East-West J. Numer. Math., Vol. 2, No. 3, pp. 00–00 (1995)

A Unified Multigrid Theory for Non-Nested Grids and/or Quadrature

C. C. DOUGLAS,* J. DOUGLAS, JR.† and D. E. FYFE‡

Received in revised form October 24, 1994

Abstract — In this paper, we extend some results from an earlier paper (SIAM J. Numer. Anal., 30 (1993), pp. 136–158) of the first two authors. We provide a unified theory for multilevel and multigrid methods when the usual assumptions are not present. For example, we do not assume that the solution spaces or the grids are nested. Further, we do not assume that there is an algebraic relationship between the linear algebra problems on different levels.

What we provide is a computationally useful theory for adaptively changing levels. Theory is provided for multilevel correction schemes, nested iteration schemes, and one way (i.e., coarse to fine grid with no correction iterations) schemes. We include examples showing the applicability of this theory: finite element examples using quadrature in the matrix assembly and finite volume examples with non-nested grids. Our theory applies directly to finite difference, wavelet, and collocation based multilevel examples as well.

Keywords. multigrid, multilevel.

American Mathematical Society Classification. 65N15, 65N10.

1. INTRODUCTION

In this paper, some concepts and theory first developed in [4] are extended. We do not make many assumptions in this paper. In particular, the grids do not necessarily have to be nested. The norms correspond to inner products on a grid, but the inner products do not have to be identical from level to level. We do not assume that there is an algebraic relationship between the linear algebra problems on different levels.

We provide what is really three level analysis rather than the more traditional two level theory. Among other things, this provides a rigorous basis for adaptively changing levels.

First, assume that there are j spaces $\mathcal{M}_k$, $1 \le k \le j$, approximating some solution space $\mathcal{M}$. Further, assume that $\dim \mathcal{M}_k \le \dim \mathcal{M}_{k+1}$. In typical multigrid applications, the growth rate is geometric at a rate related to the dimension of the spatial domain.

While the solution to the linear problem

$$Au + f = 0, \quad u, f \in \mathcal{M}, \quad A \in \mathcal{L}(\mathcal{M}),$$

*Mathematical Sciences Department, IBM Research Division, Thomas J. Watson Research Center, P. O. Box 218, Yorktown Heights, NY 10598-0218, and CERFACS, 42 Avenue Coriolis, 31057 Toulouse CEDEX, France, and Department of Computer Science, Yale University, P. O. Box 208285, New Haven, CT 06520-8285, USA

†Department of Mathematics, Purdue University, Mathematical Sciences Building, West Lafayette, IN 47907, USA

‡Laboratory for Computational Physics and Fluid Dynamics, Naval Research Laboratory, Washington, DC 20375, USA (This was supported under ONR N0001494WX35116)

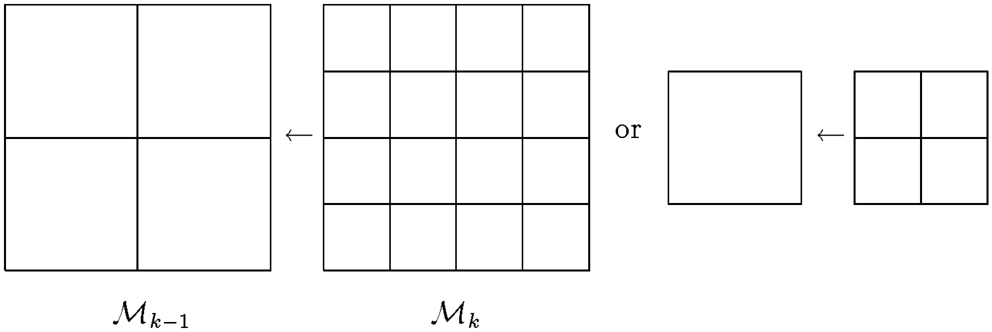

Figure 1: Square grid elements for finite differences or elements

is really sought, a set of approximate problems

$$A_k u_k + f_k = 0, \quad u_k, f_k \in \mathcal{M}_k, \quad A_k \in \mathcal{L}(\mathcal{M}_k), \tag{1.1}$$

will be solved approximately instead.

