SIAM REVIEW
Vol. 33, No. 3, pp. 405–419, September 1991

003

THE ROLE OF THE STRENGTHENED CAUCHY–BUNIAKOWSKII–SCHWARZ INEQUALITY IN MULTILEVEL METHODS*

VICTOR EIJKHOUT† AND PANAYOT VASSILEVSKI‡

Abstract. The basic theory of the strengthened Cauchy–Buniakowskii–Schwarz inequality and its applications in multilevel methods for the solution of linear systems arising from finite element or finite difference discretisation of elliptic partial differential equations is surveyed. Proofs are given both in a finite element context, and in purely algebraic form.

Key words. Cauchy–Buniakowskii–Schwarz inequality, multilevel finite element methods, Schur complements, preconditioned iterative methods.

AMS(MOS) subject classifications. 65F10, 65N20, 65N30

1. Introduction. Over the last decade a number of authors have presented multigrid methods in the context of multilevel finite element spaces. The main tool in the analysis of such methods is the extended Cauchy–Buniakowskii–Schwarz (C.B.S.) inequality. This is a refinement of the usual Cauchy–Buniakowskii–Schwarz (or Cauchy–Schwarz) inequality

$$|(u,v)| \le \sqrt{(u,u)}\sqrt{(v,v)}$$

in that it states in finite-dimensional spaces the existence for

$$u \in U, v \in V, \quad U, V \text{ linear subspaces}, \quad U \cap V = \{0\}$$

of a constant $\gamma \in [0,1)$, depending only on the spaces U and V, that is, independent of individual choices of u and v, such that

$$|(u,v)| \le \gamma\sqrt{(u,u)}\sqrt{(v,v)}.$$

The smallest such quantity γ may be called the cosine of the angle between the spaces U and V.

The strengthened C.B.S. inequality has been used in two-level methods by Bank and Dupont [7], Axelsson [1], Axelsson and Gustafsson [4], Braess [9], [10], and Maitre and Musy [16]. Recently, in connection with the two-grid FAC-preconditioner of McCormick [18], McCormick and Thomas [19], and the preconditioner of Bramble et al. [11], the inequality has also been used in Mandel and McCormick [17].

Multilevel theory without the strengthened C.B.S. inequality, using hierarchical finite element bases, has been proposed by Yserentant [21], Bank, Dupont, and Yserentant [8], and Vassilevski [20]. These methods, however, are near optimal only in the two-dimesnional case, and not optimal in the three-dimensional case. For a regularity-free multigrid theory that applies to two-dimensional and three-dimensional problems we refer to the recent paper by Bramble et al. [12]. Here the main tool is an

* Received by the editors August 28, 1989; accepted for publication (in revised form) December 6, 1990.

† Center for Supercomputing Research and Development, University of Illinois at Urbana-Champaign, 104 S. Wright Street, Urbana, Illinois 61801.

‡ Department of Mathematics, University of Wyoming, P.O. Box 3036, University Station, Laramie, Wyoming 82071; on leave from the Institute of Mathematics and Center for Informatics and Computer Technology, Bulgarian Academy of Sciences, 1113 Sofia, Acad. G. Bontchev Street, Block 8, Bulgaria.

abstract "product" iteration algorithm. This theory gives nearly optimal order methods. Regularity-free methods of optimal order have been proposed by Axelsson and Vassilevski [6], [5], and used for nine-point stencils in Axelsson and Eijkhout [3]. These methods are based on a corollary to the strengthened C.B.S. inequality discussed in this paper.

We would like to mention an additional advantage of the algebraic multilevel theory, namely that the strengthened C.B.S. inequality is verified locally. In particular, no regularity assumptions on the differential operator are made, as they are for the classical analysis of multigrid methods (cf. [15]). An implication of this is that the resulting bounds of the relative condition numbers of the multilevel preconditioners with respect to the corresponding stiffness matrices are independent of possible jump discontinuities in the coefficients of the differential operator if these occur across edges (or faces in three dimensions) of elements of the initial grid.

In the finite difference context an extended C.B.S. inequality plays an important role in the derivation of the two-grid FAC-preconditioners and the two-grid preconditioners of Bramble et al. [11]. See also Ewing, Lazarov, and Vassilevski [14] and Ewing et al. [13] where even certain nonsymmetric problems have been handled.

Here we summarise the results about the role of the strengthened C.B.S. inequality in two-level and multilevel methods contained in the abovementioned papers, and we give purely algebraic proofs whenever possible. In §2 we give the basic theorems needed for the analysis of two-level methods. Section 3 contains a purely algebraic treatment of the strengthened C.B.S. inequality. One justification for such a treatment is that multilevel methods can also be used for general linear systems with a symmetric positive definite coefficient matrix. More important, however, is the fact that for a quantitative analysis of general multilevel finite element methods it is necessary to investigate the element matrices involved. Thus we pay particular attention to the strengthened C.B.S. inequality in the presence of semidefinite matrices. The topic of §4 is the derivation of global bounds for the quantity γ in the case of multilevel methods. In §5 we briefly discuss the application of the strengthened C.B.S. inequality in multilevel methods, giving a proof of order 1 convergence of a simple two-level method, and indicating the extension to many levels.

2. Two basic theorems. In this section we prove the strengthened C.B.S. inequality, and the basic theorem for convergence of two-level methods.

