ALGEBRAIC MULTIGRID ON UNSTRUCTURED MESHES*

PETR VANĚK , JAN MANDEL , AND MARIAN BREZINA†

December 1994

Abstract. An algebraic multigrid algorithm is developed based on prolongations by smoothed aggregation. Coarse levels are generated automatically. Heuristic principles to guide the choice of the coarseninng are introduced. Almost optimal convergence bounds are proved for uniformly elliptic problems, shape regular triangulations on the finest level, and a general class of coarse problem hierarchy. Coarsening is by the factor of about three, which guarantees low complexity and bounded energy of the coarse shape functions. Numerical experiments confirm the theory and demonstrate that the method performs well also on a general class of problems with highly variable coefficients and strong anisotropies.

Key words. Algebraic multigrid, unstructured meshes, autoamtic coarsening, smoothed aggregations.

AMS(MOS) subject classifications. 65F10,65N55

1. Introduction. Multigrid methods are fast iterative methods for the solution of large sparse systems arising from discretization of partial differential equations. The main idea of multigrid methods is to complement the local exchange of information in point-wise iterative methods by utilizing several related systems, called *coarse levels,* with a smaller number of variables. The coarse levels interact with the original system to propagate information globally. Multigrid methods are very efficient iterative solvers for elliptic problems if the iterative method can control the discretization so that the coarse levels are naturally obtained as a hierarchy of discretizations with different characteristic meshsizes. In fact, generally available multigrid packages are either tightly integrated with the finite element engine such as PLTMG by Bank [2], or they use uniform grids such as Dendy's black box multigrid code [1, 9]. In practice, however, one is often required to solve linear systems produced by existing finite element software. Then the multigrid method must take the system of equations *as is* and take advantage of the functional analysis of the problem only indirectly. We need to construct an *automatic coarsening* that produces an artificial hierarchy of coarse problems, and to do so in such a way that the nice computational complexity and approximation properties of multigrid on nested meshes are preserved. There is a significant cost associated with gathering information about the type of shape functions, degrees of freedom, geometry, etc. Every specialization restricts the scope of the finite element codes the multigrid code can be used with, and the information may not be even available from vintage finite element codes without a very significant effort and is prone to errors. So, algebraic multigrid should use only information in the matrix of the system and as little extra information as possible.

* This research was supported by NSF grants ASC-9121431 and ASC-9217394.

† Center for Computational Mathematics, University of Colorado at Denver, Denver, CO 80217-3364. The first author has been on a visit from University of West Bohemia, Americká 42, 306 14 Plzeň, Czech Republic

The first approach to such matrix based multigrid was the AMG code of Ruge, et al. [14, 7, 15]. This code introduced the concept of *strongly coupled nodes* and attempted to find a coarsening from the condition that interpolation should prefer to proceed along strong couplings. The associated interpolations were based on the condition of flux preservation as in [1] exploiting the condition that the kernel of the matrix consists of constant functions (aside from boundary conditions). Since constant functions can be expressed easily without knowing much about the problem, the resulting solver could be formulated in terms of the matrix only. We use the concept of strong connections to select our aggregates of nodes, which play a role similar to Ruge's coarse points. The theory for AMG was based on two-level estimates [6] which do not give convergence bounds independent of the number of levels. The deterioration of convergence was indeed observed, and to defeat this, relatively many coarse points had to be selected, resulting in less attractive computational complexity.

Many authors have taken a geometrical approach: the coarse mesh is selected independently of the original fine mesh, and prolongation from the coarse to the fine mesh points is defined as some kind of linear interpolation using the geometry, see, e.g., Mavriplis [13]. Chan and Smith [8] define a coarse grid by graph theoretical means and then use a geometry based interpolation.

Our approach is based on the concept of *smoothed aggregation* introduced recently by Vaněk [18, 17]. First the set of nodes is decomposed into small mutually disjoint subsets. A tentative piecewise constant interpolation (in the discrete sense) is then defined using this decomposition. The final interpolation operator is obtained by smoothing the result of such piecewise constant interpolation. The coarse level operators are then defined variationally. The resulting method converges very fast for a wide range of problems including those with strongly anisotropic and discontinuous coefficients. In addition, the new method has remarkably low computational complexity since our typical coarsening ratio is about three in each dimension.

The theory for our method is based on the multilevel regularity free approach of Bramble, Pasciak, Wang, and Xu [5], cf. also Wang [20, 19]. While two-level convergence bounds are available even for the case of prolongations defined by aggregation only [6, 11], two-level bounds do not imply optimal multi-level bounds. To verify the assumptions of the theory from [5], we only need to assume that *the coarsening is by about the factor of three*, and that the aggregates of the nodes are based on aggregated elements that form a *reasonable mesh of macroelements*.

The key theoretical tool is a *bound on the energy of coarse basis functions* used in a compactness argument. We obtain similar estimates as in the classical finite element theory but without the usual assumption that the basis functions are mapped shape functions from a reference element. This allows us to treat general coarsenings of unstructured grids.

The paper is organized as follows: in Section 2, we review the basic multigrid algorithm, followed by formulation of the model problem in Section 3. Practical requirements for building prolongation operators on unstructured grids and the smoothed aggregation algorithm for building such prolongations are the subject of Section 4. Extensions to more general problems are discussed in Section 5. Section 6 contains the theoretical bound on the speed of convergence of our method. Numerical results are presented in Section 7.

2. Basic Multigrid Algorithm. An algebraic multigrid algorithm for the solution of the system of linear algebraic equations

$$Ax = b \tag{2.1}$$

consists of a preprocessing stage and the iterative process itself.

The preprocessing stage creates full rank prolongation matrices P_l of size $n_l \times n_{l+1}$, $l = 1, \ldots, L-1$ by an automatic coarsening process to be specified later. The coarse matrices are then defined variationally,

$$A_{l+1} = P_l^T A_l P_l,$$

where $A = A_1$. Since the number of levels L is not known in advance and the finest level is the only one given in the input data, we find it natural to number the levels so that *the finest level is* 1 *and the coarsest level is* L.

Once the matrices A_l and P_l are known, the standard multigrid algorithm proceeds as follows. The algorithm will be fully specified later by the choice of the cycle parameters ν_1, ν_2, γ, the hierarchy of the prolongation operators P_l, and the smoothing operators $\mathcal{S}^l(x^l, b^l)$, which are consistent iterative methods for the problem $A_l x^l = b^l$.

Algorithm 1 (Basic multigrid). *To solve the system $A_l x^l = b^l$, do:*

Pre-smoothing: *do ν_1 times $x^l \leftarrow \mathcal{S}^l(x^l, b^l)$*

Coarse grid correction:

- *let $b^{l+1} \leftarrow P_l^T(b^l - A_l x^l)$*
- *If $l+1 = L$, solve $A_{l+1}x^{l+1} = b^{l+1}$ by a direct method, otherwise apply γ iterations of this algorithm on level $l+1$, starting with initial guess $x^{l+1} = 0$*
- *correct the solution on level l by $x^l \leftarrow x^l + P_l x^{l+1}$*

Post-smoothing: *do ν_2 times $x^l \leftarrow \mathcal{S}^l(x^l, b^l)$.*

The parameters we use to obtain the theoretical estimates are $\nu_1 = \nu_2 = \gamma = 1$ with polynomial smoothers $\mathcal{S}^l(.,.)$. The numerical experimants use $\nu_1 = \nu_2 = \gamma = 1$ with the pre-smoothing iteration consisting of one forward iteration of the Gauss-Seidel followed by one iteration of backward SOR. The post-smoothing iteration consists of one forward SOR iteration followed by an iteration of backward Gauss-Seidel. The over-relaxation parameter used is 1.85 in both pre and post-smoothing.