As usual in multigrid procedures, two sets of mappings between neighboring spaces are assumed to exist. The prolongation (or interpolation) mappings are

$$\mathcal{P}_{k-1} : \ \mathcal{M}_{k-1} \to \mathcal{M}_k.$$

The restriction (or projection) mappings are

$$\mathcal{R}_k : \ \mathcal{M}_k \to \mathcal{M}_{k-1}.$$

In some cases, each A_k is related to A_{k+1} by

$$A_k = \mathcal{R}_{k+1} A_{k+1} \mathcal{P}_k; \tag{1.2}$$

however, the basic theorem in this paper, Theorem 1, does not assume this relation.

For partial differential equations that are discretized in a standard fashion, there can be natural definitions for $\mathcal{R}_{k+1}$ and $\mathcal{P}_k$. Some of these are shown graphically in [4].

For most finite difference schemes, $\mathcal{P}_k$ is related to a conventional linear interpolation scheme; $\mathcal{R}_{k+1}$ usually corresponds to a weighted average of neighboring points. Many times, $\mathcal{R}_{k+1} = c\mathcal{P}_k^T$, where $c \in \mathbb{R}$. See Figure 1 for an example set of grids.

For many finite element schemes, a Galerkin method is usually employed. Here, $\mathcal{P}_k$ corresponds to an interpolation scheme that makes sense for changing from the basis associated with grid k to the basis associated with grid $k+1$; $\mathcal{R}_{k+1}$ normally is given by $\mathcal{P}_k^T$. Once again, see Figure 1 for an example set of grids or [4].

For many finite volume schemes, $\mathcal{P}_k$ and $\mathcal{R}_{k+1}$ are similar to the ones in the finite element case. While the grids are nested, the location of the unknowns in the matrix problem are not necessarily nested. For example, in Figure 2, the dots represent the location of the unknowns in a finite volume discretization for the grid lines shown.

Associated with each level is a norm, $\|\cdot\|_k$. Assume that

$$C_{1,k}\|u\|_k \le \|\mathcal{P}_k u\|_{k+1} \le C_{2,k}\|u\|_k, \quad \forall u \in \mathcal{M}_k, \tag{1.3}$$

where the forms of $C_{1,k}$ and $C_{2,k}$ are known; these constants can depend on the coefficients in the differential problem and on the grid. A large value of $C_{2,k}$ will inhibit the rate of convergence.

Now, define a k-level *standard* correction multilevel algorithm:

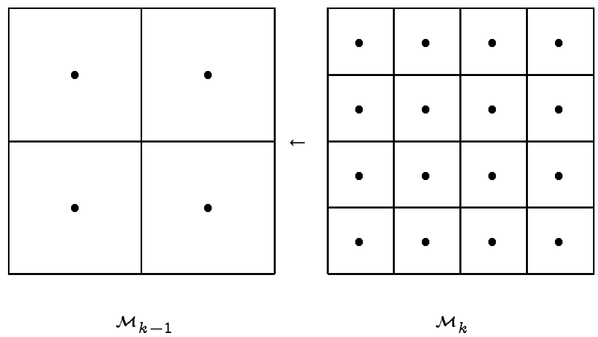

Figure 2: Non-nested grids for finite volumes

ALGORITHM MG(k, $\{\mu_\ell\}_{\ell=1}^{j}$, x_k, f_k)
(1) If $k=1$ or $\mu_k=0$, then solve $A_k x_k + f_k = 0$ to some accuracy.
(2) If $k>1$ and $\mu_k>0$, then repeat (2a)–(2c) for $i=1,\cdots,\mu_k$:
  (2a) Update x_k using the pre-solver.
  (2b) Solve a residual correction problem on level $k-1$:
    $x_k \leftarrow x_k + \mathcal{P}_{k-1}$ MG($k-1$, $\{\mu_\ell\}_{\ell=1}^{j}$, 0, $\mathcal{R}_k($ $A_k x_k + f_k$ $)$).
  (2c) Update x_k using the post-solver.
(3) Return x_k.

It is assumed that $0 \le \mu_1, \mu_j \le 1$ in this definition. In practice, $\mu_j > 1$ is common, but this can be interpreted as the repetition μ_j times of the algorithm for the case of $\mu_j = 1$.