THEOREM 1. *Given a finite-dimensional Hilbert space H, an inner product $(\cdot,\cdot)$ on H and two subspaces H_1, H_2 of H such that*

$$H_1 \cap H_2 = \{0\},$$

there exists

$$\gamma = \gamma(H_1, H_2) \in (0,1)$$

such that for all $\underline{h}_1 \in H_1$ and $\underline{h}_2 \in H_2$ the following strengthened C.B.S.*-inequality holds:*

$$|(\underline{h}_1, \underline{h}_2)| \leq \gamma \, \|\underline{h}_1\| \; \|\underline{h}_2\| \tag{1}$$

where the norm is induced by the inner product

$$\|\underline{h}\| = \sqrt{(\underline{h}, \underline{h})}.$$

Proof. Let

$$\gamma = \sup_{\underline{h}_1 \in H_1, \underline{h}_2 \in H_2} \frac{|(\underline{h}_1, \underline{h}_2)|}{\|\underline{h}_1\| \, \|\underline{h}_2\|}.$$

The ordinary C.B.S.-inequality tells us that $\gamma \leq 1$. Assume that

$$\gamma = 1.$$

Then there exist sequences

$$\{\underline{h}_1^{(n)}\}_{n=1}^{\infty} \subset H_1, \quad \{\underline{h}_2^{(n)}\}_{n=1}^{\infty} \subset H_2$$

such that $\|\underline{h}_1^{(n)}\| = \|\underline{h}_2^{(n)}\| = 1$ for all n, and $(\underline{h}_1^{(n)}, \underline{h}_2^{(n)}) \to 1$. Since in finite-dimensional spaces the unit sphere is compact, we can assume (by taking subsequences, if necessary) that there exist $\underline{h}_1 \in H_1$ and $\underline{h}_2 \in H_2$ such that

$$\underline{h}_1^{(n)} \to \underline{h}_1 \quad \text{and} \quad \underline{h}_2^{(n)} \to \underline{h}_2,$$

and such that

$$\|\underline{h}_1\| = \|\underline{h}_2\| = 1. \tag{2}$$

We then find

$$1 = (\underline{h}_1, \underline{h}_2) = \|\underline{h}_1\| \, \|\underline{h}_2\| \, .$$

From the disjointness of H_1 and H_2, however, it follows that this last equality can only be satisfied for $\underline{h}_1 = \underline{h}_2 = \underline{0}$. As this contradicts (2), we conclude that

$$\gamma < 1. \qquad \square$$

COROLLARY 1. *Let M be a symmetric positive definite matrix, and let U and V be disjoint vector subspaces of the space on which M operates. Then there exists a $\gamma \in [0, 1)$ such that*

$$(\underline{u}^t M \underline{v})^2 \leq \gamma^2 \underline{u}^t M \underline{u}\, \underline{v}^t M \underline{v} \quad \forall \underline{u} \in U, \ \underline{v} \in V. \tag{3}$$

Remark 1. The finite-dimensionality of the subspaces H_1 and H_2 can be somewhat relaxed: the above theorem also holds if one space is finite dimensional and the other closed.

The following theorem is the basis for the two-level (or multilevel) methods.

THEOREM 2. *Let M be a symmetric, positive definite 2×2 block matrix*

$$M = \begin{pmatrix} B & C^t \\ C & A \end{pmatrix}.$$

Let U and V be the spaces of vectors with only nonzero first and second components, respectively, i.e., $\underline{u} \in U$ is of the form $\underline{u} = \binom{\underline{u}_1}{0}$ and $\underline{v} \in V$ is of the form $\underline{v} = \binom{0}{\underline{v}_2}$. If $\gamma \in [0, 1)$ is such that

$$(\underline{u}^t M \underline{v})^2 \leq \gamma^2 \underline{u}^t M \underline{u}\, \underline{v}^t M \underline{v} \quad \forall \underline{u} \in U, \ \underline{v} \in V,$$

and if we write

$$S = A - CB^{-1}C^t,$$

then

$$\underline{v}_2^t S \underline{v}_2 \geq (1-\gamma^2)\underline{v}_2^t A \underline{v}_2 \tag{4}$$

for all $\underline{v}_2$ *of the appropriate dimension.*

Proof. Note that both A and B are invertible; rewriting the definition of γ then gives

$$\gamma^2 \geq \sup_{\underline{u}_1, \underline{v}_2} \frac{(\underline{u}_1^t C^t \underline{v}_2)^2}{\underline{u}_1^t B \underline{u}_1 \, \underline{v}_2^t A \underline{v}_2};$$

since for all symmetric positive definite matrices K and vectors $\underline{u}$

$$\sup_{\underline{v}} \frac{(\underline{u}^t \underline{v})^2}{\underline{v}^t K \underline{v}} = \underline{u}^t K^{-1} \underline{u},$$

we find

$$\gamma^2 \geq \sup_{\underline{v}_2} \frac{\underline{v}_2^t C B^{-1} C^t \underline{v}_2}{\underline{v}_2^t A \underline{v}_2}.$$

Therefore, for all $\underline{v}_2$ of the appropriate dimension

$$\gamma^2 \underline{v}_2^t A \underline{v}_2 \geq \underline{v}_2^t C^t B^{-1} C \underline{v}_2,$$

which is just (4). □

Remark 2. The nonnegative definiteness of $C^t B^{-1} C$ implies that

$$\underline{v}_2^t A \underline{v}_2 \geq \underline{v}_2^t S \underline{v}_2 \tag{5}$$

for all vectors $\underline{v}_2$.

