UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

An Energy-Minimizing Interpolation for Multigrid Methods

W. L. Wan

April 1997

CAM Report 97-18

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

(Presented at the Eighth Copper Mountain Conference on Multigrid Methods, April 6-11, Colorado, 1997)

AN ENERGY-MINIMIZING INTERPOLATION FOR MULTIGRID METHODS

W. L. WAN

Abstract. We shall study multigrid methods from energy minimizations and approximations. Through the analysis of an multigrid method in 1D, we introduce the concepts of stability and the approximation property in the classical theory. Based on them, we derive an energy-minimizing interpolation and present a two level analysis for it. Issues on coarsening are also addressed. Finally, we demonstrate the effectiveness of the multigrid method by applying it to unstructured grids computations and discontinuous coefficient problems.

1. Introduction. Multigrid methods have been widely used as an efficient solver for second order elliptic PDE's yet it is still not completely understood. This is partially revealed by the poor convergence of the standard multigrid methods applying to PDE's whose coefficients are anisotropic [16], discontinuous [1, 6, 9] or oscillatory [12, 31]. Special techniques, for instance, block smoothing [6], semi-coarsening [10, 30], matrix-dependent interpolations [24, 26], frequency decomposition [17, 11, 31], homogenization [12] were introduced to handle some of these cases. In this paper, we shall study multigrid methods from the energy minimization and the approximation point of views which hopefully gives an insight to a robust multigrid algorithm.

The success of multigrid hinges on the choice of coarse grids, the smoothing and the interpolation operators. In standard multigrid, full coarsening, Jacobi or Gauss-Seidel smoothing and linear interpolation are usually used. Classical convergence theory [5, 36, 37, 4, 16, 22, 29] shows that these simple ingredients are already optimal for smooth and some rough coefficient problems. In general, however, these choices may lead to slow convergence. To remedy the situation in one dimension, a well-known robust interpolation [16, 34, 24] was used. It is obtained by solving local homogeneous PDE's which are actually equivalent to minimizing the energy of the coarse grid basis functions. Through the 1D analysis in section 2, we introduce two important concepts for optimal convergence in the classical theory: stability and the approximation property.

The extension to higher dimensions is not obvious since no natural correspondence exists between the 1D case and the 2D case. Nonetheless, many attempts [15, 18, 1, 9, 16, 34, 25] have been made to set up similar local PDE's for defining the interpolation which, however, is not as robust as that in 1D. In section 3, we shall derive another interpolation by solving a minimization problem in place of local PDE's with special emphasis on stability and the approximation property.

The highlights of the present paper are the applications of our method to two kinds of problems: unstructured grids computations and discontinuous coefficient PDE's with irregular interface. The discussion and the numerical results are presented in section 4. Finally, we summarize our experience by several remarks in section 5.

2. One Dimension. We begin with the following model problem:

$$
\begin{aligned}
-\frac{d}{dx}a(x)\frac{d}{dx}u(x) &= f \qquad \text{in } (0,1) \\
u &= 0 \qquad \text{at } x = 0 \text{ and } x = 1,
\end{aligned} \tag{1}
$$

where $a(x)$ and $f(x)$ are integrable and $a(x)$ is uniformly positive on $[0,1]$.

We discretize (1) by finite element and solve the resulting linear system by a multigrid method with a special interpolation. Given a grid of size $h = 1/n$: $x_j^h = jh, j = 0, \ldots, n.$, let $V^h \subset H_0^1(0,1)$ be the usual piecewise linear finite element space. We define a coarse subspace: $V^H = \text{span}\{\phi_i^H\}$, for multigrid where $\{\phi_i^H\}$ is the set of coarse grid nodal basis functions to be defined. By definition, $\phi_i^H(x_{2i}^h) = 1$ and $\phi_i^H(x_{2i-2}^h) = \phi_i^H(x_{2i+2}^h) = 0$. Interpolation essentially is a way to define $\phi_i^H(x_{2i-1}^h)$ and $\phi_i^H(x_{2i+1}^h)$. For example, if we let them equal $1/2$, it is just linear interpolation. In fact, there is one to one correspondence between the set of coarse grid basis functions and the interpolation.

2.1. Formulation of Energy-Minimizing Interpolation. By the energy-minimization principle [28], we define $\phi_i^H(x)$ by solving the following local PDE in $[x_{2i-2}^h, x_{2i}^h]$:

$$-\frac{d}{dx}a(x)\frac{d}{dx}\phi_i^H(x) = 0 \qquad \text{in } [x_{2i-2}^h, x_{2i}^h], \tag{2}$$
$$\phi_i^H(x_{2i-2}^h) = 0, \quad \phi_i^H(x_{2i}^h) = 1.$$

The solution of $\phi_i^H(x)$ on $[x_{2i-2}^h, x_{2i}^h]$ defines $\phi_i^H(x_{2i-1}^h)$ implicitly. The homogeneous PDE looks for a "harmonic" function ϕ_i^H which essentially minimizes the energy on its support. If $a(x) \equiv 1$, for instance, $\phi_i^H(x_{2i-1}^h) = \phi_i^H(x_{2i+1}^h) = 1/2$, and we will reproduce the linear interpolation. In this case, ϕ_i^H is harmonic in the usual sense and it has minimum energy. In general, we have

$$\phi_i^H(x_{2i-1}^h) = -\frac{a(\phi_{2i-1}^h, \phi_{2i}^h)}{a(\phi_{2i-1}^h, \phi_{2i-1}^h)} = -\frac{\mathcal{A}_{2i-1,2i}^h}{\mathcal{A}_{2i-1,2i-1}^h}, \tag{3}$$

where $(\mathcal{A}_{ij}^h)$ is the stiffness matrix.

2.2. A Convergence Analysis. Classical convergence analysis cannot be applied directly to this multigrid algorithm since all the coarse grids basis functions, except on the finest grid, are not standard finite element basis which are usually assumed in the theory. Thus the following analysis is worthwhile in that finite element basis may not be necessary for optimal convergence, which provides the basic ground for the construction in two dimensions. The analysis is based on the subspace correction framework [37], where convergence analysis is possible by estimating two algebraic quantities K_0 and K_1 defined later.

We first set up some notations. Let $V = V^h$ and $V_1 \subset V_2 \subset \cdots \subset V_J = V$ be a sequence of nested subspaces of V defined by the span of $\{\phi_i^k\}_{i=1}^{n_k}, k = 1, \ldots, J$, which are the nodal basis functions at level k. The operator $A : V \to V$ is defined by: $(Au, v) = a(u, v), u, v \in V$ and the A-inner product by: $(\cdot, \cdot)_A \equiv (A\cdot, \cdot)$. Also, we define $A_i : V_i \to V_i$ by $(A_i u_i, v_i) = (Au_i, v_i), u_i, v_i \in V_i$, which is a restriction operator of A on V_i. Correspondingly, we have $R_i : V_i \to V_i$, which is an approximate inverse of A_i. In the following analysis, the generic constant C is independent of the mesh size h.