On all but level 1 (the coarsest grid level), two solvers are associated with a level: a pre-solver and a post-solver. These surround the coarser level correction (2b). In most real applications, only one solver is associated with a level (one of the pre- or post-solvers is the identity operation). The solvers can be smoothers, roughers, or direct solvers.

In order to analyze Algorithm MG from an iterative method viewpoint, we transform it into a *nonstandard* form similar to that introduced in [4]. First, add an additional level $j+1$, which is just a repetition of level j:

$$\mathcal{M}_{j+1} = \mathcal{M}_j, \quad \mathcal{P}_j = \mathcal{R}_{j+1} = I, \quad A_{j+1} = A_j, \quad C_{1,j} = C_{2,j} = 1.$$

The initial residual z_{j+1} is then given by

$$A_{j+1} x_{j+1} + f = z_{j+1}.$$

All analysis can now be done using residual correction problems.

Define the following:

z_{k+1}	The residual on level $k+1$ at some step.
$x_k^{(-1)}$	The initial guess for level k; this is normally 0, except on level $j+1$.

Now, define a k-level *nonstandard* correction multilevel algorithm:

ALGORITHM NSMG(k, $\{\mu_\ell\}_{\ell=1}^j$, z_{k+1}, $x_k^{(-1)}$)

(1) Initial residual: $\mathcal{R}_{k+1} z_{k+1} \in \mathcal{M}_k$.

(2) Initial pre-solve: Update $x_k^{(-1)}$ to get $x_k^{(0)}$ such that
$$A_k x_k^{(0)} + \mathcal{R}_{k+1} z_{k+1} = z_k^{(0)}, \text{ where } \|z_k^{(0)}\| \le \rho_k^{(1)} \|z_{k+1}\|.$$

(3) Let $\hat{x}_k^{(1)} = x_k^{(0)}$, $\hat{z}_k^{(1)} = z_k^{(0)}$, and $\gamma_1^{(1)} = 0$.

(4) If $\mu_k > 0$, then repeat $i = 1, \cdots, \mu_k$:

- (4a) If $i > 1$, then
  - (4a1) Residual: $A_k x_k^{(i-1)} + \mathcal{R}_{k+1} z_{k+1} = \hat{\theta}_k^{(i)}$.
  - (4a2) Pre-solve: Update $x_k^{(i-1)}$ to get $\hat{x}_k^{(i)}$ such that
  $$A_k \hat{x}_k^{(i)} + \mathcal{R}_{k+1} z_{k+1} = \hat{z}_k^{(i)}, \text{ where } \|\hat{z}_k^{(i)}\| \le \rho_k^{(i)} \|\hat{\theta}_k^{(i)}\|.$$
- (4b) If $k > 1$, then
  - (4c1) Correction: $\gamma_k^{(i)} = \mathcal{P}_{k-1} \bar{x}_{k-1}^{(i)}$, where
  $$\bar{x}_{k-1}^{(i)} = \text{NSMG}(k-1, \{\mu_\ell\}_{\ell=1}^j, \hat{z}_k^{(i)}, 0) \text{ and}$$
  $$A_{k-1} \bar{x}_{k-1}^{(i)} + \mathcal{R}_k \hat{z}_k^{(i)} = \bar{z}_{k-1}^{(i)}.$$
- (4c) Calculate $\sigma_k^{(i)}$:
$$\|\hat{z}_k^{(i)} + A_k \mathcal{P}_{k-1} \bar{x}_{k-1}^{(i)}\| \le \sigma_k^{(i)} \|\hat{z}_k^{(i)} + \mathcal{P}_{k-1} A_{k-1} \bar{x}_{k-1}^{(i)}\|.$$
- (4d) Residual: $A_k(\hat{x}_k^{(i)} + \gamma_k^{(i)}) + \mathcal{R}_{k+1} z_{k+1} = \theta_k^{(i)}$.
- (4e) Post-solve: Update $\hat{x}_k^{(i)} + \gamma_k^{(i)}$ to get $x_k^{(i)}$ such that
$$A_k x_k^{(i)} + \mathcal{R}_{k+1} z_{k+1} = z_k^{(i)}, \text{ where } \|z_k^{(i)}\| \le \epsilon_k^{(i)} \|\theta_k^{(i)}\|.$$

(5) Return $x_k^{(\mu_k)}$.