3. Local estimates. The theorems in the previous section ensure the convergence of two-level methods using properties of the global matrix. In this section we will present tools for the computation of the quantity γ using only local analysis, i.e., by inspection of the element matrices. Bounds obtained for γ on the element matrices then carry over to the global matrix (see also the next section; this argument was employed for proving spectral equivalence in [2]).

One essential point that has to be taken into account when considering element matrices is the fact that, even if the global (finite element) matrix is positive definite, element matrices tend to be only semidefinite. Thus we need a new proof for the existence of the γ in the strengthened C.B.S. inequality for the case of semidefinite matrices. Also, we will give a necessary condition on the null space $\mathcal{N}(M)$ of the matrix M for a $\gamma < 1$ to exist.

THEOREM 3. *Let* M *be a symmetric, positive semidefinite matrix and let* U *and* V *be two disjoint subspaces of the space on which* M *operates such that*

$$V \supset \mathcal{N}(M). \tag{6}$$

Then there exists a $\gamma \in [0,1)$ *satisfying*

$$(\underline{u}^t M \underline{v})^2 \leq \gamma^2 \underline{u}^t M \underline{u}\, \underline{v}^t M \underline{v} \quad \forall \underline{u} \in U, \quad \underline{v} \in V. \tag{7}$$

Proof. A γ can be found from Corollary 1, based on the subspace U and the subspace of V containing the vectors that are orthogonal to the null space of M. However, this γ also satisfies the defining relation (7) even if we include in V any vectors that may lie in $\mathcal{N}(M)$, as for those vectors both sides of the equation are zero. □

THEOREM 4. *Let M be a symmetric, positive semidefinite* 2×2 *block matrix*

$$M = \begin{pmatrix} B & C^t \\ C & A \end{pmatrix}$$

with B invertible, let U and V be the spaces of vectors with only nonzero first and second block components, respectively, and assume that (6) *holds. Then the* $\gamma \in [0,1)$ *guaranteed by Theorem* 3 *also satisfies*

$$\underline{v}_2^t S \underline{v}_2 \geq (1-\gamma^2)\underline{v}_2^t A \underline{v}_2 \tag{8}$$

for all $\underline{v}_2$, *where*

$$S = A - CB^{-1}C^t.$$

Proof. Obviously, if $A\underline{v}_2 = \underline{0}$, then (8) is satisfied for all γ. Now let $\gamma \in [0,1)$ be such that it satisfies (7). Then

$$\begin{aligned} \gamma^2 &\geq \sup_{\substack{\underline{u}_1 \neq \underline{0} \\ \underline{v}_2 \notin \mathcal{N}(A)}} \frac{(\underline{u}_1^t C^t \underline{v}_2)^2}{\underline{u}_1^t B \underline{u}_1\, \underline{v}_2^t A \underline{v}_2} \\ &= \sup_{\underline{v}_2 \notin \mathcal{N}(A)} \frac{\underline{v}_2^t C^t B^{-1} C \underline{v}_2}{\underline{v}_2^t A \underline{v}_2}. \end{aligned}$$

This gives the desired result. □

LEMMA 1. *Let M be as in the above theorem, but without* (6) *necessarily holding, and suppose there exists an* $\alpha > 0$ *such that*

$$\text{for all } \underline{v}: \quad \underline{v}^t S \underline{v} \geq \alpha \underline{v}^t A \underline{v}.$$

Then

$$\mathcal{N}(M) = \left\{ \begin{pmatrix} 0 \\ \underline{v} \end{pmatrix} : C^t \underline{v} = \underline{0}, \underline{v} \in \mathcal{N}(A) \right\}. \tag{9}$$

Proof. Let $\underline{x} = \binom{\underline{u}}{\underline{v}}$ be a vector such that $M\underline{x} = 0$. Then

$$B\underline{u} + C^t \underline{v} = 0 \Rightarrow \underline{u} = -B^{-1}C^t \underline{v}$$

and

$$C\underline{u} + A\underline{v} = 0 \Rightarrow S\underline{v} = A\underline{v} - CB^{-1}C^t \underline{v} = 0.$$

From the existence of a positive α we see that

$$S\underline{v} = 0 \Rightarrow A\underline{v} = 0.$$

Hence, $\underline{v} \in \mathcal{N}(A)$. Furthermore, from

$$\underline{0} = C\underline{u} + A\underline{v} = C\underline{u} \Rightarrow \underline{0} = C\underline{u} = -CB^{-1}C^t\underline{v}$$

$$\Rightarrow \underline{0} = (C^t\underline{v})^t B^{-1}(C^t\underline{v}) = (B\underline{u})^t B^{-1}(B\underline{u})$$

it follows that

$$\underline{u} = \underline{0} \quad \text{and} \quad C^t\underline{v} = \underline{0}.$$

Conversely, assume that $\underline{v} \in \mathcal{N}(A)$. Then from

$$0 \leq \underline{v}^t S\underline{v} \leq \underline{v}^t A\underline{v} = 0$$

it follows that $\underline{v} \in \mathcal{N}(S)$, and by choosing $\underline{u} = -B^{-1}C^t\underline{v}$ it follows that $(\underline{u}, \underline{v})^t \in \mathcal{N}(M)$.