3. Model Problem. Our theoretical analysis applies to the follwing model problem, although the algorithm is formulated more generally.

Consider the solution of algebraic systems from the discretization by standard conforming linear finite elements of a second order elliptic variational problem

$$u \in V: \quad a(u,v) = f(v) \qquad \forall v \in V \tag{3.1}$$

where $V = H^1_{\Gamma_D}(\Omega)$ denotes the Sobolev space of H^1 functions vanishing on $\Gamma_D \subset \partial\Omega$, $\mu(\Gamma_D) > c\mu(\partial\Omega)$, Ω a domain in $\mathbb{R}^2$. The bilinear form

$$a(u,v) = \int_\Omega \sum_{i,j} a_{ij}\partial_i u \partial_j v \tag{3.2}$$

is assumed to be symmetric, V-elliptic, and bounded,

$$c_1\|u\|^2_{H^1(\Omega)} \leq a(u,u) \leq c_2\|u\|^2_{H^1(\Omega)}, \qquad \forall u \in V \tag{3.3}$$

In particular, there is no absolute term in (3.2), so the kernel of $a(u,u)$ in the absence of boundary conditions consists of constant functions. For the case with an abolute term, see Section 5.3.

The model problem (3.1) is discretized by conforming P1 or Q1 elements on a shape regular grid. The shape regularity of the mesh and the uniform ellipticity and boundedness (3.3) are needed for our theoretical estimates, though the algorithm itself is developed to handle well the case of numerical anisotropies and strongly varying coefficients.

The space of finite element functions $V_l \subset V$ and the space of the vectors of degrees of freedom $\widehat{V}_l = \mathbb{R}^{n_l}$ are associated with each level. The space $\widehat{V}_l$ is the space of coordinates of functions in V_l with respect to the basis $\{\varphi^l_i\}^{n_l}_{i=1}$. Let $\{\varphi^1_i\}^{n_1}_{i=1}$ be the finite element basis of $V_1 = V_h$. Then

$$\begin{bmatrix} \varphi^{k+1}_1 \\ \vdots \\ \varphi^{k+1}_{n_{k+1}} \end{bmatrix} = P^T_k \begin{bmatrix} \varphi^{k}_1 \\ \vdots \\ \varphi^{k}_{n_{k}} \end{bmatrix} \qquad k = 1, \ldots, L-1, \tag{3.4}$$

Finally, we assume that the finite element basis on V_1 satisfies the decomposition of a constant property

$$\sum_{i=1}^{n_1} \varphi^1_i = C \tag{3.5}$$

on all elements aside from essential boundary condition.

4. Construction of Prolongations. The selection of the prolongation operators is dictated by the desire to achieve good theoretical estimates of the convergence of the iterations as well as by bounds on the computational complexity of the algorithm.

4.1. Requirements on prolongations. We seek prolongations that satisfy the following properties. First we specify the desired properties of the support of the coarse shape functions (or, equivalently, the allowed nonzeros of the prolongation matrices), and then the numerical values of the nonzero entries.

(AMG1) Coarse supports should follow strong couplings. Two nodes i and j on level l are strongly coupled if $|a^l_{ij}|$ is relatively large compared with $\sqrt{|a^l_{ii}a^l_{jj}|}$. We require that every two nodes in the support of a coarse basis function can be connected by a path of strong couplings. For example, strongly anisotropic equations can be solved in such a way that coarsening is done in the direction of anisotropy only (so-called *semi-coarsening*, [10], [16]). Algebraically, the anisotropy is reflected in the coefficients of the stiffness matrix in the sense that the neighboring nodes are strongly coupled in the direction of anisotropy.

(AMG2) Bounded intersection. There exists a constant K such that for any basis function its support intersects at most K supports of other basis functions of the same coarse space. This property guarantees sparsity of the resulting coarse-level matrices; the number of nonzeros per column in the matrix A_l is at most $K + 1$.

(AMG3) Decomposition of unity. Every coarse space V_l should represent the constant function exactly, aside from essential boundary condition. This requirement is motivated by the need to bound locally the error of a coarse grid approximation $P_l v^{l+1}$ of a fine grid function u^l in terms of the energy $(u^l)^T A_l u^l$ and by the fact that the constant function has zero energy because of (3.2). This requirement is present in two-level theories based on strengthened Cauchy inequality and its variants [3, 4, 6] and a similar requirement can be proved to be necessary for some methods [12]. Since the basis of the finest space is assumed to decompose a constant, cf., (3.5), we only need to require that the columns of each prolongation matrix form a decomposition of unity

$$\sum_{j=1}^{n_{l+1}} P_{ij} = 1, \qquad l = 1, \dots, L-1. \tag{4.1}$$

for all rows i that do not correspond to degrees of freedom adjacent to an essential boundary condition. For generalizations, see Sections 5.2 and 5.3.

(AMG4) Small energy of coarse basis functions. In the model 2D problem, we need that all basis functions satisfy the inverse inequality

$$\|\varphi_i^l\|^2_{H^1(\Omega)} \leq C\mu(\mathrm{supp}\varphi_i^l)^{-1}\|\varphi_i^l\|^2_{L^2(\Omega)} \tag{4.2}$$

In the case of a more general problem with anisotropies and strongly variable coefficients, we require that the energy of the coarse space basis functions be almost minimal in the sense that

$$\frac{a(\varphi_i^l, \varphi_i^l)}{||\varphi_i^l||^2_{L^2(\Omega)}} \leq C \inf_{u \in H_0^1(\mathrm{supp}\varphi_i^l)} \frac{a(u,u)}{||u||^2_{L^2(\Omega)}}.$$

(AMG5) Uniform L^2 equivalence. The discrete l^2 and the continuous L^2 norms should be equivalent on all coarse spaces:

$$||\sum_{i=1}^{n_l} x_i \varphi_i^l||_{L^2(\Omega)} \approx \left(\sum_{i=1}^{n_l} \mu_i^l x_i^2\right)^{1/2}, \qquad \forall x \in \mathbb{R}^{n_l},\ l = 1, \dots, L, \tag{4.3}$$

where the constants of equivalence are independent of the level and $\mu_i^l = \mu(\mathrm{supp}\varphi_i^l)$

We now proceed to the construction of prolongations P_l based on the matrix A_l, attempting to satisfy the above requirements.

4.2. Tentative prolongation by aggregation. First we specify a disjoint decomposition of the degrees of freedom on each level. Every component of the decomposition on level l, called also an *aggregate*, gives rise to a degree of freedom on level $l+1$. The decomposition will be done in such a way that the piecewise constant interpolation based on the aggregates would satisfy all of the properties above except for the energy bound in (AMG4).

According to the requirement (AMG1) above, for a given ε define the *strongly-coupled neighborhood of node* i as

$$N_i^l(\varepsilon) = \{j : \ |a_{ij}| \geq \varepsilon\sqrt{m_i m_j}\} \cup \{i\}, \tag{4.4}$$

where $m_j = |a_{jj}|$.