Now we define the parameters K_0 and K_1 as follows:

K_0: For any $v \in V$, there exists a decomposition $v = \sum_{i=1}^J v_i$ for $v_i \in V_i$ such that

$$\sum_{i=1}^J (R_i^{-1} v_i, v_i) \le K_0 (Av, v),$$

where R_i is usually known as the smoother in the multigrid context.

K_1: For any S *subset*$\{1, \ldots, J\} \times \{1, \ldots, J\}$ and $u_i, v_i \in V$ for $i = 1, \ldots, J$,

$$\sum_{(i,j)\in S} (T_i u_i, T_j u_j) \le K_1 (\sum_{i=1}^J (T_i u_i, u_i)_A)^{\frac{1}{2}} (\sum_{j=1}^J (T_j v_j, v_j)_A)^{\frac{1}{2}},$$

where $T_i = R_i A_i P_i$ and $P_i : V \to V_i$ is the A-orthogonal projection onto V_i.

These two quantities give the lower and upper bound for the spectrum of the preconditioned system respectively. Because of limited space, we shall give an optimal estimate for K_0 only which is enough to illustrate the concepts of stability and the approximation property.

2.2.1. Stability. In [35], we observe that the coarse grid basis functions contain a hierarchy of A-orthogonal basis functions. We start from the following two level result.

LEMMA 2.1. *For any $k = 2, \ldots, J$, $i = 1, \ldots, n_{k-1}$, $j = 1, \ldots, n_k/2$, we have $a(\phi_i^{k-1}, \phi_{2j-1}^k) = 0$.*

Proof. It is the direct consequence of the fact that the equivalent variational formulation of (2) implies $a(\phi_i^{k-1}, \phi_{2i-1}^k) = 0$, and the support of ϕ_i^{k-1} is only on $[x_{2i-2}^k, x_{2i+2}^k]$. □

LEMMA 2.2. *For any $l < k, i = 1, \ldots, n_l, j = 1, \ldots, n_k/2$, we have $a(\phi_i^l, \phi_{2j-1}^k) = 0$.*

Proof. The case l=$k-1$ is proved in Lemma 2.1. Suppose it is true for $l = \bar{k}$. By definition, $\phi_i^{\bar{k}-1} = \sum_{j=2i-1}^{2i+1} \alpha_j \phi_j^{\bar{k}}$, where $\alpha_{2i} = 1$, $\alpha_{2i-1} = \phi_i^{\bar{k}-1}(x_{2i-1}^{\bar{k}})$ and $\alpha_{2i+1} = \phi_i^{\bar{k}-1}(x_{2i+1}^{\bar{k}})$ are given by (3). Thus

$$a(\phi_i^{\bar{k}-1}, \phi_{2j-1}^k) = \alpha_{2i-1} a(\phi_{2i-1}^{\bar{k}}, \phi_{2j-1}^k) + a(\phi_{2i}^{\bar{k}}, \phi_{2j-1}^k) + \alpha_{2i+1} a(\phi_{2i+1}^{\bar{k}}, \phi_{2j-1}^k) = 0,$$

since all the terms vanish by assumption and the result follows from induction. □

Lemma 2.2 implies that the interpolation algorithm generates implicitly a set of A-orthogonal hierarchical basis functions. In some sense, this is a generalization of the hierarchical basis method [38, 3, 23, 2]. The orthogonality property immediately implies the *stability* of the nested subspaces.

LEMMA 2.3. *For any $v \in V$, there is a nontrivial decomposition $v = \sum_{i=1}^{J} v_i$ with $v_i \in V_i$ such that*

$$\sum_{i=1}^{J} (v_i, v_i)_A = (v, v)_A. \tag{4}$$

Proof. For any $v \in V$, Lemma 2.2 implies that there exists an orthogonal hierarchical decomposition of v constructed as follows: 1) $v_1 = \sum_{i=1}^{n_1} v(x_i^1)\phi_i^1$. 2) $v_k = \sum_{i=1}^{n_k} w_k(x_i^k)\phi_i^k, k = 2, \ldots, J$, where $w_k = v - \sum_{i=1}^{k-1} v_i$. By construction, we have $v_k(x_j^k) = w_k(x_j^k) = 0, j =$ even. Therefore, by Lemma 2.2, the v_k's are A-orthogonal since if $l < k$,

$$a(v_l, v_k) = a(\sum_{i=1}^{n_l} w_l(x_i^l)\phi_i^l, \sum_{j=1}^{n_k} w_k(x_j^k)\phi_j^k) = \sum_{i=1}^{n_l}\sum_{j=1}^{n_k/2} w_l(x_i^l) w_k(x_{2j-1}^k) a(\phi_i^l, \phi_{2j-1}^k) = 0.$$

The equality (4) follows immediately from the orthogonality of v_k's. □

COROLLARY 2.4. *Let $W_1 = V_1, W_i = V_i \ominus V_{i-1}, i = 2, \ldots, J$, in the A-inner product. Then V can be expressed as a direct sum of W_i's: $V = W_1 \oplus W_2 \oplus \cdots \oplus W_J$.*

Proof. Let $v \in V$. By the construction given in the proof of Lemma 2.3, we have $v = \sum_{i=1}^{J} v_i$, $v_i \in W_i$. The representation is unique since v_i's are A-orthogonal. □

Corollary 2.4 induces a projection operator $\tilde{Q}_i : V \to V_i$ defined by: $\tilde{Q}_i v = v_1 + v_2 + \cdots + v_i$, where $v = v_1 + \cdots + v_J, v_i \in W_i$, is the unique representation of v.

2.2.2. Approximation. In order to prove optimal convergence, we need the following *approximation property*:

$$\|(\tilde{Q}_k - \tilde{Q}_{k-1})v\| \le C h_k \|\tilde{Q}_k v\|_A, \qquad k = 2, \ldots, J, \tag{5}$$

where $\tilde{Q}_k$ is defined previously. From the classical analysis, the approximation property (5) is ultimately related to preserving constant functions.