Finally, from $C\underline{u} + A\underline{v} = 0$ and $\underline{v} \in \mathcal{N}(A)$ it follows that $C\underline{u} = 0$. Since $\underline{u} = -B^{-1}C^t\underline{v}$, we get

$$-CB^{-1}C^t\underline{v} = 0,$$

which implies

$$(C^t\underline{v})^t B^{-1}C^t\underline{v} = 0,$$

that is, $C^t\underline{v} = 0$ and $\underline{u} = 0$. Hence, $\mathcal{N}(M)$ is given by (9). □

From this lemma we immediately find that the sufficient condition (6) for the existence of a $\gamma \in [0, 1)$ is also necessary.

THEOREM 5. *Suppose M is a symmetric positive semidefinite 2×2 partitioned matrix where the $(2, 2)$ diagonal block is invertible. Suppose U and V are the corresponding subspaces of vectors with nonzero first and second components only. Then a $\gamma \in [0, 1)$ satisfying (7) exists if and only if the null space of M is completely contained in V.*

Proof. In Theorem 3 we proved that a $\gamma \in [0, 1)$ exists provided that $\mathcal{N}(M) \subset V$. For the converse assertion note that for the given space V we have

$$\mathcal{N}(M) \subset V \Leftrightarrow \mathcal{N}(M) = \left\{ \begin{pmatrix} \underline{0} \\ \underline{v} \end{pmatrix} : C^t\underline{v} = \underline{0}, v \in \mathcal{N}(A) \right\}. \tag{10}$$

From Lemma 1 it follows that the null space of M indeed has this form if a $\gamma \in [0, 1)$ exists. □

For the actual computation of γ in the case of a semidefinite matrix satisfying the conditions of Theorem 4, we consider the generalised eigenvalue problem

$$S\underline{u} = \lambda A\underline{u}.$$

In order to deflate out the vectors in the null space of S we can use the following lemma.

LEMMA 2. *If two matrices A and B have the same nontrivial null space, all solutions $\lambda \neq 0$ of the generalised eigenvalue problem*

$$A\underline{x} = \lambda B\underline{x}$$

are also solutions of

$$\tilde{A}\underline{\tilde{x}} = \lambda \tilde{B}\underline{\tilde{x}},$$

where $\tilde{A}$ and $\tilde{B}$ are formed by deleting for some (but not arbitrary) i the ith row and column from A and B.

Proof. Let the pair $(\underline{x}, \lambda)$ be a solution of the generalised eigenvalue problem $A\underline{x} = \lambda B\underline{x}$, and let $\underline{e} \in \mathcal{N}(A)$. Write $\underline{x} = \underline{\hat{x}} + (x_i/e_i)\underline{e}$ where i is such that the ith component e_i of e is nonzero (similarly x_i is the ith component of $\underline{x}$), and let $\hat{A}$ and $\hat{B}$ be the matrices formed by zeroing the ith row and column of A and B. Then, for all $j \neq i$,

$$(A\underline{x})_j = (A\underline{\hat{x}})_j = (\hat{A}\underline{\hat{x}})_j$$

and, similarly,

$$(B\underline{x})_j = (B\underline{\hat{x}})_j = (\hat{B}\underline{\hat{x}})_j.$$

Now, leaving out the ith components altogether gives the desired result. $\square$

As was indicated at the end of the previous section, where we presented a simple two-level method, we need the quantity γ to find eigenvalue bounds for the preconditioning of the Schur complement $S = A - CB^{-1}C^t$ by the principal minor A of M. In some cases where the subspaces U and V do not satisfy condition (6), the next lemma shows that we can compare the Schur complement S of the semidefinite matrix M and the principal minor $\bar{A}$ of a matrix $\bar{M}$ obtained after a suitable transformation of M, provided $\bar{M}$ is such that a $\gamma \in [0, 1)$ does exist. Such a situation arises, for instance, in multilevel methods for finite elements, when the standard nodal basis element stiffness matrix is transformed into the so-called two-level hierarchical basis element stiffness matrix, for which a $\gamma \in [0, 1)$ exists.

LEMMA 3. *Suppose the 2×2 block matrix*

$$M = \begin{pmatrix} B & C \\ C^t & A \end{pmatrix}$$

is semidefinite with B invertible, and that there exists a matrix

$$J = \begin{pmatrix} I & K \\ 0 & I \end{pmatrix}$$

such that for the matrix

$$\bar{M} = J^t M J = \begin{pmatrix} B & \bar{C} \\ \bar{C}^t & \bar{A} \end{pmatrix}$$

the spaces U and V satisfy (6). Then there exists a $\gamma \in [0, 1)$ satisfying

$$\underline{u}_2^t \bar{A} \underline{u}_2 \geq \underline{u}_2^t S \underline{u}_2 \geq (1 - \gamma^2) \underline{u}_2^t \bar{A} \underline{u}_2 \quad \forall \underline{u}_2. \tag{11}$$

Proof. Simple computation shows that

$$S = A - C^t B^{-1} C$$

is equal to

$$\bar{S} = \bar{A} - \bar{C}^t B^{-1} \bar{C}.$$

Relation (11) now follows from this and from relations (5) and (4), which are guaranteed for $\bar{M}$ and $\bar{S}$. □

4. Finite element analysis. Direct application of the strengthened C.B.S. inequality to multilevel finite element spaces is impractical because the component spaces H_1 and H_2 depend on the discretisation, and therefore γ might also; that is, γ might tend to 1 as the mesh is refined. The purpose of this section is to show that this is not the case: $\gamma < 1$ uniformly in the discretization.