ALGORITHM 2 (AGGREGATION). *Let the matrix A_l of order n_l and $\varepsilon \in [0,1)$ be given. Generate a disjoint covering $\{C_i^l\}_{i=1}^{n_{l+1}}$ of the set $\{1,\ldots,n_l\}$ as follows.*

Initialization *Set $R = \{1,\ldots,n_l\}$ and $j = 0$.*

Step 1 *Select disjoint strongly coupled neighborhoods as the initial attempted covering: If there exists a strongly coupled neighborhood $N_i^l(\varepsilon) \subset R$, set $j \leftarrow j+1$, $C_j^l \leftarrow N_i^l(\varepsilon)$, $R \leftarrow R \setminus C_j^l$. Repeat until R does not contain any strongly coupled neighborhood.*

Step 2 *Add each remaining $i \in R$ to one of the sets already selected to which it is strongly connected, if possible:*
Copy $\tilde{C}_k^l = C_k^l$, $k = 1,\ldots,j$
If there exists $i \in R$ and k such that $N_i^l(\varepsilon) \cap \tilde{C}_k^l \neq \emptyset$ then set $C_k^l \leftarrow C_k^l \cup \{i\}$.
Repeat until no such i exists.

Step 3 *Make the remaining $i \in R$ into aggregates that consist of subsets of strongly coupled neighborhoods: If there exists $i \in R$, set $j \leftarrow j+1$ and $C_j^l = R \cap N_i^l(\varepsilon)$. Repeat until $R = \emptyset$.*

Define the tentative prolongation $\tilde{P}_l$ by the aggregates C_l^i:

$$(\tilde{P}_l)_{ij} = \begin{cases} 1 \text{ if } i \in C_j^l \\ 0 \text{ otherwise} \end{cases} \tag{4.5}$$

4.3. Final smoothed prolongation. The piecewise constant prolongation $\tilde{P}_l$ will now be improved by a smoothing to get the final prolongation matrix P_l. This smoothing will preserve the desired properties of the prolongation while decreasing the energy of the coarse basis functions. We choose a simple Jacobi smoother, giving the prolongation matrix

$$P_l = (I - \omega D^{-1} A_l^F)\tilde{P}_l \tag{4.6}$$

where $A_l^F = (a_{ij}^F)$ is the *filtered matrix* given by

$$a_{ij}^F = \left\{ \begin{array}{cl} a_{ij} \text{ if } & j \in N_i^l(\varepsilon) \\ 0 & \text{otherwise} \end{array} \right\} \text{ if } i \neq j, \qquad a_{ii}^F = a_{ii} - \sum_{j=1, j\neq i}^{n_l} (a_{ij} - a_{ij}^F). \tag{4.7}$$

The purpose of the filtering (4.7) instead of just using the matrix A_l in (4.6) is to decrease the number of the nonzeros in the prolongation matrix P_l. This effect is pronounced in anisotropic problems, where the smoothed prolongation will turn out to be piecewise constant in the direction of weak couplings, but there is little or no effect due to filtering for isotropic problems and small ε.

It is easy to see that the prolongation by aggregation satisfies the decomposition of unity property (4.1) (in the absence of boundary conditions). From the construction of the filtered matrix (4.7), it follows that then A_l^F also vanishes on constant functions, so from (4.6), the smoothed prolongation still satisfies the decomposition of unity property (4.1).

Fig. 4.1 shows the coarse basis functions given by prolongation by aggregation and the smoothed aggregation in 1D. Note that for the 1D Laplace operator and the choice of $\omega = 2/3$ in (4.6), the smoothed coarse space basis is exactly the one of $P1$-finite elements. Fig. 4.2 shows the typical aggregates obtained on an unstructured grid. The corresponding supports are formed by adding one belt of elements to the aggregates. The smoothing adds at most one more belt of adjacent elements.

Numerical experiments suggest the choice of the parameters

$$\varepsilon = 0.08\,(\frac{1}{2})^{l-1}, \qquad \omega = \frac{2}{3}.$$

Note that ε decreases as level index increases. The motivation for this choice of ε is that large values of ε make the coarsening process slow down. Our choice of ε proves to preserve the pace of coarsening. The computations show the convergence rate to be insensitive to this variation of ε on the coarse levels even in case of problems with large anisotropies.

Algorithm 2 applied to uniformly elliptic problems typically creates aggregates C_k^l, with $\mathrm{card}(C_k^l)$ of about 9. In particular, we can expect coarsening by about a factor of 3 in each dimension. For anizotropic problems, the coarsening will be by about a factor of 3 in the direction of the anizotropy only. The resulting coarse level matrix A_{l+1} tends to follow in either case (due to the matrix filtration) the nonzero pattern of the 9-point stencil.

5. Generalizations. We now summarize several observations and extensions useful in practice, though we will not dwell on their complete theoretical justification.

5.1. Vector problems. The coarsening process can easily be generalized to the case of non-scalar problems with a constant number of degrees of freedom per node (e.g., 2D or 3D elasticity problems) satisfying the decomposition of a constant property (3.5) field-by-field. Let n_d denote the number of degrees of freedom per node. The communication between the neighboring nodes k, l can now be expressed in the form of a matrix selection A_{kl} of order n_d

$$A_{kl} = A(df(k), df(l)),$$

where $df(i)$ denotes the list of degrees of freedom associated with node i. This so called *block approach* consists in the simple idea of replacing the scalar operations by their

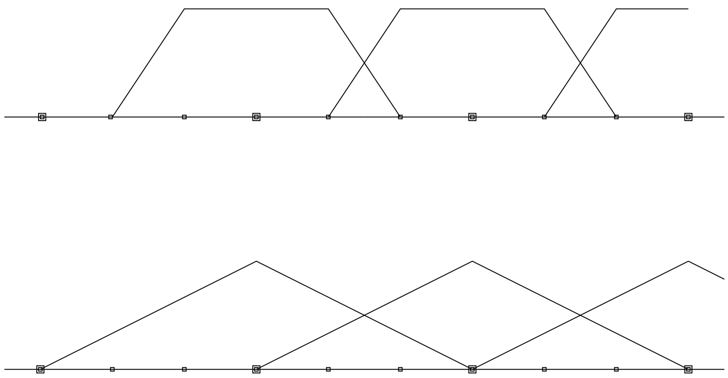

FIG. 4.1. *The basis functions given by aggregation and the corresponding smoothed basis for 1D Laplacian, using the smoother $I - 2/3D^{-1}A$*

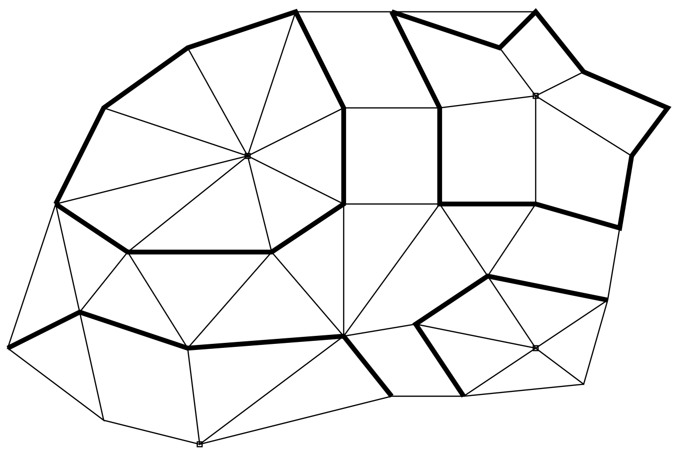

FIG. 4.2. *Typical 2D aggregates*

corresponding block-matrix counterparts.

