{\frac{1}{2}} + \delta^{\mu}[g(u)]^{\frac{1}{2}} \le \delta \tag{2.28a}$$

$$(1-\bar{\delta}^{\mu})[t^*(u)e(u)]^{\frac{1}{2}} + \bar{\delta}^{\mu}[e(u)]^{\frac{1}{2}} \le \bar{\delta}\ . \tag{2.28b}$$

Then

$$\|\bar{\varepsilon}_j\|_A \le \min\,(\delta, \bar{\delta})\|\varepsilon_j\|_A\ . \tag{2.29}$$

Proof: For definiteness, assume that $\delta = \min\,(\delta, \bar{\delta})$. As in the discussion above, let δ_j be a bound for the convergence factor of $MG(j,\cdot,\cdot)$. Since $\delta_0 = 0$ we may make the inductive assumption

$$\delta_{j-1} \le \delta\ . \tag{2.30}$$

Since $t^*(w) \le 1,\ e(w) \le 1$ the estimate (2.26) yields

$$\begin{aligned}\frac{\|\bar{\varepsilon}_j\|_A}{\|\varepsilon_j\|_A} &\le \sup_{u} \left\{ (1-\delta_{j-1}^{\mu})[t(u)g(u)]^{\frac{1}{2}} + \delta_{j-1}^{\mu}[g(u)]^{\frac{1}{2}} \right\} \\ &\le \sup_{u} \left\{ (1-\delta^{\mu})[t(u)g(u)]^{\frac{1}{2}} + \delta^{\mu}[g(u)]^{\frac{1}{2}} \right\} \le \delta\ . \qquad \blacksquare\end{aligned}$$

Remark: Since

$$(1-\delta^{\mu})[t(u)g(u)]^{\frac{1}{2}} + \delta^{\mu}[g(u)]^{\frac{1}{2}} \le (1-\delta^{\mu})[t(u)g(u)]^{\frac{1}{2}} + \delta^{\mu}$$

we obtain an upper bound for δ by finding a root x of the equation

$$(1-x^{\mu})\max_{u}\,[t(u)g(u)]^{\frac{1}{2}} + x^{\mu} = x$$

which satisfies $0 < x < 1$. Observe that in the case of the W-cycle ($\mu \geq 2$) this result requires

$$\|T_j \bar{G}_j\|_A^2 = \sup\ \left[t(u)g(u)\right] < \frac{1}{4}$$

while the result of [10] given in Theorem 2.3 for the special case where $\bar{E}_j = I$ (which implies $t^*(u) = e(u) = 1$) only requires

$$\|T_j \bar{G}_j\|_A^2 = \sup\ [t(u)g(u)] < \frac{1}{2}\ .$$

While Theorem 2.4 is of interest for the one-sided W-cycles the duality argument yields stronger results in the (truly) symmetric case.

Theorem 2.5: Consider the case where

$$\bar{E}_j = \bar{G}_j^*\ . \tag{2.31}$$

Let $\hat{\delta} = \hat{\delta}(\bar{G}_j, \mu)$ be independent of j and satisfy

$$0 < \hat{\delta} < 1 \tag{2.32a}$$

and, for every $u \in S_j$

$$(1 - \hat{\delta}^\mu)[t(u)g(u)] + \hat{\delta}^\mu[g(u)]\ \leq\ \hat{\delta}\ . \tag{2.32b}$$

Then

$$\|\bar{\varepsilon}_j\|_A \leq \hat{\delta}\|\varepsilon_j\|_A\ . \tag{2.33}$$

Proof: Once more, we may assume $\delta_{j-1} \leq \hat{\delta}$. Then (2.26) and Schwartz's inequality yields

$$\delta_j \leq \sup_{u,w}\ [(1-\hat{\delta}^\mu)[t^*(w)e(w)] + \hat{\delta}^\mu e(w)]^{\frac{1}{2}}[(1-\hat{\delta}^\mu)[t(u)g(u)] + \hat{\delta}^\mu g(u)]^{\frac{1}{2}}\ .$$

However $t^* = t$ and $e = g$. Hence (2.32b) gives (2.33). ∎

Remark: Since $0 \leq t(u), g(u) \leq 1$ we see that, for every $x \in [0,1]$

$$(1-x)t(u)g(u) + xg(u) \leq (1-x)[t(u)g(u)]^{\frac{1}{2}} + x[g(u)]^{\frac{1}{2}}\ .$$