Consider now an elliptic bilinear form $a(\cdot,\cdot)$:

$$a(u,v) = \int_\Omega \sum_{i,j} a_{ij}(\underline{x}) \frac{\partial}{\partial x_i} u \frac{\partial}{\partial x_j} v\, dx, \qquad u,v \in H^1(\Omega) \tag{12}$$

with

$$u, v = 0 \quad \text{on } \Gamma_D \subset \Gamma = \partial\Omega,$$

$$\text{meas}(\Gamma_D) \neq 0.$$

We assume that Ω is a plane polygon. Let τ_1 be an initial triangulation of Ω. We associate with τ_1 a discrete finite element space V_1 of piecewise polynomials on the triangles of τ_1, which are continuous on $\bar{\Omega}$ and which vanish on Γ_D.

We generate a sequence of triangulations $\{\tau_k\}_{k=1}^{\ell}$ that form successive refinements of τ_1, and let V_k be the corresponding finite element spaces of piecewise polynomials on the triangles of τ_k, continuous on $\bar{\Omega}$ and vanishing on Γ_D. We also require that

$$V_k \subset V_{k+1}, \qquad k = 1, 2, \cdots, \ell - 1.$$

This is, for instance, the case where the degree of the piecewise polynomials is the same on each level.

In addition to the spaces V_k we define the hierarchical spaces $\tilde{V}_k$ as follows. Let $\{\phi_i^{(k)}\}$ be a nodal basis in V_k associated with the set N_k of points at level k, excluding Dirichlet boundary points. Now

$$\tilde{V}_k = \text{Span}\{\phi_i^{(k)}\}_{i: x_i \in N_k \setminus N_{k-1}}.$$

Since functions from $\tilde{V}_k$ are zero at points $x_i \in N_{k-1}$, we conclude that, for every $\phi \in \tilde{V}_k$,

$$\phi|_T = \text{const} \Rightarrow \phi|_T = 0 \quad \text{if } T \in \tau_{k-1}.$$

Using these spaces we define the so-called two-level hierarchical basis element stiffness matrices A_T for elements $T \in \tau_{k-1}$ as

$$A_T = \begin{pmatrix} A_{T;11} & A_{T;12} \\ A_{T;21} & A_{T;22} \end{pmatrix}.$$

Here we use the restricted bilinear form

$$a_T(u,v) := \int_T \sum_{i,j} a_{ij}(\underline{x}) \frac{\partial}{\partial x_i} u \frac{\partial}{\partial x_j} v \, dx \tag{13}$$

to define the components of A_T as follows:

$$A_{T;11} = (a_T(\phi_i^{(k)}, \phi_j^{(k)}))_{x_i, x_j \in (N_k \setminus N_{k-1}) \cap T},$$
$$A_{T;12} = (a_T(\phi_i^{(k)}, \phi_j^{(k-1)}))_{\substack{x_i \in (N_k \setminus N_{k-1}) \cap T \\ x_j \in N_{k-1} \cap T}} = A_{T;12}^t,$$

and

$$A_{T;22} = (a_T(\phi_i^{(k-1)}, \phi_j^{(k-1)}))_{x_i, x_j \in N_{k-1} \cap T}.$$

Note that $A_{T;22}$ is the ordinary nodal basis element stiffness matrix at level $k-1$. Furthermore, we note that $A_{T;11}$ is positive definite because the only constant vector that can be constructed using functions from the index set $N_k \setminus N_{k-1}$ is the zero vector.

LEMMA 4. *The element Schur complement S_T of A_T given by*

$$S_T = A_{T;22} - A_{T;21} A_{T;11}^{-1} A_{T;12} \tag{14}$$

has the null space

$$\mathcal{N}(S_T) = \text{Span}\{(1, 1, \cdots, 1)^t\}.$$

Proof. Let $S_T \underline{v}_T = 0$ for some $\underline{v}_T \in V_{k-1}$. Then

$$0 = \underline{v}_T^t S_T \underline{v}_T = \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix}^t A_T \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix}$$

with $\underline{u}_T = -A_{T;11}^{-1} A_{T;12} \underline{v}_T$. Now let u and v be any two functions such that their restrictions to T, $u|_T$ and $v|_T$, have coefficient vectors $\underline{u}_T$ and $\underline{v}_T$, respectively. Let w be such that $w|_T = u|_T + v|_T$. Then

$$a_T(w,w) = \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix}^t A_T \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix} = 0,$$

so $w|_T =$ constant. But by construction we have

$$w(x_i) = v(x_i), \qquad x_i \in N_{k-1},$$

which shows that

$$\underline{v}_T = c \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

for some $c \in \mathbb{R}$. $\square$

LEMMA 5. *The matrices S_T and $A_{T;22}$ have the same null space.*

Proof. In Lemma 4 we proved that the null space of $S|_T$ consists of the constant vectors; since

$$A^{(k)}_{T;22} = A^{(k-1)}_T$$

is a nodal basis stiffness matrix it has the same null space. □

We can now prove the uniform strengthened C.B.S. inequality.