The strongly coupled neighborhood of node i will be redefined as

$$N_i^l = \{ \, j : \ \ \|A_{ij}\| \geq \epsilon \sqrt{m_i \, m_j} \} \cup \{i\}, \tag{5.1}$$

where $m_j = \|A_{jj}\|$, and $\|.\|$ is a matrix norm. Further, in the definition of auxiliary prolongations (4.5), we replace the numbers 1 and 0 by identity and zero matrices of order n_d, respectively. The efficiency of this generalization is demonstrated by Experiments No. 1 and No. 2 in Section 7.

5.2. High order elements and unscaled problems. Condition (3.5) seems restrictive (e.g., finite elements of high order may violate it). This difficulty, however, can be circumvented easily. To this end, we need to have information about the coefficients of the decomposition of unity by the finite element basis of V_1

$$\sum_{i=1}^{n_1} \alpha_i \varphi_i^l = 1,$$

where $\alpha \in \mathbb{R}^{n_1}$ is user input data. As in case of (3.5), the decomposition is understood away from essential boundary conditions. The definition (4.5) of the auxiliary prolongators will remain in place for all levels except for level 1; we define $\tilde{P}_1$ as

$$\tilde{P}_{1_{ij}} = \begin{cases} \alpha_i \text{ if } i \in C^1_j \\ 0 \text{ otherwise} \end{cases} \tag{5.2}$$

Note that this definition will cause the representation of a constant to change only on $\widehat{V}_1$; on $\widehat{V}_2, \ldots, \widehat{V}_L$, the representation of a constant function will be a constant vector of coordinates as in the original case when (3.5) is satisfied. Generalization to the nonscalar case is straightforward. It was applied to the problem from Example No. 1 of Section 7 modified by scaling the basis functions randomly in the interval $[0.01, 1]$. The results are summed up in Example No. 7.

5.3. Absolute term. If (3.1) is replaced by

$$a(u,v) = \int_\Omega \sum_{i,j} a_{ij} \partial_i u \partial_j v + quv, \qquad q > 0,$$

then the constant is no longer locally in the kernel of A_l and the prolongation smoothers will not preserve the constant. The matrix, however, will be better conditioned, which compensates for the loss of the preservation of a constant. The influence of the variation of the absolute term is demonstrated in Experiment No. 6 in Section 7.

The presence of a very large absolute term as well as certain ways of handling the Dirichlet boundary conditions cause the diagonal dominance to be very strong in some columns. The nodes corresponding to these columns behave as isolated nodes in the coarsening algorithm and may prevent further coarsening without reaching the desired size of the coarsest level. The smoothers $\mathcal{S}^l$, however, are very efficient in approximating the values in these nodes. Therefore, these nodes can be taken out of the set R in the initialization step of aggregation Algorithm 2. This modification prevents stalling of coarsening and has no adverse effect on the convergence of the method.

6. Convergence Analysis. In this section, we prove an almost optimal bound on the convergence of the method for the case of the model problem from Section 3 discretized on a shape regular mesh (that is, no strong anisotropies). We also need to make some natural assumptions on the properties of the aggregates C^l_i and the supports of the coarse shape functions, which tend to be satisfied for the aggregates generated by Algorithm 2.

The constants in the estimates in this section are independent of the number of elements on the finest grid and of element sizes but they do depend on the H^1 equivalence constants (3.3) of the form $a(\cdot,\cdot)$ and the shape regularity of the mesh. Thus, our estimates are uniform on a class of fine grids that satisfy a uniform bound on the condition of the Jacobians of the map from the reference element to an arbitrary element. In particular, the estimates are independent of the number of levels.

In the rest of this section, the prolongation are defined by

$$P_l = S_l \tilde{P}_l, \tag{6.1}$$

where

$$S_l = I - \sigma_l A_l. \tag{6.2}$$

and the parameter σ_l is given as follows. Let $\bar{\lambda}_1$ be a given bound of $\varrho(A_1)$ such that

$$\varrho(A_1) \le \bar{\lambda}_1 \le C\varrho(A_1) \tag{6.3}$$

and for $l = 1, \ldots, L-1$,

$$M_l = \max_{i=1,\ldots,n_{l+1}} \operatorname{card}(C_i^l).$$

Then let

$$\bar{\lambda}_{l+1} = \frac{M_l}{9}\bar{\lambda}_l, \qquad \sigma_l = \frac{4}{3\bar{\lambda}_l}. \tag{6.4}$$

With the definition (6.1) – (6.4) of the prolongation, we first prove a bound on the spectral radii of the coarse level matrices, which also gives a bound on the energy of the coarse basis functions.

LEMMA 6.1. *For all $l = 1, \ldots, L$, it holds that $\varrho(A_l) \le \bar{\lambda}_l$.*

Proof. The case $l = 1$ follows from (6.3). Suppose the proposition is true for some $l \ge 1$. Then

$$\langle A_{l+1}x, x\rangle = \langle A_l S_l \tilde{P}_l x, S_l \tilde{P}_l x\rangle \le \varrho(S_l^2 A_l)\langle \tilde{P}_l x, \tilde{P}_l x\rangle \le M_l \varrho(S_l^2 A_l)\langle x, x\rangle$$

Because $S_l = I - \sigma_l A_l$,

$$\varrho(S_l^2 A_l) \le \max_{\xi\in[0,\bar{\lambda}_l]} (1 - \frac{4}{3}\frac{1}{\bar{\lambda}_l}\xi)^2 \xi = \frac{1}{9}\bar{\lambda}_l,$$

concluding the proof for $l+1$. □

The next theorem follows immediately. Since our coarsening is approximately by the factor of three, Theorem 6.2 gives a uniform bound on the coarse matrices, which will be assumed in the next section.

THEOREM 6.2. *Let the prolongation be given by (6.1) – (6.4), and let*

$$M_1 M_2 \cdots M_l \le C_0 9^l, \qquad l = 1, \ldots, L-1.$$

Then $\varrho(A_l) \le \bar{\lambda}_1 C_0$, $l = 2, \ldots, L$.

We are now ready to formulate the assumptions on the coarse space hierarchy

$$V_1 \supset V_2 \supset \ldots \supset V_L,$$

where $V_1 \subset V$ is the $P1$ (or $Q1$)-finite element space given by a shape regular triangulation of Ω.

ASSUMPTION 1. *The coarse space basis functions $\{\varphi_i^l\}$ form a system of* macroelements *$\{T_i^l\}$ with the following properties:*

1. $\bar{\Omega} = \bigcup T_i^l$

2. The boundary of each T_i^l *consists of the edges of the finite elements on* V_1.

3. Each macroelement T_i^l *can be mapped onto one of a finite number of* polygonal master macroelements *so the condition number of the Jacobian of this mapping is uniformly bounded.*

4. Each macroelement T_i^{l+1} *consists of at most* C_M *macroelements* T_i^l.

5. There is a basis function φ_i^l *associated with every vertex* v_i^l *in the interior of* Ω. *On the interior macroelements, these basis functions satisfy*

$$\sum_{k=1}^{n} \varphi_{i_k}^l = 1.$$

6. For any macroelement T, *its edge* e *is either an edge of another macroelement, or part of* $\partial\Omega$. *In the latter case, if* $\mu(e \cap \Gamma_D) \neq \emptyset$,

$$\mu(\partial T \cap \Gamma_D) > C\mu(\partial T). \tag{6.5}$$

7. The restriction of the coarse space function to the macroelement is spanned by the associated basis functions.

8. Basis function associated with a vertex vanishes on the edges which are not adjacent to that vertex.

9. The energy of all coarse shape functions is uniformly bounded:

$$|\varphi_i^l|_{H^1(\Omega)} \leq C_E. \tag{6.6}$$

(This is satisfied from Theorem 6.2.)