LEMMA 2.5. *For any $k = 1, \ldots, J$, $\sum_{i=1}^{n_k} \phi_i^k(x) = 1$.*

Proof. The case $k = J$ is trivial. For $k < J$, consider the interval $[x_{2i-1}^k, x_{2i}^k]$. Only ϕ_i^{k-1} and ϕ_{i-1}^{k-1} are nonzero. Let $\psi^{k-1} = \phi_i^{k-1} + \phi_{i-1}^{k-1}$. By (2), ψ^{k-1} satisfies the following equations:

$$-\frac{d}{dx}a(x)\frac{d}{dx}\psi_i^{k-1}(x) = 0 \qquad \text{in } [x_{2i-2}^k, x_{2i}^k],$$
$$\psi^{k-1}(x_{2i-2}^k) = 1, \quad \psi^{k-1}(x_{2i}^k) = 1.$$

By uniqueness, $\psi^{k-1} \equiv 1$ on $[x_{2i-1}^k, x_{2i}^k]$ and hence the result. □

LEMMA 2.6. *For any $v \in V$ and any $k = 2, \ldots, J$, we have $\|(\tilde{Q}_k - \tilde{Q}_{k-1})v\| \le C h_k \|\tilde{Q}_k v\|_A$.*

Proof. Since we can always change v to $\tilde{Q}_k v$, we assume $v \in V_k$, i.e. $v = \sum_{i=1}^{n_k} \nu_i \phi_i^k$. By definition, $\tilde{Q}_{k-1} v = \sum_{i=1}^{n_{k-1}} \nu_{2i}\phi_i^{k-1}$. Let $w = v - \tilde{Q}_{k-1}v = \sum_{i=1}^{n_k} \omega_i \phi_i^k$. Then $\omega_{2i} = 0$ and $\omega_{2i-1} = \nu_{2i-1} - (\alpha\nu_{2i-2} + \beta\nu_{2i}), i = 1, \ldots, n_k/2$, where $\alpha = \phi_{i-1}^{k-1}(x_{2i-1}^k)$ and $\beta = \phi_i^{k-1}(x_{2i-1}^k)$. By Lemma 2.5, $\alpha + \beta = 1$ and hence $\omega_{2i-1} = \alpha(\nu_{2i-1} - \nu_{2i-2}) - \beta(\nu_{2i} - \nu_{2i-1})$. Now we estimate the L^2 norm of w on $[x_{2i-2}^k, x_{2i}^k]$:

$$\int_{x_{2i-2}^k}^{x_{2i}^k} w^2 dx = \int_{x_{2i-2}^k}^{x_{2i}^k} (\omega_{2i-1}\phi_{2i-1}^k)^2 dx = \omega_{2i-1}^2 \int_{x_{2i-2}^k}^{x_{2i}^k} (\phi_{2i-1}^k)^2 dx = \omega_{2i-1}^2 \mathcal{M}_{2i-1,2i-1}^k,$$

where $\mathcal{M}_{j,j}^k$ is the (j,j) entry of $\mathcal{M}_k$ (the mass matrix with respect to $\{\phi_i^k\}_{i=1}^{n_k}$). By the formula of ω_{2i-1} and the Cauchy-Schwarz inequality,

$$\int_{x_{2i-2}^k}^{x_{2i}^k} w^2 dx \le \mathcal{M}_{2i-1,2i-1}^k [\alpha(\nu_{2i-1} - \nu_{2i-2})^2 + \beta(\nu_{2i} - \nu_{2i-1})^2]. \tag{6}$$

By a direct computation, the A-norm of v is given by,

$$\int_{x_{2i-2}^k}^{x_{2i}^k} a(x)(v')^2 dx = [\alpha(\nu_{2i-1} - \nu_{2i-2})^2 + \beta(\nu_{2i} - \nu_{2i-1})^2]\mathcal{A}_{2i-1,2i-1}^k,$$

where $\mathcal{A}_{j,j}^k$ is the (j,j) entry of $\mathcal{A}_k$ (the stiffness matrix with respect to $\{\phi_i^k\}_{i=1}^{n_k}$). Combining with (6),

$$\int_{x_{2i-2}^k}^{x_{2i}^k} w^2 dx \le \frac{\mathcal{M}_{2i-1,2i-1}^k}{\mathcal{A}_{2i-1,2i-1}^k} \int_{x_{2i-2}^k}^{x_{2i}^k} a(x)(v')^2 dx \le C h_k^2 \int_{x_{2i-2}^k}^{x_{2i}^k} a(x)(v')^2 dx, \tag{7}$$

since $\mathcal{M}_{j,j}^k = O(h_k)$ and $\mathcal{A}_{j,j}^k = O(h_k^{-1})$. Summing over i in (7), the result follows. □

COROLLARY 2.7. *For any $v \in V$ and any $k = 2, \ldots, J$, $\|(\tilde{Q}_k - \tilde{Q}_{k-1})v\| \le C h_k \|(\tilde{Q}_k - \tilde{Q}_{k-1})v\|_A$.*

Proof. Let $\tilde{v} = (\tilde{Q}_k - \tilde{Q}_{k-1})v$ and apply Lemma 2.6. □

Now we are in a position to put all the results together and give an optimal estimate for K_0.

THEOREM 2.8. *Let $\omega_0 = \min_{2 \le i \le J}(\rho(A_i)\lambda_{\min}(R_i))$. Then $K_0 \le \frac{C}{\omega_0}$.*

Proof. For any $v \in V$, let $v = \sum_{i=1}^J v_i$, $v_i = (\tilde{Q}_i - \tilde{Q}_{i-1})v$. By Lemma 2.3, we have

$$\|\tilde{Q}_1 v\|_A^2 + \sum_{i=2}^J \|(\tilde{Q}_i - \tilde{Q}_{i-1})v\|_A^2 = \|v\|_A^2. \tag{8}$$

In view of the definition of K_0, for each $i > 1$, we have

$$(R_i^{-1} v_i, v_i) \le \frac{1}{\lambda_{\min}(R_i)}(v_i, v_i) \le \frac{\rho(A_i)}{\omega_0}\|(\tilde{Q}_i - \tilde{Q}_{i-1})v\|^2 \le \frac{C}{\omega_0}\|(\tilde{Q}_i - \tilde{Q}_{i-1})v\|_A^2,$$

by Corollary 2.7 and $\rho(A_i) = O(h_i^{-2})$. For $i = 1$, since $R_1 = A_1^{-1}$, we get instead

$$(R_1^{-1} v_1, v_1) = ((R_1 A_1)^{-1} v_1, A_1 v_1) = \|v_1\|_A^2 = \|\tilde{Q}_1 v\|_A^2.$$

Combining with (8), $\sum_{i=1}^J (R_i^{-1} v_i, v_i) = \|\tilde{Q}_1 v\|_A^2 + \sum_{i=2}^J (R_i^{-1} v_i, v_i) \le \max(1, \frac{C}{\omega_0})\|v\|_A^2$. □

2.3. Implications. The analysis verifies that stability and the approximation property are useful characterizations for good subspaces. We rewrite them here again but in a slightly different format.

$$\|\tilde{Q}_1 v\|_A^2 + \sum_{k=2}^J \|(\tilde{Q}_k - \tilde{Q}_{k-1})v\|_A^2 \le C_0 \|v\|_A^2, \tag{9}$$

$$\|(\tilde{Q}_k - \tilde{Q}_{k-1})v\| \le C_1 h_k \|\tilde{Q}_k v\|_A, \qquad \forall k > 1. \tag{10}$$

Recall that the coarse grid basis functions determine uniquely the subsequent subspaces. The two inequalities are actually two requirements for the basis functions.

One more remark is that, as mentioned before, the approximation property (10) is usually implied by preserving constant functions. Sometimes it may be convenient to change (10) to: $\sum_{i=1}^{n_k} \phi_i^k(x) = 1, \forall k$. These two conditions are our guidelines for defining multigrid methods in higher dimensions.