THEOREM 6. *Let the local ellipticity condition number be defined by*

$$\sigma = \sup_{T \in \tau_1} \frac{\mu_2(T)}{\mu_1(T)}$$

where

$$\mu_1(T)\,|\xi|^2 \leq \sum_{i,j} a_{ij}(x)\xi_i\xi_j \leq \mu_2(T)\,|\xi|^2 \tag{15}$$

for all $\xi \in \mathbb{R}^2$ and $x \in T$. We assume that $\mu_1(T)$ and $\mu_2(T)$ are the largest and smallest number, respectively, satisfying (15). *Note that $\sigma \geq 1$. Then there exists a constant $\gamma \in [0,1)$, depending only on the geometry of the initial triangulation τ_1 and σ, such that*

$$a(u,v) \leq \gamma\sqrt{a(u,u)}\,\sqrt{a(v,v)}$$

for all $v \in V_{k-1}$, $u \in \tilde{V}_k$, and $k \geq 1$.

Proof. The proof consists of three parts. First we prove the existence of a $\gamma \in [0,1)$ for each level and each element on that level, then we prove the existence of γ uniformly on each level, and finally the existence uniformly over all levels.

Let T be an element of the triangulation τ_{k-1}. Choose $v \in V_{k-1}$ and $u \in \tilde{V}_k$. We have

$$v|_T = \sum_{x_i \in N_{k-1} \cap T} v_i \phi_i^{(k-1)}, \qquad v_i = v(x_i)$$

and

$$u|_T = \sum_{x_i \in (N_k \setminus N_{k-1}) \cap T} u_i \phi_i^{(k)}, \qquad u_i = u(x_i).$$

Let us denote by $\underline{u}_T$ and $\underline{v}_T$ the coefficient vectors of $u|_T$ and $v|_T$, respectively, in the above expansions. By (13) we have

$$a_T(u,v) = \underline{u}_T^t\, A_{T;1,2}\underline{v}_T$$

where $A_{T;1,2}$ is the off-diagonal block of the two-level hierarchical basis element stiffness matrix A_T.

Consider the generalised eigenvalue problem

$$S_T\underline{v}_T = \lambda A_T^{(k-1)}\underline{v}_T, \qquad \underline{v}_T \notin \mathcal{N}(S_T) = \mathcal{N}(A_T^{(k-1)}).$$

Since

$$S_T = A_T^{(k-1)} - A_{T;21}A_{T;11}^{-1}A_{T;12}$$

and S_T is positive definite for vectors not in its null space, we have that any generalized eigenvalue λ satisfies $0 < \lambda \leq 1$. Let λ_T be the smallest eigenvalue. Then

$$\lambda_T = 1 - \gamma_T^2$$

where

$$\gamma_T^2 = \sup_{\underline{v}_T \notin \mathcal{N}(A_T^{(k-1)})} \frac{\underline{v}_T^t A_{T;21} A_{T;11}^{-1} A_{T;12} \underline{v}_T}{\underline{v}_T^t A_T^{(k-1)} \underline{v}_T}. \tag{16}$$

We shall show that $\gamma_T \in [0,1)$ and that γ_T is related to the local strengthened C.B.S. inequality. Indeed there exists a constant $\gamma_T' \in [0,1)$ such that

$$|a_T(u,v)| \leq \gamma_T' \sqrt{a_T(u,u)} \sqrt{a_T(v,v)}$$

because

$$\tilde{V}^{(k)}|_T \cap V^{(k-1)}|_T = \{0\} \quad \text{and} \quad \mathcal{N}(A_T) \subset V^{(k-1)}$$

by construction. In matrix form this translates to

$$\left|\underline{u}_T^t A_{T;12} \underline{v}_T\right| \leq \gamma_T' \sqrt{\underline{u}_T^t A_{T;11} \underline{u}_T} \sqrt{\underline{v}_T^t A_T^{(k-1)} \underline{v}_T}.$$

Choosing

$$\underline{u}_T = A_{T;11}^{-1} A_{T;12} \underline{v}_T,$$

we get

$$\underline{v}_T^t A_{T;21} A_{T;11}^{-1} A_{T;12} \underline{v}_T \leq {\gamma_T'}^2 \underline{v}_T^t A_T^{(k-1)} \underline{v}_T$$

for all $\underline{v}_T$, whence

$$\gamma_T^2 \leq {\gamma_T'}^2 < 1.$$

We will next prove that γ_T can be estimated uniformly on T. Denote the element matrices corresponding to the Laplace operator by $\hat{A}_T$, with elements $\hat{A}_{T;i,j}$ and Schur complements $\hat{S}_T$ defined analogous to (14). The Schur complements S_T and $\hat{S}_T$ are related by

$$\begin{aligned}
\underline{v}_T^t S_T \underline{v}_T &= \inf_{\underline{u}_T} \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix}^t A_T \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix} \\
&\geq \mu_1(T') \inf_{\underline{u}_T} \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix}^t \hat{A}_T \begin{pmatrix} \underline{u}_T \\ \underline{v}_T \end{pmatrix} \\
&= \mu_1(T') \underline{v}_T^t \hat{S}_T \underline{v}_T
\end{aligned}$$

where μ_1 is defined in (15) and where T' is an element from τ_1 containing T. Similarly, we find that A_T and $\hat{A}_T$ are related by

$$\underline{v}_T^t A_T^{(k-1)} \underline{v}_T \leq \mu_2(T') \underline{v}_T^t \hat{A}_T^{(k-1)} \underline{v}_T.$$