This is almost the same algorithm as was analyzed in [4]. The difference is in step (4c). Here we calculate the norm of the difference between the effect of two similar operators on the correction with respect to the residual before the correction was computed.

Consider the example of adaptively changing levels based on reducing the residual norm adequately. We can calculate $\sigma_k^{(i)}$ while computing a correction in step (4b). Based on the size of $\sigma_k^{(i)}$, we can determine if the current candidate for $\bar{x}_{k-1}^{(i)}$ is sufficient in order to maintain convergence on level k (or a fast enough convergence rate). Should $\sigma_k^{(i)}$ be too large, more corrections on level $k-2$ or a better approximation on level $k-1$ might be appropriate.

In order to consider a priori analysis, the actual forms for $\rho_k^{(i)}$ and $\epsilon_k^{(i)}$ should be substituted. Examples of these forms can be found for various elliptic partial differential equations and iterative solvers in [1] and [2].

A second multigrid variant is a nested iteration scheme, which begins computation on level 1 and traverses the levels to some level j, using each level k, $k < j$, to generate an initial guess for level $k+1$ and possibly for solving residual correction problems. Define a k-level *standard* nested iteration multigrid scheme by

ALGORITHM NI($j, \{\mu_k\}_{k=1}^j, x_1, \{f_k\}_{k=1}^j$)

(1) For $k = 1, \cdots, j$, do
- ($1a$) If $k > 1$, then $x_k \leftarrow \mathcal{P}_{k-1} x_{k-1}$
- ($1b$) $x_k \leftarrow$MG($k, \{\mu_\ell\}_{\ell=1}^k, x_k, f_k$)

(2) Return x_j

Note that $\mu_\ell = 1$, all ℓ, corresponds to full multigrid (or nested iteration V cycle). Choosing $\mu_\ell = 0$, all ℓ, corresponds to one way multigrid, i.e., no correction cycles whatsoever (see [2] and [6]).

Define a *nonstandard* nested iteration multilevel algorithm by

Algorithm NSNI(j, $\{\mu_\ell\}_{\ell=1}^{j}$, $x_1^{(-1)}$)
(1) repeat $k = 1, \cdots, j$:
  (1a) Initial guess: If $k > 1$, then
    (1a1) $x_k^{(-1)} = \mathcal{P}_{k-1} x_{k-1}^{(\mu_{k-1})}$.
  (1b) Residual: $z_k = A_k x_k^{(-1)} + f_k$.
  (1c) Solve: $x_k^{(\mu_k)}$ =NSMG(k, $\{\mu_\ell\}_{\ell=1}^{j}$, z_k, $x_k^{(-1)}$).
(2) Return $x_j^{(\mu_j)}$.

Theory for the nonstandard formulations is in Section 2. Examples are in Sections 3 and 4.

2. THEORY

In this section, we prove some basic theorems, based on a simple theory that is computationally useful, including for adaptively changing levels.

The basic theorem for Algorithm NSMG is the following:

Theorem 1. *Assume the following for all levels $1 \le k \le j$:*

1. *z_{j+1} is the residual on level $j+1 \ge 2$.*

2. *$z_k^{(i)}$ is the residual on level k at step i.*

3. $\|\hat{z}_k^{(i)} + A_k \mathcal{P}_{k-1} \bar{x}_{k-1}^{(i)}\| \le \sigma_k^{(i)} \|\hat{z}_k^{(i)} + \mathcal{P}_{k-1} A_{k-1} \bar{x}_{k-1}^{(i)}\|$.

4. $\|(I - \mathcal{P}_{k-1}\mathcal{R}_k) z_k^{(i)}\| \le \delta_k^{(i)} \|z_k^{(i)}\|$.

Let

$$E_1^{(1)} = \epsilon_1^{(1)} \rho_1^{(1)} \quad \textit{and} \quad E_k^{(\mu_k)} = \prod_{i=1}^{\mu_k} \left(\epsilon_k^{(i)} \rho_k^{(i)} \sigma_k^{(i)} \left[\delta_k^{(i)} + C_{2,k-1} E_{k-1}^{(\mu_{k-1})} \right] \right).$$

Then,

$$\|z_k^{(\mu_k)}\|_k \le E_k^{(\mu_k)} \|z_{k+1}\|_{k+1}. \tag{2.1}$$

Proof: The proof of (2.1) is a double induction argument. The result is trivial when $k = 1$. Assume that the result is true for all levels $\ell < k$.