Note that we have made no assumptions on quasiuniformity of the mesh.

The main difference between the assumptions above and the usual case of nested finite element spaces typically assumed in multigrid methods is that we make no assumption about the actual form of the coarse shape functions. In particular, the coarse shape functions are not finite element shape functions on some coarse elements.

Multigrid theories rely on the fact that the discrete l^2 norms on the spaces of the vectors of degrees of freedom are, after proper scaling, uniformly equivalent to the L^2 norm of the associated finite element functions. This is trivial if the coarse space are defined by coarse finite elements with the usual shape functions. Here we prove a similar property under our assumptions. We only need to treat the the case of a reference macroelement.

Lemma 6.3. *Let* K *be a polygon with edges* e_1 *to* e_n *and* $C_0 > 0$. *Then there exists constants* $C_2 > C_1 > 0$ *with the following property: If* $f_1, \ldots, f_n \in H^1(K)$ *satisfy*

(i) $\sum_{i=1}^n f_i = 1$ *in* K,

(ii) $f_i = 0$ *in the sense of traces on all edges of* K *except* e_i, e_{i+1} *(with the cyclic numbering* $e_{n+1} = e_1$*),*

(iii) $\|f_i\|_{H^1(K)} \leq C_0$, *for all* i,

then

$$C_1 \sum_{i=1}^n \alpha_i^2 \leq \|\sum_{i=1}^n \alpha_i f_i\|_{L^2(K)} \leq C_2 \sum_{i=1}^n \alpha_i^2, \qquad \forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}.$$

Proof. Denote by G the set of all n-tuples $(f_1, \ldots, f_n)$ that satisfy the conditions (i)—(iii). The set G is bounded in $(H^1(K))^n$, hence weakly precompact. Since G is closed and convex, it is weakly closed, so G is weakly compact. Because of the compact imbedding of H^1 in L^2, G is compact in $(L^2(K))^n$.

Now we show that every n-tuple in G is linearly independent. Assume there is a linear combination $\sum_{i=1}^n c_i f_i = 0$ in K, and, say, $c_1 \neq 0$. Since all functions except f_1 and f_2 are zero on the edge e_2, we have

$$f_1 + f_2 = 1, \quad c_1 f_1 + c_2 f_2 = 0, \qquad \text{on } e_2.$$

It follows that f_1 is a nonzero constant on e_2. But $f_1 = 0$ on e_3 by (ii), so the trace norm $\|f_1\|_{H^{1/2}(\partial K)} \geq \|f_1\|_{H^{1/2}(e_2 \cup e_3)}$ is not finite, which is a contradition.

The entries of the Gram matrix $M = \{(f_i, f_j)_{L_2(K)}\}_{i,j=1}^n$ are continuous functions on G in the $(L^2(K))^n$ norm. Since the extreme eigenvalues are continuous functions of the entries of the matrix, the function $g(f_1, \ldots, f_n) = \lambda_{\min}(M)$ is continous on G in the $(L^2(K))^n$ norm. Because G is compact, g attains a minimum C_1 and a maximum C_2 on G. Because all n-tuples in G are linearly independent, $g > 0$ on G, so $C_1 > 0$. ☐

On the coarse spaces V_l, denote the l^2 norm

$$\|u\|_l = \left(\sum_{i=1}^{n_l} u_i^2 \right)^{1/2}, \qquad u \in V_l, \quad u = \sum_{i=1}^{n_l} u_i \varphi_i^l.$$

We now construct an operator $\Pi_l : V_1 \rightarrow V_l$ satisfying H^1-stability and local optimal approximation property in the $\|\cdot\|_l$ norm. The construction is by an averaged interpolation, and, except for the fact that we do not work with the usual finite element shape functions, it is standard, e.g. [21].

Let $u \in V$. For every basis function φ_i^l corresponding to a vertex v_i^l in the macroelement triangulation, define the averaged value of u on a neighborhood of v_i^l by

$$\alpha_i^l = \frac{1}{\mu(N)} \int_N u,$$

where N is the circle with the center v_i^l and the radius equal to $1/3$ of the minimum of the diameters of all macroelements v_i^l is adjacent to. Then let

$$\Pi_l u = \sum_{i=1}^{n_l} \alpha_i^l \varphi_i^l.$$

Lemma 6.4. *There exists a constant C independent of the number of levels and number of elements such that for all $u \in V_1$,*

$$|\Pi_l u|_{H^1(\Omega)} \leq C|u|_{H^1(\Omega)} \tag{6.7}$$

$$||(\Pi_l - \Pi_{l+1})u||_l \leq C|u|_{H^1(\Omega)} \tag{6.8}$$

Proof. Consider the union A of a macroelement T on level l along with all neighboring macroelement, which is the image of a configuration of reference macroelements $\hat{A}$ by a Lipschitz continuous mapping $F : \hat{A} \rightarrow A$ such that the Jacobian ∂F is uniformly well conditioned:

$$\|\partial F(x)\| \|\partial F^{-1}(\hat{x})\| \leq C, \quad \hat{x} = F(x), \tag{6.9}$$

and such that the macroelement T is the image of a reference macroelement $\hat{T} \subset \hat{A}$. There is only a finite number of such reference configurations $\hat{A}$. Because of (6.9), the mapping F preserves the H^1 seminorms

$$|u|_{H^1(A)} \approx |\hat{u}|_{H^1(\hat{A})}, \quad \hat{u}(\hat{x}) = u(F(\hat{x})) \tag{6.10}$$

with uniform constants.

Mapping the integrals in the construction of Π, we obtain the local mapped averaged interpolation $\hat{\Pi} : H^1(\hat{A}) \rightarrow H^1(\hat{T})$. Using the definition of $\hat{\Pi}$, boundedness of the basis functions in the H^1 seminorm and the Poincaré inequality $||\hat{u}||_{L^2(\hat{A})} \leq C||\hat{u}||_{H^1(\hat{A})}$ when $\partial T \cap \Gamma_D \neq \emptyset$ (or the invariance of both sides of the inequality to adding a constant in the case of $\partial T \cap \Gamma_D = \emptyset$) yields

$$|\hat{\Pi}\hat{u}|_{H^1(\hat{T})} \leq C|\hat{u}|_{H^1(\hat{A})}.$$

Mapping back to the original macroelements and using (6.10) we get

$$|\Pi_l u|_{H^1(T)} \leq C|u|_{H^1(A)}.$$

The proof of (6.7) is finished by summation over all macroelements on level l.

To prove (6.8), consider the level $l-1$ macroelements contained in A mapped on the reference configuration of macroelements in $\hat{A}$. There is again a finite number of such two-level reference configurations. Define $\hat{\Pi}_l = \hat{\Pi}$ as above and let $\hat{\Pi}_{l-1}$ be defined by mapped averaged interpolation on the level $l-1$ macroelements. First consider the case when $\partial T \cap \Gamma_D = \emptyset$. Then, since both sides are invariant to adding a constant to $\hat{u}$, and the operator $\hat{\Pi}_l$ is L^2 bounded,

$$||\hat{\Pi}_l \hat{u} - \hat{\Pi}_{l-1}\hat{u}||_{L^2(\hat{T})} \leq C|\hat{u}|_{H^1(\hat{A})}. \tag{6.11}$$

When $\partial A \cap \Gamma_D \neq \emptyset$, (6.5) implies Poincaré inequality on $\hat{A}$. From the fact that $\hat{\Pi}_l$ is L^2 bounded, it follows that

$$\|\hat{\Pi}_l \hat{u}\|_{L^2(\hat{T})} \leq C\|\hat{u}\|_{L^2(\hat{A})} \leq C|\hat{u}|_{H^1(\hat{A})}$$

and analogously for $\hat{\Pi}_{l+1}$. Hence, we have again (6.11). Using Lemma 6.3, we can replace in (6.11) the L^2 norm by the l^2 norm of the expansion in terms of basis functions $F^{-1}\varphi_i^l$ of level l on $\hat{T}$. Mapping the inequality (6.11) back from $\hat{A}$ to A, adding over all macroelements T on level l. □

We are now ready for the convergence bound.