3. Two Dimensions. To facilitate our understanding of the 2D case, we shall first focus on the standard structured triangular grids on the square domain Ω: $[0,1] \times [0,1]$ and postpone the unstructured grids case to section 4.1. The 2D model problem is:

$$\begin{aligned} -\nabla \cdot a(x,y)\nabla u(x,y) &= f(x,y), \quad && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{11}$$

with the same assumptions on $a(x,y)$ and $f(x,y)$ as before.

3.1. Formulation of Energy-Minimizing Interpolation. The extension to two dimensions is difficult since there is no definite or natural analogue between the 1D and the 2D case. Several possibilities of setting up local PDE's exist, for instance, the stencil or the so-called blackbox approach [18, 1, 9, 16, 39, 34], Schur complement approach [25, 19, 13] and the algebraic multigrid approach [32], each of which mimics the 1D case in some way.

The analysis of the 1D case shows that stability and the approximation property are two basic requirements for the coarse grid basis functions. Thus, the basis functions should have small energy and

approximate fine grid functions with at least first order of accuracy. Suppose a maximal independent set M is chosen as coarse grid points. Let the coarse grid nodal basis function $\phi_i^H = \sum_{j\in\{A^h_{i,j}\neq 0\}\setminus M} \alpha_j^i \phi_j^h$. We define the interpolation by solving a constrained minimization problem for $\{\alpha_j^i\}$:

$$\min \ \frac{1}{2}\sum_{i=1}^{m} \|\phi_i^H\|_A^2 \qquad \text{subject to} \qquad \sum_{i=1}^{m} \phi_i^H(x) = 1 \quad \text{in } \Omega, \tag{12}$$

where $m = (n/2+1)^2, h = 1/n$. Notice that the coarse grid nodal basis functions on the boundary with Dirichlet condition are also constructed so that the constraint makes sense.

The solution method for the minimization problem is briefly explained as follows. For each i, $\phi_i^H = \sum_{j=1}^{n} \varphi_j^i \phi_j^h$ and $\varphi^i = (\varphi_1^i \cdots \varphi_n^i)^T$ is a sparse vector with at most seven nonzeros. Let $\Phi = [\varphi^1; \cdots; \varphi^m]$ be an $mn \times 1$ vector appending all the φ^i's. Noting that $\|\phi_i^H\|_A^2 = (\varphi^i)^T \tilde{A}^h \varphi_i$, where $\tilde{A}^h$ is the assembled stiffness matrix on the fine grid without boundary conditions, (12) can be written as the following equivalent linear constrained quadratic minimization problem:

$$\min_{\Phi} \ \frac{1}{2}\Phi^T \mathcal{Q} \Phi \qquad \text{subject to} \qquad \mathcal{B}^T \Phi = \mathbf{1}. \tag{13}$$

The symbol $\mathbf{1}$ is understood as a vector of all 1's. The $mn \times mn$ SPD matrix $\mathcal{Q}$ is block diagonal with each block equals $\tilde{\mathcal{A}}_i^h$ where

$$(\tilde{\mathcal{A}}_i^h)_{kl} = \begin{cases} \tilde{\mathcal{A}}^h_{kl} & \text{if } \varphi_k^i \neq 0 \text{ and } \varphi_l^i \neq 0 \\ \delta_{kl} & \text{otherwise.} \end{cases}$$

The $n \times mn$ rectangular matrix $\mathcal{B}^T = [\mathcal{J}_1^T \cdots \mathcal{J}_m^T]$, where $\mathcal{J}_i = \mathcal{J}_i^T$ is a near identity matrix corresponding to the restriction operator that maps v to v_i such that $v = v_i$ on $\text{supp}(\phi_i^H)$ and $v_i = 0$ otherwise. More precisely,

$$(\mathcal{J}_i)_{kl} = \begin{cases} 1 & \text{if } k = l \text{ and } \varphi_k^i \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that $\mathcal{J}_i^T \varphi^i = \varphi^i$ and hence $\mathcal{B}^T\Phi = \sum_{i=1}^m \mathcal{J}_i^T \varphi^i = \sum_{i=1}^m \varphi^i = \mathbf{1}$. By the Lagrange multiplier formulation, (13) is equivalent to:

$$\begin{bmatrix} \mathcal{Q} & \mathcal{B} \\ \mathcal{B}^T & 0 \end{bmatrix} \begin{bmatrix} \Phi \\ \Lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{14}$$

where Λ is an $n \times 1$ vector of Lagrange multipliers. After solving for Λ and substituting back into (14), the system is reduced to:

$$\mathcal{Q}\Phi = \mathcal{B}(\mathcal{B}^T \mathcal{Q}^{-1} \mathcal{B})^{-1} \mathbf{1}.$$

The solution process for Φ is divided into two steps: 1) Solve $\mathcal{B}^T\mathcal{Q}^{-1}\mathcal{B}g = \mathbf{1}$. 2) Solve $\mathcal{Q}\Phi = \mathcal{B}g$.

Step 1: Since we do not want to form the huge systems $\mathcal{Q}^{-1}$ and $\mathcal{B}^T\mathcal{Q}^{-1}\mathcal{B}$ explicitly, direct solve is not preferred. As $\mathcal{B}^T\mathcal{Q}^{-1}\mathcal{B}$ is SPD, conjugate gradient is used instead. The only nontrivial procedure is the matrix-vector multiply. Recall that $\mathcal{B}$ is a block of near identity matrices $\mathcal{J}_i$. There is no operation involved in forming $\mathcal{B}g = \bar{g} = [\bar{g}^1; \cdots; \bar{g}^m]$. By the definition of $\mathcal{J}_i$, the nonzeros of $\bar{g}$ equal those of φ^i. Thus forming $\tilde{g} = \mathcal{Q}^{-1}\bar{g}$ only involves converting a little 7×7 submatrix of $\tilde{\mathcal{A}}_i^h$ for each $\bar{g}^i$. Finally $\mathcal{B}^T\tilde{g} = \sum_{i=1}^m \tilde{g}^i$, $\tilde{g} = [\tilde{g}^1; \cdots; \tilde{g}^m]$. The overall computations only involve $O(n)$ operations.

It is interesting to note that if $\tilde{A}^h$ is definite, for instance, the equation in (11) has a positive zeroth order term, $\tilde{A}^h$ itself is a natural preconditioner for $\mathcal{B}^T\mathcal{Q}^{-1}\mathcal{B}$ since

$$\mathcal{B}^T\mathcal{Q}^{-1}\mathcal{B} = \sum_{i=0}^{n/2} \mathcal{J}_i^T (\tilde{\mathcal{A}}_i^h)^{-1} \mathcal{J}_i = \sum_{i=0}^{n/2} \mathcal{R}_i (\mathcal{R}_i^T \tilde{\mathcal{A}}^h \mathcal{R}_i)^{-1} \mathcal{R}_i^T,$$

where $\mathcal{R}_i^T$ is the submatrix of nonzero rows of $\mathcal{J}_i$, and it is now clear that $\mathcal{B}^T\mathcal{Q}^{-1}\mathcal{B}$ is the additive Schwarz preconditioner for $\tilde{A}^h$.