We first assume that $\mu_k = 1$ and estimate $\|\theta_k^{(1)}\|$. The norm comparison (1.3) enters in the argument below, as well as the hypotheses listed above. Thus,

$$
\begin{aligned}
\|\theta_k^{(1)}\| &\leq \|A_k(x_k^{(0)}+\gamma_k^{(1)})+\mathcal{R}_{k+1}z_{k+1}\| \\
&= \|z_k^{(0)}+A_k\gamma_k^{(1)}\| = \|z_k^{(0)}+A_k\mathcal{P}_{k-1}\bar{x}_{k-1}^{(1)}\| \\
&\leq \sigma_k^{(1)}\|z_k^{(0)}+\mathcal{P}_{k-1}A_{k-1}\bar{x}_{k-1}^{(1)}\| \\
&= \sigma_k^{(1)}\|(I-\mathcal{P}_{k-1}\mathcal{R}_k)z_k^{(0)}+\mathcal{P}_{k-1}\bar{z}_{k-1}^{(1)}\| \\
&\leq \sigma_k^{(1)}\left(\delta_k^{(1)}\|z_k^{(0)}\|+C_{2,k-1}E_{k-1}^{(\mu_{k-1})}\|z_k^{(0)}\|\right) \\
&\leq \sigma_k^{(1)}\rho_k^{(1)}\left(\delta_k^{(1)}+C_{2,k-1}E_{k-1}^{(\mu_{k-1})}\right)\|z_{k+1}\|_{k+1}.
\end{aligned}
$$

Hence,

$$
\|z_k^{(1)}\|_k \leq E_k^{(1)}\|z_{k+1}\|_{k+1}.
$$

Now, assume that the result is true for $\mu_k = 1,\cdots,i-1$. Then,

$$
\begin{aligned}
\|\theta_k^{(i)}\| &\leq \|A_k(\hat{x}_k^{(i)}+\gamma_k^{(i)})+\mathcal{R}_{k+1}z_{k+1}\| \\
&= \|\hat{z}_k^{(i)}+A_k\gamma_k^{(i)}\| = \|\hat{z}_k^{(i)}+A_k\mathcal{P}_{k-1}\bar{x}_{k-1}^{(i)}\| \\
&\leq \sigma_k^{(i)}\|(I-\mathcal{P}_{k-1}\mathcal{R}_k)\hat{z}_k^{(i)}+\mathcal{P}_{k-1}\bar{z}_{k-1}^{(i)}\| \\
&\leq \sigma_k^{(i)}\left(\delta_k^{(i)}\|\hat{z}_k^{(i)}\|+C_{2,k-1}E_{k-1}^{(\mu_{k-1})}\|\hat{z}_k^{(i)}\|\right) \\
&\leq \sigma_k^{(i)}\rho_k^{(i)}\left(\delta_k^{(i)}+C_{2,k-1}E_{k-1}^{(\mu_{k-1})}\right)\|z_k^{(i-1)}\|,
\end{aligned}
$$

so that

$$
\|z_k^{(i)}\|_k \leq E_k^{(i)}\|z_{k+1}\|_{k+1}. \qquad \blacksquare
$$

A more precise analysis, based on an affine space decomposition of each $\mathcal{M}_k$, would follow the analysis in [4].

The basic theorem for Algorithm NSNI is the following:

THEOREM 2. *Make the same assumptions as in Theorem 1. Further assume that (1.1) is approximated by some ξ_k such that*

$$
A_k\xi_k + f_k = \theta_k \tag{2.2}
$$

starting from some initial guess $x_k = \mathcal{P}_{k-1}\xi_{k-1}$. Given some $\{\zeta_k\}_{k=1}^{j}$, we want

$$
\begin{cases}
\|\theta_1\| \leq \zeta_1\|A_1x_1+f_1\|, \\
\|\theta_k\| \leq \zeta_k\|\theta_{k-1}\|, \quad 1<k\leq j.
\end{cases} \tag{2.3}
$$

Then

$$
E_k^{(\mu_k)} \leq \zeta_k \quad \text{for } 1\leq k\leq j \tag{2.4}
$$

for an appropriate choice of $\{\{\rho_k^{(i)},\epsilon_k^{(i)}\}_{i=1}^{\mu_k}\}_{k=1}^{j}$.