THEOREM 6.5. *Let the Assumption 1 be satisfied and let $e_k \in \widehat{V}_1$ denote the error after k-th iteration. Then for the problem (3.1) with the prolongation (4.6) and consistent smoothers $\mathcal{S}^l(.,.)$ with the linear part given by (6.2) and (6.4) (the same choice as for the prolongation smoothers), it holds that*

$$||e_{k+1}||_{A_1} \leq (1 - \frac{C}{L})||e_k||_{A_1},$$

where the constant C does not depend on the number of levels or on the number of elements.

Proof. The proposition follows immediately from Theorem 1 in [5], which only assumes the existence of a mapping $\Pi_l : V_1 \to V_l$ such that

$$\begin{array}{rcl} |\Pi_l u|_{H^1(\Omega)} & \leq & C|u|_{H^1(\Omega)}, \\ ||(\Pi_l - \Pi_{l+1})u||_l & \leq & C(\varrho(A_l))^{-1/2}|u|_{H^1(\Omega)}. \end{array}$$

which we have proved as (6.7) and (6.8) in view of (6.6). □

7. Numerical Experiments . The robustness and efficiency of the algorithm is demonstrated by the seven experiments below. The code is available through anonymous ftp to `tiger.denver.colorado.edu` , directory /pub/faculty/pvanek. The experiments were performed on an IBM RS-6000/360 with 128 MBytes of memory.

The residual was measured in the l^2 norm. By algebraic complexity we mean the number of nonzero entries in the matrices on all the levels divided by the number of nonzeros in the matrix on finest level.

The rate of convergence is computed as an average reduction of l^2-norm of residual per iteration.

We used V(1,1) cycle in all the experiments below.

Experiment No. 1:

Problem:	planar elasticity
Poisson ratio:	0.3
Boundary conditions:	Dirichlet and Neumann
Mesh:	see Fig. 7.2
Number of nodes:	10670
Number of equations:	21358
Stopping condition:	relative residual smaller than 10^{-5}
Algebraic complexity:	1.23
Rate of convergence:	0.08
Real time of computation:	5 s
CPU time of computation:	5 s

Experiment No. 2:

Problem:	planar elasticity
Poisson ratio:	0.3
Boundary conditions:	Dirichlet and Neumann
Mesh:	see Fig. 7.3 and Fig. 7.4
Number of nodes:	6473
Number of equations:	12946
Stopping condition:	relative residual smaller than 10^{-5}
Algebraic complexity:	1.58
Rate of convergence:	0.14
Real time of computation:	5 s
CPU time of computation:	4 s

Experiment No. 3:

Problem:	anisotropic problem (7.1) with $q(x,y) = 0$ and the jumps in coefficients as in Fig. 7.1
Boundary conditions:	Dirichlet
Mesh:	regular square mesh
Number of nodes:	10^6
Number of equations:	10^6
Stopping condition:	relative residual smaller than 10^{-5}
Algebraic complexity:	1.56
Rate of convergence:	0.10
Real time of computation:	7892 s
CPU time of computation:	768 s

Note the difference in the real and CPU times. This difference was caused by massive memory swapping needed to accommodate the problem of one million degrees of freedom. Thus it reflects insufficient memory rather than a deficiency of the algorithm.

The experiments were conducted for the operators

$$-\frac{\partial}{\partial x}\, a(x,y)\frac{\partial u}{\partial x} - \frac{\partial}{\partial y}\, b(x,y)\frac{\partial u}{\partial y} + q(x,y)\, u = f(x,y) \tag{7.1}$$

and

$$-\sum_{i,j=1}^{3} \frac{\partial}{\partial x_i}(w_{ij}(x,y)\frac{\partial u}{x_j}) = f(x,y), \tag{7.2}$$

see the experiment descriptions for the operator used.

Experiment No. 4:

Problem:	3D problem (7.2) with random coefficients $w_{11} = w_{22} = w_{33} = \exp(rn)$, $w_{ij} = 0$ for $i \neq j$ rn is a random number uniformly distributed in the interval $[\ln(10^{-2}), \ln(10^2)]$
Boundary conditions:	Dirichlet
Mesh:	regular square mesh
Number of nodes:	68921
Number of equations:	68921
Stopping condition:	relative residual smaller than 10^{-5}
Algebraic complexity:	1.15
Rate of convergence:	0.07
Real time of computation:	146 s
CPU time of computation:	116 s

Experiment No. 5:

Problem:	3D problem (7.2) with random coefficients $w_{11} = \exp(rn_1), w_{22} = \exp(rn_2), w_{33} = \exp(rn_3)$ $w_{ij} = 0$ for $i \neq j$ rn_i is a random number uniformly distributed in the interval $[\ln(10^{-2}), \ln(10^2)]$
Boundary conditions:	Dirichlet
Mesh:	regular square mesh
Number of nodes:	68921
Number of equations:	68921
Stopping condition:	relative residual smaller than 10^{-5}
Algebraic complexity:	1.14
Rate of convergence:	0.21
Real time of computation:	233 s
CPU time of computation:	134 s

Experiment No. 6:

Problem:	2D anisotropic problem (7.1), with $q(x, y) = 0, 1$ and 10 with jumps in coefficients as in Fig. 7.1
Boundary conditions:	Dirichlet
Mesh:	regular square mesh
Number of nodes:	160000
Number of equations:	160000
Stopping condition:	relative residual smaller than 10^{-5}

	q = 0	q = 1	q = 10
Algebraic complexity:	1.65	1.65	1.65
Rate of convergence:	0.11	0.11	0.10
Real time of computation:	95 s	96 s	91 s
CPU time of computation:	85 s	83 s	83 s

Experiment No. 7:

Problem:	planar elasticity discretized by finite elements with randomly scaled basis
Poisson ratio:	0.3
Boundary conditions:	Dirichlet and Neumann
Mesh:	see Fig. 7.2
Number of nodes:	10679
Number of equations:	21358
Stopping condition:	relative residual smaller than 10^{-5}
Algebraic complexity:	1.24
Rate of convergence:	0.09
Real time of computation:	13 s
CPU time of computation:	13 s