Step 2: It is similar to *Step 1* and is not discussed.

3.2. Special Features. The above formulation is consistent with the 1D case described in section 2.1 in several aspects:

1) In 1D, the constraint exists implicitly since it is automatically satisfied if $\{\phi_i^H\}$ solve the local PDE's (see Lemma 2.5).
2) Since no explicit constraint is needed in 1D, the minimization of $\sum_{i=0}^{n/2} \|\phi_i^H\|_A^2$ can be splitted into m independent problems each of which is equivalent to (2).
3) If $a(x,y) \equiv 1$, we almost reproduce the linear interpolation in the sense that numerical experiments show the interpolation values are close to 1/2. The minor discrepancy probably comes from the boundary where the interaction between minimizing energy and preserving constant functions is different from the interior.

This interpolation also possesses some other advantages that make it flexible and applicable to complicated problems:

1) Like algebraic multigrid, the construction of the interpolation operator is purely algebraic. In other words, geometry and in particular the grid information are not needed. Unlike algebraic multigrid, however, the hidden geometric information in the stiffness matrix is exploited by the constrained minimization problem.
2) The interpolation is still valid if the coarse grid points do not form an independent set. Independent sets are certainly beneficial to efficiency but nevertheless necessary. In some situations, we may want to remove this requirement. Further discussions on coarsening will be given in section 4.2.

3.3. A Two Level Analysis. Since our interpolation is implicitly defined by the minimization problem, analysis is tricky and has not yet been fully investigated. Nevertheless, a two level analysis, which is essentially due to Xu [35] [27], is presented. In this special case, only the preserving constant property is actually used. The role of the minimization is explained afterwards.

We shall verify the stability and the approximation properties (9), (10) which are equivalent to the following in the two level case.

LEMMA 3.1. *There exists an average interpolant I^H such that for any $v \in V^h \subset H_0^1(\Omega)$,*

$$|I^H v|_{1,\Omega} \leq C_0 |v|_{1,\Omega} \tag{15}$$

$$\|v - I^H v\|_{0,\Omega} \leq C_1 h |v|_{1,\Omega}, \tag{16}$$

where $\|\cdot\|_{0,\Omega}$, $\|\cdot\|_{1,\Omega}$ and $|\cdot|_{1,\Omega}$ are the usual L^2, H^1 and H^1 semi norms respectively.

Proof. For any coarse grid point x_j^H, let e_j be an edge that contains x_j^H. Let ψ_j be the linear function on e_j such that $(v,\psi_j)_{0,e_j} = v(x_j^H), \forall v \in P_1(e_j)$, where $P_1(e_j)$ is the set of piecewise linear functions on e_j. Define $I^H : V^h \to V^H$ by: $(I^H v)(x) = \sum_j (v,\psi_j)_{0,e_j} \phi_j^H(x)$.

We now prove (16). Let $S_i \subset \Omega$ denote the interior of the support of ϕ_i^H and $\mathcal{S}^h$ be the set of intersections of $\{S_i\}$. Then it is not hard to see that there exists $\mathcal{G}^h \subset \mathcal{S}^h$ such that G_j^h's $\in \mathcal{G}^h$ are disjoint and $\overline{\Omega} = \cup \overline{G_j^h}$. In the agglomeration multigrid context, G_j^h's are just agglomerated macro elements. Let $G_i^h \in \mathcal{G}^h$ and consider an auxiliary triangle K_i^h that contains G_i^h and all its neighboring elements G_j^h, i.e. $\overline{G_j^h} \cap \overline{G_i^h} \neq \phi$. Let $F_{K_i^h} : \hat{K} \to K_i^h$ be an affine mapping that maps the standard reference element $\hat{K}$ to K_i^h. Define $\widehat{G_i} = F_{K_i^h}^{-1}(G_i^h) \subset \hat{K}$; $\hat{v}(\hat{x}) = v(F_{K_i^h}(\hat{x})), \hat{x} \in \hat{K}$ and $\widehat{I^H}\hat{v} = \widehat{I^H v}$. (Note: v is extended by zero if K_i^h is outside Ω). By the trace theorem, it can be proved that $\|\hat{v} - \widehat{I^H}\hat{v}\|_{0,\widehat{G_i}} \leq C\|\hat{v}\|_{1,\hat{K}}$. Since constant functions are invariant under I^H, so are under $\widehat{I^H}$. Hence

$$\|\hat{v} - \widehat{I^H}\hat{v}\|_{0,\widehat{G_i}} = \inf_{\hat{c}\in\Re} \|\hat{v} + \hat{c} - \widehat{I^H}(\hat{v}+\hat{c})\|_{0,\widehat{G_i}} \leq C \inf_{\hat{c}\in\Re} \|\hat{v}+\hat{c}\|_{1,\hat{K}} \leq C|\hat{v}|_{1,\hat{K}}.$$

Transforming back to K_i^h, we have $\|v - I^H v\|_{0,G_i^h} \leq Ch|v|_{1,K_i^h}$. Thus

$$\|v - I^H v\|_{0,\Omega}^2 = \sum_i \|v - I^H v\|_{0,G_i^h}^2 \leq Ch^2 \sum_i |v|_{1,K_i^h}^2 \leq Ch^2 |v|_{1,\Omega}^2.$$

Inequality (15) can be proved similarly using the H^1 norm instead and hence is omitted. □

The role of the minimization can be revealed by the estimation of $\|I^H v\|_A$ in the case when (11) has a positive term $b(x,y)u$ and I^H is the nodal value interpolant:

$$\|I^H v\|_A = \|\sum_i v(x_i^H)\phi_i^H\|_A \le \sum_i \|\phi_i^H\|_A |v(x_i^H)| \le (\sum_i \|\phi_i^H\|_A^2)^{1/2} (\sum_i |v(x_i^H)|^2)^{1/2}$$
$$\le \frac{C}{h}(\sum_i \|\phi_i^H\|_A^2)^{1/2}\|v\|_2 \le \frac{C}{h \min b(x,y)}(\sum_i \|\phi_i^H\|_A^2)^{1/2}\|v\|_A.$$

Despite that the bound may not be sharp, the minimization problem essentially minimizes the constant in the stability inequality (9).

3.4. Numerical Results. We verify the optimal convergence behavior of our multigrid method by comparing it to the standard one when they are used to solve the Poisson equation on $[0,1] \times [0,1]$ with homogeneous Dirichlet boundary condition and $f(x) = 1$. Two pre and post Gauss-Seidel smoothings are used and the iteration is stopped when the relative residual norm is less than 10^{-10}. Table 1 shows the number of multigrid iterations to convergence with different number of levels (2-5). (Entries with no data are denoted by $*$). The previous analysis only predicts optimal convergence in two level but optimal results are indeed observed in the multilevel case as well.

grid	Linear				Energy-min.			
	2	3	4	5	2	3	4	5
16 × 16	9	10	*	*	9	11	*	*
32 × 32	9	10	11	*	9	10	11	*
64 × 64	9	10	11	11	9	10	11	12

TABLE 1

Convergence of multigrid using linear and energy-minimizing interpolations.