The proof of Theorem 2 is obvious (see [2] for example). Note that by calculating $\delta_k^{(i)}$ and $\sigma_k^{(i)}$ as a computation progresses, the choice of $\rho_k^{(i)}$ and $\epsilon_k^{(i)}$ can be chosen adaptively to ensure that (2.4) is satisifed.

The one way multigrid method is a common computational method in engineering applications. It has been used for decades as a method for producing an initial guess on the grid in which a solution to a problem is actually wanted. This process is described in [6] for a procedure that he first saw in the 1920's.

Consider a typical partial differential equation problem to be solved numerically. It is discretized on a set of grids Ω_k, $1 \le k \le j$, with some notion of grid spacing (or a mesh diameter) h_k.

The basic theorem for one way multigrid is the following:

THEOREM 3. *Make the same assumptions as in Theorem 2. Further assume $\mu_k = 0$, $1 \le k \le j$, and that*

$$\theta_k = C h_k^q, \quad C, q, h > 0 \in \mathbb{R}.$$

Then

$$\zeta_k = C C_{2,k} (h_k / h_{k-1})^q$$

is adequate to ensure that (2.2) is satisfied. Hence, (2.4) is satisfied with $\rho_k^{(1)} = \zeta_k$.

Once again the proof is obvious. Note Theorem 3 gives a simple bound for one way multigrid that is independent of the solver used on each level.

3. AN EXAMPLE OF THE FAILURE OF $A_k = \mathcal{R}_{k+1} A_{k+1} \mathcal{P}_k$

Consider the Dirichlet problem

$$\begin{cases} -\sum_{i,j=1}^{2} (a_{ij}(x) u_{x_i})_{x_j} + b_i(x) U_{x_i} c(x) u &= f(x), \qquad x \in \Omega = [0,1]^2, \\ u &= 0, \qquad x \in \partial\Omega. \end{cases} \tag{3.1}$$

Let the k-level partition $\mathcal{S}_k$ of Ω consist of squares of side length $2^{-(k+\ell)}$, where ℓ is independent of k. Let the k-level finite element space $\mathcal{M}_k$ consist of C^0-bilinear functions over $\mathcal{S}_k$ that vanish on $\partial\Omega$. Then, the natural k-level Galerkin equations,

$$A_k u_k = \varphi_k, \tag{3.2}$$

would be generated by seeking a function $\tilde{u}^k \in \mathcal{M}_k$ such that

$$\sum_{i,j=1}^{2} (a_{ij} \tilde{u}^k_{x_i}, \tilde{v}^k_{x_j}) + \sum_{i=1}^{2} (b_i \tilde{u}^k_{x_i}, \tilde{v}^k) + (c\tilde{u}^k, \tilde{v}^k) = (f, \tilde{v}^k), \quad \tilde{v}^k \in \mathcal{M}_k, \tag{3.3}$$

where $(\cdot,\cdot)$ indicates the inner product on $L^2(\Omega)$; but exact integration is not, in general, feasible. Thus, it is usually necessary to invoke a quadrature rule to approximate the integrals in (3.3). It is well-known that a (2×2)-Gauss quadrature rule suffices to maintain unique solvability of the resulting linear equations, along with the proper asymptotic order of accuracy of the k-level approximation to the solution of (3.1). Denote by $(\cdot,\cdot)_G$ the (2×2)-Gauss quadrature approximation to $(\cdot,\cdot)$, and define the k-level equations (3.1) through the approximation

$$\sum_{i,j=1}^{3} (a_{ij} \tilde{u}^k_{x_i}, \tilde{v}^k_{x_j})_G + \sum_{i=1}^{2} (b_i \tilde{u}^k_{x_i}, \tilde{v}^k)_G + (c\tilde{u}^k, \tilde{v}^k)_G = (f, \tilde{v}^k)_G, \quad \tilde{v}^k \in \mathcal{M}_k.$$

Let us consider the feasibility of the relation $A_k = \mathcal{R}_{k+1} A_{k+1} \mathcal{P}_k$ by making a simple parameter count. If the prolongation and restriction operators are defined in terms of the parameters related to the vertex values of a single element in the coarser partition and the vertex values of the corresponding four squares in the finer partition, it suffices to consider a unit square S^1 for the coarser element (associated with index 1) and its partition (associated with index 2) into four squares, $S_j^2,\ j = 1, \ldots, 4$, for the finer elements. Note that the sixteen quadrature points on S_j^2 are distinct from the four quadrature points on S^1; thus, different values of the coefficients in the differential equation enter into the formation of the equations (3.1).