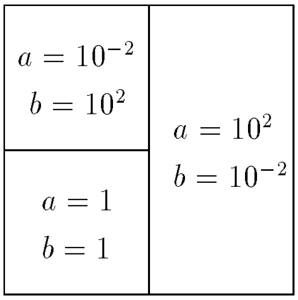

FIG. 7.1. *The coefficients* $a(x, y)$, $b(x, y)$

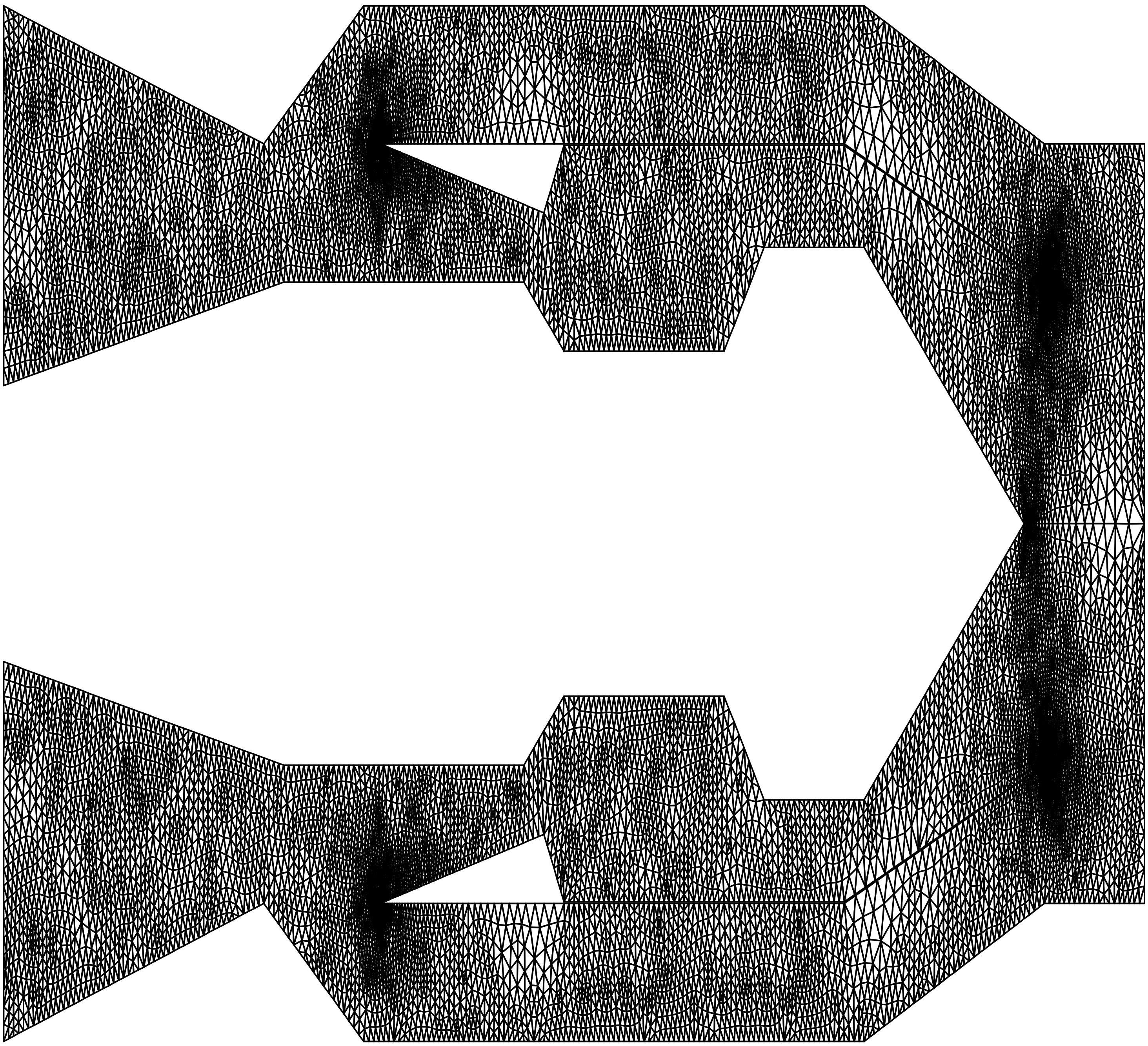

Fig. 7.2. *Mesh 1*

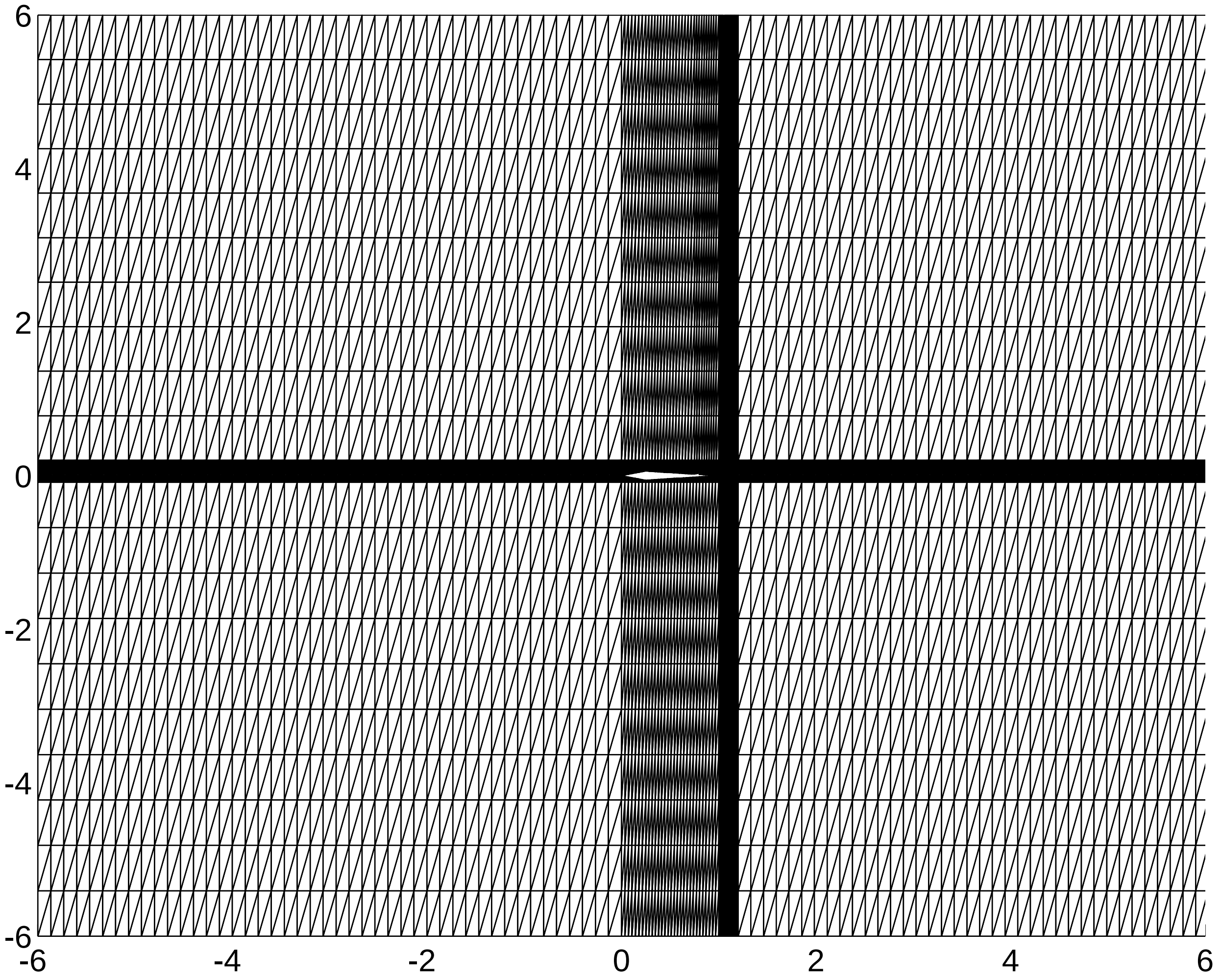

Fig. 7.3. *Mesh 2*

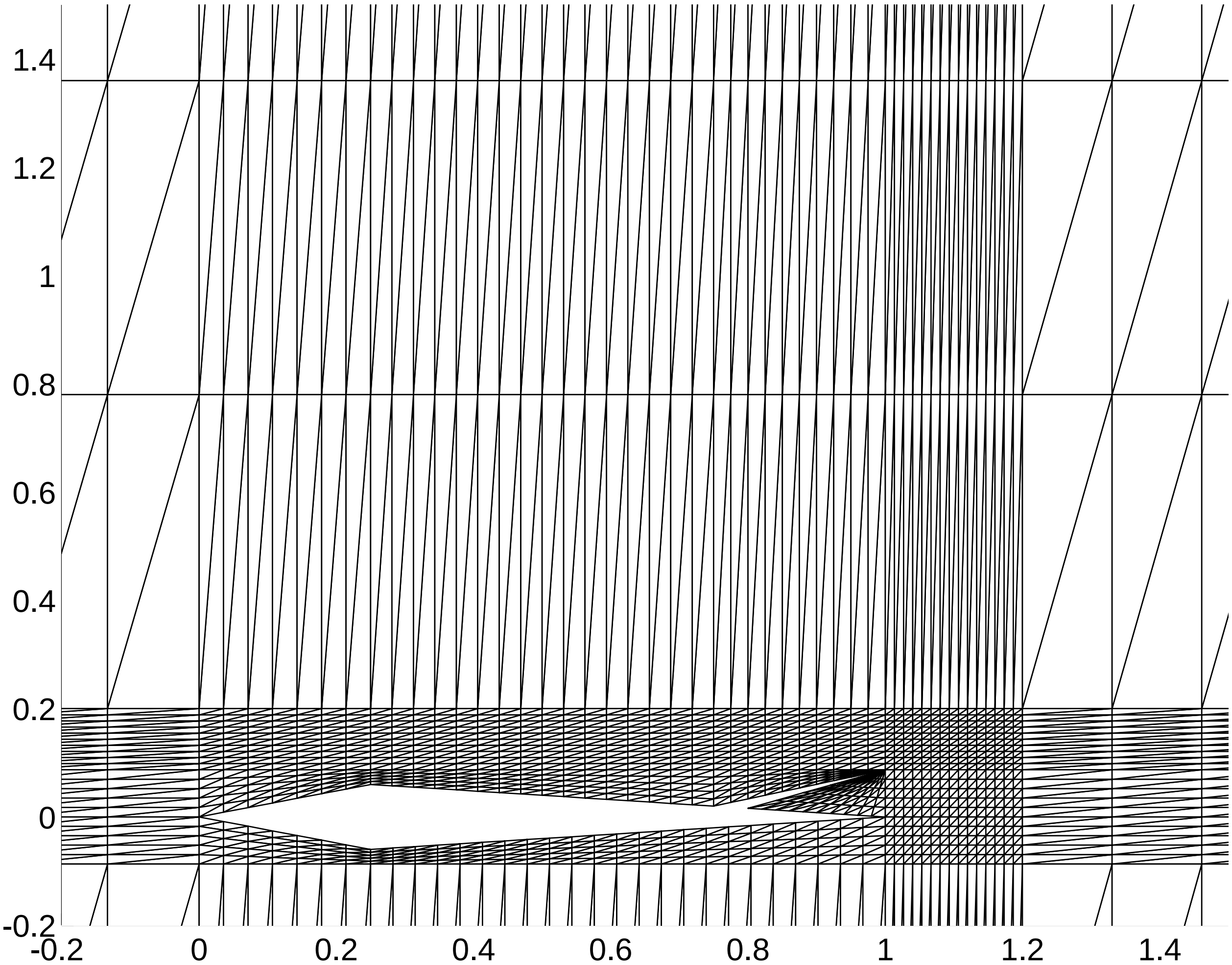

FIG. 7.4. *Mesh 2 - detail*

REFERENCES

[1] R. E. ALCOUFFE, A. BRANDT, J. E. DENDY, AND J. W. PAINTER, *The multi–grid methods for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Stat. Comput., 2 (1981), pp. 430–454.

[2] R. E. BANK, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations – Users' Guide 6.0*, SIAM Books, Philadelphia, 1990.

[3] R. E. BANK AND T. DUPONT, *An optimal order process for solving elliptic finite element equations*, Math. Comp., 36 (1981), pp. 35–51.

[4] D. BRAESS, *The contraction number of a multigrid method for solving the Poisson equation*, Numer. Math., 37 (1981), pp. 387–404.

[5] J. H. BRAMBLE, J. E. PASCIAK, J. WANG, AND J. XU, *Convergence estimates for multigrip algorithms without regularity assumptions*, Math. Comp., 57 (1991), pp. 23–45.

[6] A. BRANDT, *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comput., 19 (1986), pp. 23–56.

[7] A. BRANDT, S. F. MCCORMICK, AND J. W. RUGE, *Algebraic multigrid (AMG) for sparse matrix equations*, in Sparsity and Its Applications, D. J. Evans, ed., Cambridge Univ. Press,

Cambridge, 1984.

[8] T. F. Chan and B. Smith, *Domain decomposition and multigrid algorithms for elliptic problems on unstructured meshes*, Tech. Report 93-42, CAM, 1993.

[9] J. E. Dendy, *Black box multigrid for systems*, Appl. Math. Comput., 19 (1986), pp. 57–74.

[10] J. E. Dendy, M. P. Ida, and J. M. Rutledge, *A semicoarsening multigrid algorithm for SIMD machines*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 1460–1469.