4. Applications. The special features of our interpolation lead to applications to unstructured grids computations and discontinuous coefficient problems.

4.1. Unstructured Grids Computation. Unstructured grids are irregular grids which are not obtained by successive refinement of coarser grids. Thus there is no natural choice of coarse meshes. The challenge of multigrid is to define interpolations on such grids. One approach is by retriangulation [14, 21]. A maximal independent set for the coarse grid points is determined and the coarse mesh is formed by a retriangulation of the coarse grid points. The linear interpolation is defined naturally on the retriangulated grid. Another approach is by agglomeration [20, 33]. Groups of nearby elements are agglomerated together to form bigger macro elements. The piecewise constant interpolation is usually used but nevertheless higher order interpolations are also possible by extra costs.

The former approach has a nice theory to govern optimal convergence. However it suffers from the geometric dependence of the algorithm. For complicated geometries, special care [7, 8] is needed to handle different types of boundary conditions correctly. Besides, grid information is required for future retriangulation and interpolation. Moreover, its application to 3D problems is difficult since tetrahedralization of arbitrary points in space is nontrivial. The latter approach, on the other hand, is algebraic if injection is used for the interpolation and hence it is applicable to 2D and 3D problems. However, mesh information is still necessary if higher order interpolations, which have been noted in the literature [34] necessary for optimal convergence, are used.

Our multigrid method can be viewed as a kind of agglomeration multigrid. The elements on the fine grid are implicitly agglomerated together through the construction of the coarse grid basis functions. The constraint in the minimization problem ensures our interpolation is at least first order accurate. If $a(x,y) \equiv 1$, the optimal linear interpolation is almost recovered (cf. section 3.2) in the structured grids case. Thus close to linear interpolation is expected on the unstructured grids as verified by the numerical results shown later. Furthermore, since the algorithm is purely algebraic, the previously mentioned geometric difficulties disappear.

4.1.1. Supplements. However, the interpolation described in section 3.1 is incomplete for the unstructured grids case and a further discussion is needed.

Supplement 1: Unlike the structured grids case where the noncoarse grid points are connected to exactly two coarse grid points, it is possible that some noncoarse grid points are attached to one coarse

grid point only. The constant preserving constraint enforces the values at those points the same as the corresponding coarse grid points. In some sense, it is like piecewise constant interpolation which is not sufficient for optimal convergence. Our solution is simply to change those noncoarse grid points to coarse grid points. Independence among coarse grid points are thus broken which, however, does not affect the convergence nor the construction of the interpolation (cf. section 3.2). Since usually only a few number of such points exists, the overall complexity increases only by a bit.

Supplement 2: Another difficulty comes from the boundary. In contrast to the structured grids case, it is possible that a noncoarse grid point on the boundary with homogeneous Dirichlet condition is connected to an interior coarse grid point. The interpolated value at the noncoarse grid point, in general, is nonzero which is inconsistent with the boundary condition. Our solution is that interior coarse grid points do not interpolate boundary points with Dirichlet condition.

4.1.2. Numerical Results. We demonstrate our multigrid algorithm on unstructured meshes by solving the Poisson and a variable coefficient problems on the well-known NASA airfoil [8] with 4253 unknowns. The Poisson problem is associated with the homogeneous Dirichlet boundary condition and the variable coefficient problem:

$$\frac{\partial}{\partial x}((1+xy)\frac{\partial u}{\partial x}) + \frac{\partial}{\partial y}(\sin(3y)\frac{\partial u}{\partial y}) = (4xy+2)\sin(3y) + 9x^2\cos(6y),$$

with a mixed boundary condition: $u = 2 + x^2\sin(3y), x \leq 0.2$ and $\partial u/\partial n = 0, x > 0.2$. The stiffness matrix is solved by the conjugate gradient method. One V-cycle multigrid with two pre and post Gauss-Seidel smoothings is used as a preconditioner. For comparison purpose, the initial guess is zero and the iteration is stopped when the relative residual norm is less than 10^{-6} and 10^{-5} for the Poisson and the variable coefficient problem respectively.

	Energy-min. MG			Retriangulation MG		
MG	number of coarse	number of iter.		number of coarse	number of iter.	
levels	grid points	Poisson	Var. Coeff.	grid points	Poisson	Var. Coeff.
2	1285	6	6	1170	5	4
3	373	6	6	340	5	5
4	74	6	7	101	5	5

TABLE 2
Convergence of multigrid preconditioners on the airfoil.

The results are shown in Table 2 where the data of the retriangulation multigrid with special treatments on the Neumann boundary condition is obtained from [7, 8] (note: GMRES were used). Both methods have roughly the same number of iterations which verify that our method has optimal convergence behavior. Although our multigrid method takes more coarse grid points initially due to *Supplement 1*, it takes about the same number as the other one eventually.

4.2. Discontinuous Coefficient Problems. The standard multigrid converges slowly when the coefficient $a(x, y)$ is piecewise constant with jumps differed by several order of magnitudes. The finite element discretization is tricky in itself. Despite that, suppose the discretization is properly done. We want to derive a multigrid method which is as insensitive to the jump as possible.

The focus of the tackling strategy in the literature was to define an interpolation so that the discontinuous behavior of the derivative of the solution along the interface is correctly handled. The motivation is probably due to the effective interpolation described in section 2.1. In fact, this approach is effective for regular interface problems, for example, the interface forms a square, on structured squared grids. Not much attention has been paid to irregular interface problems which are naturally arised in unstructured grids and in other problems. We shall consider irregular interface problems on structured triangular grids and introduce the coarsening aspect for multigrid methods. The unstructured grid case is treated similarly.

As mentioned before, a successful multigrid method depends not just on the interpolation but all the procedures as a whole. Our key observation is that coarsening is crucial for interface problems. Specifically, the coarse grid points should resolve the shape of the interface in some sense described

later. Intuitively speaking, experiences reported in the literature [1] revealed that the parts of the solution on regions of different constant coefficients behave independently and connected together through a Neumann boundary condition on the interface. Theoretically speaking, convergence proofs for interface problems [36] require the discontinuities lie on all coarser meshes. In view of these, we present a coarsening algorithm for discontinuous coefficient problems.

Remarks: 1) Special coarsening techniques are quite common for anisotropic problems. To the author's knowledge, however, no coarsening strategy specific to discontinuous coefficient problems has been discussed in the literature. 2) It is interesting to notice that unstructured grids multigrid has to be used after one level even though the original grid is structured, if the following special coarsening is used to resolve the interface.