Hence,

$$\hat{\delta}(\bar{G}_j, \mu) \leq \delta(\bar{G}_j, \mu)) \,. \tag{2.34}$$

Thus, Theorem 2.4 shows that the one-sided multigrid scheme with $\mu = 2$ is convergent if

$$\|T_j \bar{G}_j\|_A^2 = \eta < \frac{1}{4} \,.$$

While Theorem 2.5 shows that the symmetric multigrid scheme with $\mu = 2$ is convergent if

$$\|T_j \bar{G}_j\|_A^2 = \eta < \frac{1}{2} \,.$$

Note that, in general, the duality arguments for the W-cycle give weaker results than the estimate of [10].

Consider the case $\mu = 1$, i.e. the V-cycle. Then

$$\hat{\delta} = \sup \frac{t(u)g(u)}{1 - g(u) + t(u)g(u)} \,. \tag{2.35}$$

Theorem 2.6: Let

$$D := \{u \,\epsilon\, S_j : g(u) = 1\} \,. \tag{2.36}$$

Suppose there is a $u_0 \,\epsilon\, S_j \,\cap\, D$ and

$$t(u_0) \neq 0 \,.$$

Then

$$\hat{\delta} = 1 \,. \tag{2.37}$$

Proof:

$$\frac{t(u_0)g(u_0)}{1 - g(u_0) + t(u_0)g(u_0)} = \frac{t(u_0)}{t(u_0)} = 1 \,. \qquad \blacksquare$$

Let us now compare the convergence proofs for the V-cycle.

Theorem 2.7: Let α be the optimal constant in (2.11b). That is

$$\alpha = \sup \frac{\|T_j \bar{G}_j u\|_A^2}{\|u\|_A^2 - \|\$_j \bar{G}_j u\|_A^2} \,. \tag{2.38}$$

Let $\hat{\delta}$ be given by (2.35). Then

$$\alpha \leq \hat{\delta} \ . \tag{2.39}$$

Proof: Since

$$\|\bar{G}_j u\|_A^2 = \|\$_j \bar{G}_j u\|_A^2 + \|T_j \bar{G}_j u\|_A^2$$

we see that

$$\alpha = \sup \ \frac{\|T_j \bar{G}_j u\|_A^2}{\|u\|_A^2 + \|T_j \bar{G}_j u\|_A^2 - \|\bar{G}_j u\|_A^2} \ .$$

We only increase the right-hand-side if we maximize $\|\bar{G}_j u\|_A^2$. Hence, we may replace $\|\bar{G}_j u\|_A^2$ by $g(u)\|u\|_A^2$ and obtain

$$\alpha \leq \sup \ \frac{\|T_j \bar{G}_j u\|_A^2}{[1 - g(u)]\|u\|_A^2 + \|T_j \bar{G}_j u\|_A^2} \ .$$

The right-hand-side is monotone increasing in $\|T_j \bar{G}_j u\|_A^2$. Hence we may replace $\|T_j \bar{G}_j u\|_A^2$ by $t(u)g(u)\|u\|_A^2$ and obtain (2.39). ∎

Corollary: Consider the general $(MG(j, u, f))$ multigrid iterative method. Let

$$\hat{\delta}_1 = \hat{\delta}(\bar{G}_j, 1), \ \hat{\delta}_2 = \hat{\delta}(\bar{E}_j^*, 1) \ . \tag{2.40}$$

As usual ε_j be the error before the multigrid cycle and $\bar{\varepsilon}_j$ be the error after that cycle. Then

$$\|\bar{\varepsilon}_j\|_A \leq [\hat{\delta}_1 \hat{\delta}_2]^{\frac{1}{2}} \ \|\varepsilon_j\|_A \ . \tag{2.41}$$

Proof: Apply Theorem 2.1, equation (2.14) and Theorem 2.7.