For the moment, let $\mathcal{M}_1$ be the span of the four bilinear basis functions associated with the vertices of S^1 and $\mathcal{M}_2$ the span of the nine bilinear basis functions associated with the vertices of $S_j^2,\ j = 1, \ldots, 4$. Then, let us generalize the question as to whether there exist $\mathcal{R}_{k+1}$ and $\mathcal{P}_k$ such that $A_k = \mathcal{R}_{k+1} A_{k+1} \mathcal{P}_k$ slightly by asking if there exist maps

$$P : \mathcal{M}_1 \longrightarrow \mathcal{M}_2 \qquad \text{and} \qquad Q : \mathcal{M}_1 \longrightarrow \mathcal{M}_2 \tag{3.4}$$

such that

$$(A_1 u, v) = (A_2 P u, Q v), \qquad u, v \in \mathcal{M}_1. \tag{3.5}$$

Let us make a simple parameter count. Each of the matrices P and Q has 36 entries. For each nontrivial coefficient $a_{ij},\ b_i$, or c, the quadrature rule associates sixteen values of the coefficient in the A_2-inner product and only four in the A_1-inner product; thus, twelve independent constraints arise for each such coefficient. Since there are seven possibly nontrivial, distinct coefficients, it is clear that it cannot always be possible to satisfy (3.5). If there were fewer coefficients to handle, the maps could exist but have rather strange relationships to standard interpolation procedures.

Let us ask a different question. Let us take reasonable definitions of P and Q and ask to what extent (3.5) fails for locally smooth coefficients. Let $P = Q$ be the embedding operator between $\mathcal{M}_1$ and $\mathcal{M}_2$, and consider the special case for which

$$a_{ij}(x) = \delta_{ij} a(x), \qquad b_i(x) = c(x) = 0. \tag{3.6}$$

It follows easily from the Bramble-Hilbert lemma that

$$(A_1 u, v) - (A_2 P u, Q v) = \mathcal{O}(\|u\| \|v\|), \qquad u, v \in \mathcal{M}_1, \tag{3.7}$$

where the norm is the norm in H^1. Then, if the analogous restriction and prolongation operators are used at each level,

$$\sigma_k^{(i)} = 1 + \mathcal{O}(h_k^2) \tag{3.8}$$

if the H^1-norm is employed at each level. Thus, using the naturally associated quadrature rule at each level is a reasonable choice for these choices for $\mathcal{R}_{k+1}$ and $\mathcal{P}_k$.

4. FINITE VOLUME EXAMPLE

Consider the two-point boundary value problem

$$\left\{ \begin{array}{rcll} -(a(x)u_x)_x + c(x)u &=& f(x), & x \in \Omega = [0,1], \\ u(0) = u(1) &=& 0. & \end{array} \right. \tag{4.1}$$

A finite volume discretization of (4.1) yields

$$a_{i-1/2} u_{i-1}^k + (\Delta x_i c_i - a_{i-1/2} - a_{i+1/2}) u_i^k + a_{i+1/2} u_{i+1}^k = \Delta x_i f_i, \quad i = 1, \ldots, N$$

on level k where $a_{i+1/2} = 2A_{i+1/2}/(\Delta x_i + \Delta x_{i+1})$, and Δx_i is the length of cell (interval) i. While the grid points $x_{i+1/2}$ in a finite volume multigrid procedure are nested, the locations of the unknowns u are not nested (see Figure 2).

Clearly, one would not use a multigrid approach to solve this problem. However, multigrid is a viable alternative for the equivalent multi-dimensional problem. The following remarks generalize to the multi-dimensional case through the use of tensor product formulations for the prolongation and restriction operators (see [5] for applications). We discuss the one-dimensional case for clarity.