4.2.1. A Coarsening Algorithm. Let Ω^+, Ω^- be disjoint open subsets of Ω such that $\Omega = \overline{\Omega^+} \cup \overline{\Omega^-}$. Let $\Gamma = \partial\Omega^+$ be the interface and $\Gamma \cap \partial\Omega = \phi$. Let $a(x,y) \equiv a^+$ in Ω^+ and $a(x,y) \equiv a^-$ in Ω^- and $a^- \leq a^+$. The case of multiple jumps is treated similarly and hence is omitted.

Algorithm

1) Determine the set of fine grid points N^+ in $\overline{\Omega^+}$.
2) Full coarsening on N^+.
3) Full coarsening on $N^- = N \backslash N^+$, N=set of fine grid points.
4) Either (a) or (b) is used. Change a noncoarse grid point x_i^h to a coarse grid point, if
   (a) i) $x_i^h \in N^+$ and $x_j^H \in N^-$ for any coarse grid point x_j^H connected to x_i^h; else if
   ii) $x_i^h \in N^-$ and $x_j^H \in N^+$ for any coarse grid point x_j^H connected to x_i^h; else if
   iii) $x_i^h \in N^+$ and more than one coarse grid point $x_j^H \in N^- \backslash \partial\Omega$ is connected to x_i^h; else if
   iv) $x_i^h \in N^-$ and more than one coarse grid point $x_j^H \in N^+$ is connected to x_i^h; else if
   v) $x_i^h \in \partial\Omega$ and no $x_j^H \in \Omega$ is connected to x_i^h.
   (b) i) $x_i^h \in N^+$ and less than two coarse grid points $x_j^H \in N^+$ connected to x_i^h; else if
   ii) $x_i^h \in N^-$ and less than two coarse grid points $x_j^H \in N^-$ connected to x_i^h; else if
   iii) $x_i^h \in \partial\Omega$ and no $x_j^H \in \Omega$ is connected to x_i^h.
   In addition, coarse grid points in N^+ do not interpolate those in N^- and vice versa.

Step 1: Notice that $\tilde{A}_{ii}^h = a^+ \int_\Omega |\nabla \phi_i^h|^2 dx$ for any $x_i^h \in N^+$. If $a^+ \gg a^-$, the set of grid points in N^+ can be easily determined by the large diagonal entries of $\tilde{A}_{ii}^h$.
Step 2, 3: They are just standard coarsenings with points in N^+ first followed by those in N^-.
Step 4: Since there is no definite choice how the interpolation is done along the interface, two possible sets of criteria are suggested to ensure that the noncoarse grid points on Γ are properly interpolated so that the discontinuous derivative behavior of the solution is captured.

4.2.2. Complexity Issue. Resolving the interface increases the total number of coarse grid points which in turn increases the overall complexity. However, if the interface forms a simple piecewise smooth curve, the increase in coarse grid points is at most the total number of grid points on the interface, which is only $O(\frac{1}{h})$ compared to $O(\frac{1}{h^2})$ for the total number of coarse grid points. The complexity, η_0, of one V-cycle multigrid with standard coarsening is estimated to be:

$$\eta_0 = O(\frac{1}{h^2} + \frac{1}{4h^2} + \frac{1}{16h^2} + \ldots) = O(\frac{4}{3h^2}),$$

and the complexity, η_1, of one V-cycle multigrid with our special coarsening is:

$$\eta_1 = O(\tfrac{1}{h^2} + (\tfrac{1}{4h^2} + \tfrac{1}{h}) + (\tfrac{1}{4}(\tfrac{1}{4h^2} + \tfrac{1}{h}) + \tfrac{1}{2h}) + (\tfrac{1}{4}(\tfrac{1}{4}(\tfrac{1}{4h^2} + \tfrac{1}{h}) + \tfrac{1}{2h}) + \tfrac{1}{4h}) + \ldots) = O(\tfrac{4}{3h^2} + \tfrac{8}{3h}).$$

Thus the increase in complexity is asymptotically small compared to the standard one.

4.2.3. Numerical Results. We demonstrate the importance of coarsening by two examples. The number $a^- = 1$ and a^+ is varied from 10 to 10^4. The iteration stops when the relative residual norm is less than 10^{-6}. Other settings are the same as in section 3.4.

Example 1: A square interface on an 8×8 grid as shown in Fig.1(a), with Ω^+=shaded region, is enough to show the importance of coarsening. For illustration purpose, only two levels are used and the set of coarse grid points (denoted by $\circ$) is manually selected. Also, in addition to criteria set 4(a),

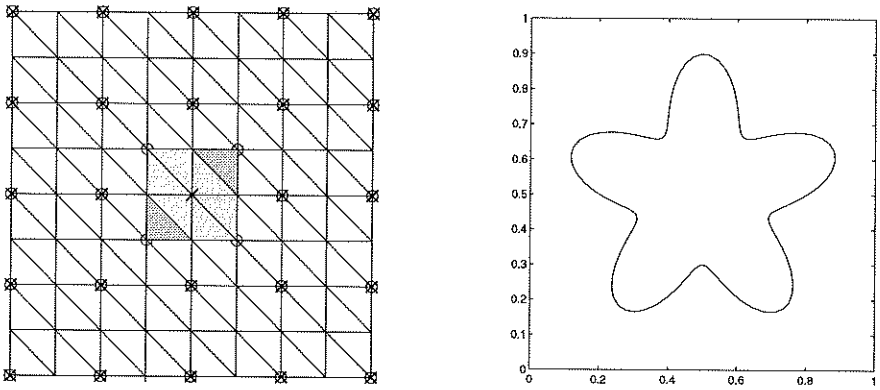

FIG. 1. *(a) Coarse grid points of different coarsenings (b) The interface is formed by:* $(x, y), x = 0.5 + r\cos\theta, y = 0.5 + r\sin\theta$ *where* $r = 0.3 + 0.1\sin(5\theta), 0 \leq \theta \leq 2\pi$.

coarse grid points only interpolate the points in the same region of coefficients. Table 3 shows the number of multigrid iterations to convergence using our special and the standard coarsenings together with energy-minimizing and linear interpolations respectively. The standard multigrid converges slowly with the increasing size of the jump because the standard coarse grid points (denoted by $\times$), never approximate the two darker corners well, if $a^+ \gg a^-$. On the other hand, an addition of four coarse grid points at the corners improves the convergence dramatically.

a^+	10	10^2	10^3	10^4
Special Coarsening	6	6	6	6
Standard Coarsening	12	21	23	23

TABLE 3

Convergence of multigrid methods with different coarsenings.

Example 2: A more complicated irregular interface shown in Fig. 1 (b) is used. The results in Table 4 shows that the convergence is independent of the mesh size and the jump but probably depends on the number of levels. Same settings have been repeated for the linear and the interpolation described in [9]. They both take more than 1000 iterations to converge with convergence rate around 0.99.