Deriving from (16) the equation

$$1-\gamma_T^2 = \inf_{\underline{v}_T \notin \mathcal{N}(A_T^{(k-1)})} \frac{\underline{v}_T^t S_T \underline{v}_T}{\underline{v}_T A_T^{(k-1)} \underline{v}_T},$$

using the previous two results, and noting that the null spaces of $A_T^{(k-1)}$ and $\tilde{A}_T^{(k-1)}$ agree, we obtain

$$\begin{aligned} 1-\gamma_T^2 &\geq \frac{\mu_1(T')}{\mu_2(T')} \inf_{\underline{v}_T \notin \mathcal{N}(\hat{A}_T^{(k-1)})} \frac{\underline{v}_T^t \hat{S}_T \underline{v}_T}{\underline{v}_T^t \hat{A}_T^{(k-1)} \underline{v}_T} \\ &\geq \frac{1}{\sigma}(1-\hat{\gamma}_T^2). \end{aligned}$$

Here, $\hat{\gamma}_T \in [0,1)$ is the quantity for $\hat{A}_T$ analogous to γ_T for A_T. From this inequality we conclude that

$$\gamma_T^2 \leq 1 - \frac{1}{\sigma}(1-\hat{\gamma}_T^2) \in [0,1).$$

Now, for the Laplace operator, $\hat{\gamma}_T^2$ depends only on the geometry of the triangulation and the polynomials used. Hence, if on all levels the triangles are geometrically similar to some fixed set of elements $\tau_0 \supset \tau_1$, then there exists a $\hat{\gamma} \in [0,1)$ such that

$$\hat{\gamma}_T^2 \leq \hat{\gamma}^2$$

for all $T \in T_{k-1}$ and all $k \geq 1$.

To prove the global strengthened C.B.S. inequality, note that

$$\begin{aligned} |a(u,v)| &= \sum_{T\in\tau_{k-1}} |a_T(u,v)| \\ &\leq \sum_{T\in\tau_{k-1}} \gamma_T \sqrt{a_T(u,u)}\sqrt{a_T(v,v)} \\ &\leq \sup_{T\in\tau_{k-1}} \gamma_T \cdot \sum_{T\in\tau_{k-1}} \sqrt{a_T(u,u)}\sqrt{a_T(v,v)} \\ &\leq \gamma \left(\sum_{T\in\tau_{k-1}} a_T(u,u)\right)^{1/2} \left(\sum_{T\in\tau_{k-1}} a_T(v,v)\right)^{1/2} \\ &= \gamma\sqrt{a(u,u)}\sqrt{a(v,v)}, \end{aligned}$$

which completes the proof. $\square$

5. Multilevel solution methods. Theorem 2 is the basis for the proof of order 1 convergence of two-level and multilevel solution methods for finite element problems. By way of example we consider a simple two-level method for the solution of

$$\begin{pmatrix} B & C^t \\ C & A \end{pmatrix} \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix} = \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix}.$$

In a finite element context such a partitioning arises if we have two nested grids $\Omega^{(2h)} \subset \Omega^{(h)}$ with corresponding finite element spaces $V^{(2h)} \subset V^{(h)}$ associated with the nodes of these grids. A two-level splitting of the finite element space is attained by letting the second block component of the coefficient vector correspond to functions from $V^{(2h)}$, and the first to functions from $V^{(h)}$ on nodes in $\Omega^{(h)} - \Omega^{(2h)}$. In this case the theory from §2 is directly applicable; if the matrix is simply the stiffness matrix of $V^{(h)}$ on $\Omega^{(h)}$, we need in addition Lemma 3 to apply that theory.

We formulate a simple stationary iterative method for the above problem:

$$\text{(1)} \qquad \text{solve} \begin{pmatrix} B & 0 \\ C & I \end{pmatrix} \begin{pmatrix} \underline{z}_1 \\ \underline{z}_2 \end{pmatrix} = \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix},$$

$$\text{(2)} \qquad \text{iterate} \begin{array}{l} \underline{z}_2^{(0)} = \underline{z}_2, \\ A(\underline{z}_2^{(n+1)} - \underline{z}_2^{(n)}) = \underline{y}_2 - S\underline{z}_2^{(n)}, \end{array}$$

$$\text{(3)} \qquad \text{solve} \begin{pmatrix} I & B^{-1}C^t \\ 0 & I \end{pmatrix} \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix} = \begin{pmatrix} \underline{z}_1 \\ \underline{z}_2^{(n_{\max})} \end{pmatrix}.$$

THEOREM 7. *The above iterative method takes a number of iterations that depends only on the value of γ, and not explicitly on the dimension of the matrix.*

Proof. The iteration matrix $I - A^{-1}S$ for the iterative process (2) has a spectral radius bounded by γ^2. □

This iterative method uses only a two-level splitting of the finite element space, thereby necessitating exact solution of systems with the finite element matrix A. We conclude this section by indicating how a multilevel solution method can be derived that does not suffer from this disadvantage. The procedure presented here was developed in [6] and [5] and used also in [3]; we refer to these papers for the details.

Consider then a multilevel splitting of the coefficient matrix

$$A^{(k)} = \begin{pmatrix} B^{(k)} & C^{(k)^t} \\ C^{(k)} & A^{(k-1)} \end{pmatrix}$$

and an incomplete factorization preconditioner based on this splitting,

$$M^{(k)} = \begin{pmatrix} B^{(k)} & 0 \\ C^{(k)} & Z^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B^{(k)^{-1}}C^{(k)^t} \\ 0 & I \end{pmatrix},$$

where $Z^{(k-1)}$ is intended as an approximation to

$$S^{(k-1)} = A^{(k-1)} - C^{(k)}B^{(k)^{-1}}C^{(k)^t}.$$

For a two-level preconditioner we would take $Z^{(k-1)} = A^{(k-1)}$; a true multilevel preconditioner is arrived at by the definition

$$Z^{(k-1)} = \left(I - P\left(M^{(k-1)^{-1}}S^{(k-1)}\right)\right) S^{(k-1)^{-1}}$$

where P is some polynomial satisfying

$$P(0) = 1.$$

Note that solving a system with $M^{(k)}$ now involves solving at least one system with $M^{(k-1)}$. Analysis of the condition number of $M^{(k)^{-1}}A^{(k)}$ now is based on a fixed-point argument, using the polynomial and the value of γ obtained for the splitting.