Remark: The inequality (2.39) is not really so surprising. The defining properties of $t(u)$, $g(u)$ involve inequalities. If these were equalities, then using the fact

$$\|T_j \bar{G}_j u\|_A^2 = \|\bar{G}_j u\|_A^2 - \|S_j \bar{G}_j u\|_A^2$$

we would find that (2.38) and (2.35) imply that

$$\hat{\delta} = \alpha \ .$$

We now consider the case where it is possible to associate a real value $\sigma \in [0,1]$ with u so that

$$\sigma = \sigma(u) , \quad 0 \leq \sigma \leq 1 , \tag{2.42a}$$

$$t(u) = J(\sigma) , \quad g(u) = \gamma(\sigma) . \tag{2.42b}$$

Remark: It is easy to see that the convergence proofs of Braess [4] and Yserentant [17] fall in this category. In the case of [4] we have $\sigma = \rho$ and

$$g(u) = \gamma(\rho) = \rho^k , \quad 0 \leq \rho \leq 1 , \tag{2.43a}$$

$$t(u) = J(\rho) = \frac{1-\rho}{2-\rho} , \quad 0 \leq \rho \leq 1 . \tag{2.43b}$$

In the case of [17] we have $\sigma = \rho$ and

$$g(u) = \gamma(\rho) = \rho^{2k} , \quad 0 \leq \rho \leq 1 , \tag{2.44a}$$

$$t(u) = J(\rho) = \min \{1, \ c(1-\rho)\}\rho^{4k} . \tag{2.44b}$$

Observe that in both cases $\gamma(\rho) = 1$ if and only if $\rho = 1$. And, $J(1) = 0$. In general $g(u) = \gamma(\sigma) < 1$, but as the dimension of S_j increases $\|g(u)\|_\infty$ approaches 1. That is

$$\sup_{\dim S_j \to \infty} \ \{g(u) : u \in S_j\} \ \to \ 1- \ . \tag{2.43}$$

As we have seen, the set D is an important set in this analysis. While the discussion above implies D may be empty, we must consider the set $D_\gamma := \{\sigma \in [0,1], \ \gamma(\sigma) = 1\}$.

Let us see how the estimate (2.15b), or estimates related to (2.15b) arise. First of all, let us recall that in most cases we have

$$\frac{\|\bar{G}_j^{r+1} u\|_A}{\|\bar{G}_j^{r} u\|_A} \geq \frac{\|\bar{G}_j^{r} u\|_A}{\|\bar{G}_j^{r-1} u\|_A} , \quad r \geq 1 . \tag{2.45}$$

Indeed, this is a basic premise of the heuristic approach to multigrid - many smoothing steps become ineffective as, for most of the classical smoothers,

$$\frac{\|\bar{G}_j^{r+1} u\|_A}{\|\bar{G}_j^{r} u\|_A} \ \to \ 1 - Ch^p \ \text{ as } r \ \to \ \infty .$$

Suppose that we have determined $J(\sigma)$ and $\gamma(\sigma)$ for a smoothing step. Consider k smoothing steps. Let

$$\hat{\sigma} = \sigma(\bar{G}_j^{k-1}u) \tag{2.46a}$$

Then (2.45) implies

$$\frac{\|\bar{G}_j^k u\|_A^2}{\|u\|_A^2} = \frac{\|\bar{G}_j^k u\|_A^2}{\|\bar{G}_j^{k-1}u\|_A^2} \cdots \frac{\|\bar{G}_j u\|_A^2}{\|u\|_A^2} \leq \gamma(\hat{\sigma})^k \, , \tag{2.46b}$$

and

$$\|T_j \bar{G}_j^k u\|_A^2 \leq J(\hat{\sigma})\|\bar{G}_j^{k-1}u\|_A^2 \leq J(\hat{\sigma})\|u\|_A^2 \, . \tag{2.46c}$$

Thus, considering the V-cycle based on k smoothing steps, we deal with

$$\hat{\delta}_k = \sup_{\hat{\sigma}} \frac{J(\hat{\sigma})[\gamma(\hat{\sigma})]^k}{1 + [\gamma(\hat{\sigma})]^k J(\hat{\sigma}) - [\gamma(\hat{\sigma})]^k} \, . \tag{2.47}$$

This may not be an optimal choice of functions $J(\hat{\sigma})$ and $\gamma(\hat{\sigma})^k$. Nevertheless, according to the theory developed above, it may provide an estimate.