Grid size	Level	# of nodes	$a^+ = 10$	$a^+ = 10^2$	$a^+ = 10^3$	$a^+ = 10^4$
1/32	2	344	12	13	14	14
	3	116	14	14	15	15
	4	40	20	22	22	22
1/64	2	1261	13	13	13	13
	3	367	14	14	14	14
	4	102	18	18	18	18
	5	28	30	37	39	39
1/128	3	1258	16	16	16	16
	4	319	19	19	19	19
	5	86	27	28	28	28
	6	25	43	52	54	54

TABLE 4

Convergence of the special coarsening multigrid with different sizes of grid and jump.

5. Conclusions. From the energy-minimizing interpretation of the 1D interpolation, we have derived an analogous 2D interpolation which preserves constant and in a way minimizes the energy of decompositions of functions. The resulting multigrid methods has been shown theoretically pleasing and numerically effective. Its purely algebraic implementation makes it particularly attractive to unstructured grids computations where geometric complications do not raise a problem.

We have also addressed the issue of coarsening for multigrid methods applied specifically to discontinuous coefficient problems. The slow convergence of standard multigrid may be due to the poor choice of coarse grid points. In fact, we have presented a heuristic coarsening algorithm which has shown to be effective for this kind of problems.

Acknowledgments. The author would like to thank Barry Smith for his introducing the concept of energy minimization and Tony Chan for his guidance throughout the paper.

REFERENCES

[1] R.E. Alcouffe, A. Brandt, J.E. Dendy Jr., and J.W. Painter. The multi-grid method for the diffusion equation with strongly discontinuous coefficients. *SIAM J. Sci. Stat. Comput.*, 2(4):430–454, 1981.

[2] R.E. Bank. Hierarchical bases and the finite element method. *Acta Numerica*, pages 1–43, 1996.

[3] R.E. Bank, T.F. Dupont, and H. Yserentant. The hierarchical basis multigrid method. *Numer. Math.*, pages 427–458, 1988.

[4] J. Bramble. *Multigrid Methods.* Longman Scientific & Technical, Essex, UK, 1993.

[5] J. Bramble, J. Pasciak, J. Wang, and J. Xu. Convergence estimates for multigrid algorithms without regularity assumptions. *Math. Comp.*, 57:23–45, 1991.

[6] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.*, 31:333–390, 1977.

[7] T.F. Chan, S. Go, and J. Zou. Multilevel domain decomposition and multigrid methods for unstructured meshes: Algorithms and theory. Technical Report CAM 95-24, Department of Mathematics, UCLA, 1995.

[8] T.F. Chan, S. Go, and J. Zou. Boundary treatments for multilevel methods on unstructured meshes. Technical Report CAM 96-30, Department of Mathematics, UCLA, 1996.

[9] J.E. Dendy Jr. Black box multigrid. *J. Comp. Phys.*, 48:366–386, 1982.

[10] J.E. Dendy Jr., M.P. Ida, and J.M. Rutledge. A semi-coarsening multigrid algorithm for SIMD machines. *SIAM J. Sci Stat. Comput.*, 13:1460–1469, 1992.

[11] J.E. Dendy Jr. and C.C. Tazartes. Grandchild of the frequency decomposition multigrid method. *SIAM J. Sci. Comput.*, 16(2):307–319, 1994.

[12] B. Engquist and E. Luo. Multigrid methods for differential equations with highly oscillatory coefficients. In *Proceedings of the Sixth Copper Mountain Conference on Multigrid Methods*, pages 175–190, 1993.

[13] J. Fuhrmann. A modular algebraic multilevel method. In *Virtual Proceedings of the Ninth International GAMM-Workshop on Parallel Multigrid Methods*, 1996.

[14] H. Guillard. Node-nested multigrid method with delaunay coarsening. Technical Report 1898, INRIA, Sophia Antipolis, France, 1993.

[15] W. Hackbusch. Ein iteratives Verfahren zur schnellen Auflosung elliptischer Randwert probleme. Technical Report 76-12, Universitat zu Koln, 1976.

[16] W. Hackbusch. *Multi-grid Methods and Applications.* Springer-Verlag, Berlin, Heidelberg, 1985.

[17] W. Hackbusch. The frequency decomposition multigrid method, part I: Application to anisotropic equations. *Numer. Math.*, 56:229–245, 1989.

[18] R. Kettler and J.A. Meijerink. A multigrid method and a combined multigrid-conjugate gradient method for elliptic problems with strongly discontinuous coefficients in general domains. Shell Publ 604, KSEPL, Rijswijk, 1981.

[19] S. Knapek. Matrix-dependent multigrid-homogenization for diffusion problems. In *1996 Copper Mountain Conference on Iterative Methods*, 1996.

[20] M.H. Lallemand, H. Steve, and A. Dervieux. Unstructured multigrdding by volume agglomeration: Current status. *Computers and Fluids*, pages 397–433, 1992.

[21] D.J. Mavriplis. Multigrid techniques for unstructured meshes. Technical Report 95-27, ICASE, NASA Langley, USA, 1995.

[22] S. McCormick. *Multigrid Methods.* SIAM, Philadelphia, Penn, 1988.

[23] E. Ong. *Hierarchical Basis Preconditioners for Second Order Elliptic problems in Three Dimen-*

sions. PhD thesis, University of Washington, Seattle, 1989.

[24] A. Reusken. Multigrid with matrix-dependent transfer operators for a singular perturbation problem. *Computing*, 50:199–211, 1993.

[25] A. Reusken. Multigrid with matrix-dependent transfer operators for convection- diffusion problems. In *Multigrid Methods, Vol. IV, Springer Lectures in Math.* Springer-Verlag, 1993.

[26] J.W. Ruge and K. Stuben. Algebraic multigrid. In S. McCormick, editor, *Multigrid Methods*, pages 73–130. SIAM, 1987.

[27] L.R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54:483–493, 1990.

[28] B. Smith. Private communication.

[29] B. Smith, P. Bjorstad, and W. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, Cambridge, 1996.

[30] R.A. Smith and A. Weiser. Semicoarsening multigrid on a hypercube. *SIAM J. Sci Stat. Comput.*, 13:1314–1329, 1992.

[31] S. Ta'asan. *Multigrid Methods for Highly Oscillatory Problems*. PhD thesis, Weizmann Institute of Science, Rehovat, Israel, 1984.

[32] P. Vanek, J. Mandel, and M. Brezina. Algebraic multigrid on unstructured mesh. Submitted to SIAM J. Sci. Comput.

[33] V. Venkatakrishnan and D.J. Mavriplis. Agglomeration multigrid for the three-dimensional euler equations. Technical Report 94-5, ICASE, NASA Langley, USA, 1994.

[34] P. Wesseling. *An Introduction to Multigrid Methods*. Wiley, Chichester, 1992.

[35] J. Xu. Private communication.

[36] J. Xu. *Theory of Multilevel Methods*. PhD thesis, Cornell University, 1989.

[37] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.

[38] H. Yserentant. On the multilevel splitting of finite element spaces. *Numer. Math.*, pages 379–412, 1986.

[39] P.M. De Zeeuw. Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *J. Comp. Appl. Math.*, pages 1–27, 1990.