In [6] it was proved that for polynomial degrees ≥ 2 such a method gives a condition number of order 1, independent of the number of levels. This holds even if $B^{(k)}$ is approximated, for instance, by a diagonal matrix; see [5]. The application of this technique to certain finite difference meshes was demonstrated in [3].

It is not hard to see that, if the splitting of the finite element space is induced by meshes of ever doubling meshwidth, we can allow polynomial degrees up to 3 for two-dimensional problems, and up to 7 for three-dimensional ones, while still obtaining a preconditioner for which the work per iteration is proportional to the number of unknowns. Thus, such multilevel preconditioners give methods of optimal order of computational labour.

Acknowledgment. The authors acknowledge constructive criticism by Jurgen Aarden and Professor Steve McCormick, which helped to improve the presentation of this paper.

REFERENCES

[1] O. Axelsson, *On multigrid methods of the two-level type*, in Multigrid methods, Proceedings, Köln-Porz, W. Hackbusch and U. Trottenberg, eds., Lecture Notes in Math., 960, 1982, pp. 352–367.

[2] O. Axelsson and A. Barker, *Finite Element Solution of Boundary Value Problems. Theory and Computation*, Academic Press, Orlando, FL, 1984.

[3] O. Axelsson and V. Eijkhout, *The nested recursive two-level factorization method for nine-point difference matrices*, Tech. Report 8936, Mathematisch Instituut, Nijmegen, 1989, SIAM J. Sci. Statist. Comput., 12 (1991), to appear.

[4] O. Axelsson and I. Gustafsson, *Preconditioning and two-level multigrid methods of arbitrary degree of approximation*, Math. Comp., 40 (1983), pp. 219–242.

[5] O. Axelsson and P. Vassilevski, *Algebraic multilevel preconditioning methods*, II, Tech. Report 1988-15, Institute for Scientific Computing, the University of Wyoming, Laramie, WY, 1988, SIAM J. Numer. Anal., 27 (1990), pp. 1569–1590.

[6] ———, *Algebraic multilevel preconditioning methods*, I, Numer. Math, 56 (1989), pp. 157–177.

[7] R. Bank and T. Dupont, *Analysis of a two-level scheme for solving finite element equations*, Tech. Report CNA-159, Center for Numerical Analysis, University of Texas, Austin, TX, 1980.

[8] R. Bank, T. Dupont, and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427–458.

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

[10] ———, *The convergence rate of a multigrid method with Gauss-Seidel relaxation for the Poisson equation*, in Multigrid Methods, Proceedings, Köln-Porz, 1981, W. Hackbusch and U. Trottenberg, eds., Lecture Notes in Math., 960, 1982, pp. 368–387.

[11] J. Bramble, R. Ewing, J. Pasciak, and A. Schatz, *A preconditioning technique for the efficient solution of problems with local grid refinement*, Comput. Methods Appl. Mech. Engrg., 67 (1988), pp. 149–159.

[12] J. Bramble, J. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with aplication to domain decomposition and multigrid*, Tech. Report 90-39, Mathematical Science Institute of Cornell University, Ithaca, NY, June 1990.

[13] R. Ewing, R. Lazarov, J. Pasciak, and P. Vassilevski, *Finite element methods for parabolic problems with time steps variable in space*, Tech. Report, Institute for Scientific Computing, University of Wyoming, Laramie, WY, 1989.

[14] R. Ewing, R. Lazarov, and P. Vassilevski, *Local refinement techniques for elliptic problems on cell-centered grids*, Tech. Report 1988-16, Institute for Scientific Computing, University of Wyoming, Laramie, WY, 1988.

[15] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, New York, 1985.

[16] J. Maitre and F. Musy, *The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems*, in Multigrid methods, Proceedings, Köln-Porz, 1981, W. Hackbusch and U. Trottenberg, eds., Lecture Notes in Math., 960, 1982, pp. 535–544.

[17] J. Mandel and S. McCormick, *Iterative solutions of elliptic equations with refinement: the two-level case*, in Domain Decomposition Methods, T. Chan, R Glowinski, J.Périaux, and O. Widlund, eds., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1989, pp. 81–92.

[18] S. McCormick, *Fast adaptive composite grid (fac) methods: theory for the variational case*, Comput. Suppl., 5 (1984), pp. 115–121.

[19] S. McCormick and J. Thomas, *The fast adaptive composite grid (fac) method for elliptic equations*, Math. Comp., 46 (1986), pp. 439–456.

[20] P. Vassilevski, *Nearly optimal iterative methods for solving finite element elliptic equations based on the multilevel splitting of the matrix*, Tech. Report 1989-9, Institute for Scientific Computing, University of Wyoming, Laramie, WY, 1989.

[21] H. Yserentant, *On the multilevel splitting of finite element spaces*, Numer. Math, 49 (1986), pp. 379–412.