Let us define a restriction matrix

$$\mathcal{R}_k = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 1 & 1 \end{pmatrix}.$$

This is just piecewise linear interpolation. We can also define a prolongation matrix $\mathcal{P}_{k-1} = 2{\mathcal{R}_k}^T$. This prolongation matrix corresponds to piecewise constant interpolation; clearly, not a very accurate choice, but a demonstrative one.

If we formulate the coarse grid matrix from $A_{k-1} = \mathcal{R}_k A_k \mathcal{P}_{k-1}$ and a restricted right-hand-side from $\mathcal{R}_k$, we obtain

$$a_{2i-3/2}u_{i-1}^{k-1} + (\Delta x_{2i}c_{2i} + \Delta x_{2i-1}c_{2i-1} - a_{2i-3/2} - a_{2i+1/2})u_i^{k-1} +$$

$$a_{2i+1/2}u_{i+1}^{k-1} = \Delta x_{2i-1}f_{2i-1} + \Delta x_{2i}f_{2i}, \quad i = 1, \dots, N/2.$$

This is a reasonable coarse grid approximation where the only difference from the finite volume discretization on the coarse grid would be in the use of the underlying fine grid to discretize the finite volume integral.

A straightforward calculation of $\|(I - \mathcal{P}_{k-1}\mathcal{R}_k)x)\|$ for arbitrary x shows that $\delta_k = 1/\sqrt{2}$.

A more practical prolongation matrix would use quadratic interpolation. On an equally spaced grid this becomes

$$\mathcal{P}_{k-1} = \frac{1}{32}\begin{pmatrix} 25 & -3 & 0 & 0 & \dots & 0 & 0 & 0 \\ 33 & 5 & 0 & 0 & \dots & 0 & 0 & 0 \\ 5 & 30 & -3 & 0 & \dots & 0 & 0 & 0 \\ -3 & 30 & 5 & 0 & \dots & 0 & 0 & 0 \\ 0 & 5 & 30 & -3 & \dots & 0 & 0 & 0 \\ 0 & -3 & 30 & 5 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 5 & 30 & -3 \\ 0 & 0 & 0 & 0 & \dots & -3 & 30 & 5 \\ 0 & 0 & 0 & 0 & \dots & 0 & 5 & 33 \\ 0 & 0 & 0 & 0 & \dots & 0 & -3 & 25 \end{pmatrix}.$$

The first and last rows of $\mathcal{P}_{k-1}$ have incorporated the uniform Dirichlet boundary conditions for the current problem. On a nonuniform grid the matrix entries involve ratios of mesh lengths. The use of this prolongation matrix in the definition of the coarse matrices would expand the bandwidth of each successive coarser matrix. This defeats the purpose of multigrid where one expects to do less work on the coarser grids. The use of this prolongation matrix with the piecewise linear interpolation restriction matrix gives $\delta_k = \sqrt{531}/32$.

5. CONCLUSIONS

In this paper, we extended the theory given in a previous paper in two ways. The theory here is more precise than the earlier one. Further, it is applicable to problems that are not nested and/or ones in which the linear systems use quadrature in their assembly. Finally, the theory here allows multigrid software (e.g., [3]) adaptively to change levels with a higher degree of precision than with the earlier theory.

REFERENCES

1. R. E. Bank and C. C. Douglas. Sharp estimates for multigrid rates of convergence with general smoothing and acceleration, *SIAM J. Numer. Anal.*, (1985) **22**, pp. 617–633.
2. C. C. Douglas. Multi–grid algorithms with applications to elliptic boundary–value problems, *SIAM J. Numer. Anal.*, (1984) **21**, pp. 236–254.
3. C. C. Douglas. Implementing abstract multigrid or multilevel methods. In: *Sixth Copper Mountain Conference on Multigrid Methods*, volume CP 3224, NASA, Hampton, VA, 1993, pp. 127–141.
4. C. C. Douglas and J. Douglas. A unified convergence theory for abstract multigrid or multilevel algorithms, serial and parallel. *SIAM J. Numer. Anal.*, (1993) **30**, pp. 136–158.
5. G. Patnaik, R. H. Guirguis, J. P. Boris, and E. S. Oran. A barely implicit correction for flux-corrected transport. *J. Comput. Phys.*, (1987) **71**, pp. 1–20.
6. R. V. Southwell. *Relaxation Methods in Engineering Science*, Oxford University Press, Oxford, 1940.