[11] J. Mandel, *On multigrid and iterative aggregation methods for nonsymmetric problems*, in Multigrid Methods II, W. Hackbusch and U. Trottenberg, eds., Berlin, 1986, Springer–Verlag, pp. 219–231.

[12] ———, *Iterative solvers by substructuring for the p-version finite element method*, Comput. Meth. Appl. Mech. Engrg., 80 (1990), pp. 117–128.

[13] D. J. Mavriplis, *Algebraic turbulence modeling for unstructured and adaptive meshes*, AIAA J., 29 (1991), pp. 2086–2093.

[14] J. W. Ruge, *Algebraic multigrid (AMG) for geodetic survey problems*, in Prelimary Proc. Internat. Multigrid Conference, Fort Collins, CO, 1983, Institute for Computational Studies at Colorado State University.

[15] J. W. Ruge and K. Stüben, *Efficient solution of finite difference and finite element equations by algebraic multigrid (AMG)*, in Multigrid Methods for Integral and Differential Equations, D. J. Paddon and H. Holstein, eds., The Institute of Mathematics and its Applications Conference Series, Clarendon Press, Oxford, 1985, pp. 169–212.

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

[17] P. Vaněk, *Fast multigrid solver*. Applications of Mathematics, to appear.

[18] ———, *Acceleration of convergence of a two-level algorithm by smoothing trarnsfer operator*, Applications of Mathematics, 37 (1992), pp. 265–274.

[19] J. Wang, *Convergence analysis without regularity assumptions for multigrid algorithms based on SOR smoothing*, SIAM J. Numer. Anal., 29 (1992), pp. 987–1001.

[20] ———, *Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems*, SIAM J. Numer. Anal., 30 (1993), pp. 275–285.

[21] O. B. Widlund, *An extension theorm for finite element spaces with three applications*, in Numerical Techniques in Continuum Mechanics, W. Hackbusch and K. Witsch, eds., vol. 16 of Notes on Numerical Fluid Mechanics, Braunschweig, 1986, Vieweg, pp. 110–122.