Theorem 2.8: Suppose $\hat{\delta}(\bar{G}_j, 1) < 1$. Hence,

$$J(\hat{\sigma}) = 0 \, , \quad \text{for all } \; \hat{\sigma} \in D_\gamma \, . \tag{2.48}$$

Indeed,

$$\sup_{D\gamma} \left\{ \frac{J'(\hat{\sigma})}{J'(\hat{\sigma}) - \gamma'(\hat{\sigma})} \right\} < 1 \, . \tag{2.49}$$

Assume that

$$C = \sup_{D_\gamma} \left\{ \frac{-J'(\hat{\sigma})}{\gamma'(\hat{\sigma})} \right\} > 0 \, . \tag{2.50}$$

Then

$$\hat{\delta}_k \geq \frac{C}{C+k} \, . \tag{2.51}$$

Proof: We apply L'Hospital's rule to this function at the points $\hat{\sigma}_\gamma \in D_\gamma$. Then

$$\lim_{\sigma \to \hat{\sigma}_\gamma} \hat{\delta}_k(\hat{\sigma}) = \frac{J'(\hat{\sigma}_\gamma)}{J'(\hat{\sigma}_\gamma) - k\gamma'(\hat{\sigma}_\gamma)} \, . \qquad \blacksquare$$

REFERENCES

1. N. S. Bahvalov, On the convergence of a relaxation method with natural constraints on the elliptic operator. *Zh. Vychisl. Mat. i Mat. Fiz.* 6, 861-883 (1966); *U.S.S.R. Comput. Math. and Math. Phys.* 6, 101-135.

2. R. Bank and Craig C. Douglas, Sharp estimates for multigrid rates of convergence with general smoothing and acceleration *SIAM J. Numer. Anal. 22*, 617-633 (1985).

3. R. Bank and T. Dupont; An optimal order process for solving elliptic finite-element equations, *Math. Comp. 36*, 35-51, (1981).

4. D. Braess, The convergence rate of a multigrid method with Gauss-Seidel relaxation for the Poisson equation, *Math. of Comp.* 42, 505-519 (1984).

5. D. Braess and W. Hackbusch, A new convergence proof for the multigrid method including the V-cycle, *SIAM J. Numer. Anal.* 20, 967-975 (1983).

6. W. Hackbusch, On the convergence of multi-grid iterations, Beiträge Numer. *Math.* 9, 213-239 (1981).

7. D. Kamowitz and S. V. Parter, A study of some multigrid ideas, Report #545, Computer Science Dept., Univ. of Wisconsin; *Appl. Math. Comput., 17*, 153-184 (1985).

8. J.-F. Maitre and F. Musy, Multigrid methods: Convergence theory in a variational framework, *SIAM J. Numer. Anal. 21*, 657-671 (1984).

9. J.-F. Maitre and F. Musy, Multigrid methods for symmetric variational problems: A general theory and convergence estimates for usual smoothers, *To Appear: Appl. Math. Computing*

10. S. McCormick and J. Ruge, Multigrid methods for variational problems, *SIAM J. Numer. Anal.* 19, 924-929 (1982).

11. S. McCormick, Multigrid methods for variational problems: Further results, *SIAM J. Numer. Anal.* 21, 255-263 (1984).

12. S. McCormick, Multigrid methods for variational problems: General theory, *SIAM J. Numer. Anal.* *22,* 634-643 (1985).

13. R. A. Nicolaides, On the l^2 convergence of an algorithm for solving finite-element equation, *Math. Comp.* 31, 892-906 (1977).

14. S. V. Parter, A note on convergence of the multigrid V-cycle, *Appl. Math. and Comp.* *17*, 137-151 (1985).

15. K. Stüben, *MGϕ1*: A multi-grid program to solve $\triangle u - c(x,y)U = f(x,y)$ on Ω, $U = g(x,y)$ on $\partial\Omega$, on nonrectangular bounded domains Ω. IMA-Report: 82.02.02 Gesellschaft för Mathematik and Datenverarbeitung MBH, Bonn (1982).

16. R. Verfürth, The contraction number of a multigrid method with mesh ratio 2 for solving Poisson's equation, *Linear Algebra and its Appl.* *60,* 113-128 (1984).

17. H. Yserentant, On the convergence of multi-level methods for strongly nonuniform families of grids and any number of smoothing steps per level, *Computing* 30, 305-313 (1983).