CHAPTER 15

Algebraic Multilevel Preconditioning Methods, III

O. Axelsson*
P. S. Vassilevski**

ABSTRACT

A new variant of the algebraic multilevel iteration method studied by the authors in previous publications is presented. In the present method it is not required to estimate the accuracy with which the block matrices corresponding to the, at every refined level, added meshpoints are approximated. As can be expected this lack of information slows down the rate of convergence somewhat. However, the rate is still of optimal order, under the same condition as for the previous methods, namely that $\nu > (1-\gamma^2)^{-1/2}$, where ν is the degree of the shifted and normalized Chebyshev matrix polynomials, used to approximate the arising Schur complements and where γ is the constant in the strengthened Cauchy-Bunyakowski-Schwarz inequality, corresponding to the bilinear form and the function spaces spanned by the nodal basis functions in the vertices and the new set of hierarchical basis functions in the edge nodepoints.

* Faculty of Mathematics and Informatics, Catholic University, Toernooiveld, 6525 ED Nijmegen, The Netherlands.

** Center of Computer Science and Technology, Bulgarian Academy of Sciences, Acad. G.Bontchev str. block 25a, 1113 Sofia, Bulgaria. The research of the second author was partly supported by Bulgarian Committee of Science under Grant # 55-26-3-87.

1. INTRODUCTION

In previous publications (see [2] and [3]) the authors have presented a multilevel iteration method to solve second order self-adjoint elliptic problems with a computational complexity proportional to the number of meshpoints on the finest level. This optimality property requires only that the coefficients of the differential operator are smooth within elements of the coarsest mesh but they can be discontinuous between elements. In particular they are valid for a domain with a reentrant corner. The preconditioner is based on an approximate factorization of matrices partitioned in a two by two block form where the inverses of the arising Schur complements are approximated by certain matrix polynomials such as shifted and normalized Chebyshev polynomials, involving the inverse of the preconditioner on the next coarser mesh level, and the Schur complement itself (version (i)) or the stiffness matrix on the next coarser level (version (ii)). Hence the preconditioner is defined by recursion from a finer level to the next coarser and on the coarsest mesh it is assumed that the matrix is solved by some direct or iterative method to full (or sufficient) precision.

The best choice of polynomials for fastest rate of convergence, the Chebyshev polynomials requires the estimation of a certain parameter (the lower bound of the eigenvalues of the matrix in the polynomial). This parameter depends on γ, the problem parameter in the strengthened Cauchy-Bunyakowski-Schwarz (C.B.S.) inequality and on a parameter which depends on the accuracy with which the matrices, corresponding to the newest meshpoints and their basisfunctions, are approximated.

While the computation of γ can be done readily from the local finite element matrices, the estimation of the second parameter requires some extra labor (see [3]). In the present method we consider a new method where we have avoided the estimation of the latter parameter. The new method is based on an auxilliary sequence of matrices which is spectrally equivalent to the original sequence of stiffness matrices, and the multilevel preconditioner is computed from the new sequence. Therefore, in addition we require that the matrix blocks corresponding to the newest meshpoints in this latter matrix sequence are approximated to a sufficient accuracy à priori. How this can be done will be shown in the paper.

The remainder of the paper is organized as follows. In Section 2 we present the prerequisites of knowledge required for the presentation of the multilevel methods and in Section 3 the new matrix sequence is presented and the spectral equivalence of it with the given matrix sequence is shown. The definition of the multilevel preconditioner and the estimate of the rate of convergence can be found in Section 4. Finally the new method is illustrated with numerical tests.

While we present the methods only for elliptic boundary value problems in two space dimensions, the method is equally applicable for problems in three space dimensions. The computational complexity is of optimal order in two space dimensions for piecewise linear basisfunctions for any initial triangulation, if this is refined uniformly. In three space dimensions the value of γ is not known for a general division of a domain in tetrahedrons. However for the standard tetrahedrons, for instance it suffices to choose $\nu = 2$ for an optimal rate of convergence.

2. PREREQUISITES

Consider the variational formulation of the second order self-adjoint elliptic operator problem

$$\mathcal{L}u = -\sum_{i,j} \frac{\partial}{\partial x_i}(a_{ij}(x)\frac{\partial u}{\partial x_j})$$

with Dirichlet boundary condition on Γ_0, a subset of the boundary Γ of the polygonal domain $\Omega \subset \mathbb{R}^2$, with meas $(\Gamma_0) \neq 0$ and homogeneous Neuman boundary conditions on $\Gamma \backslash \Gamma_0$. Hence, given $f \in L_2(\Omega)$ we wish to find $u \in H_g^1(\Omega)$ such that

$$a(u,v) = (f,v), \text{ all } v \in H_0^1(\Omega),$$

where $H_0^1(\Omega)$ and $H_g^1(\Omega)$ are the Sobolev spaces of functions with trace equal to zero and g, respectively on Γ_0 and

$$a(u,v) = \int_\Omega \sum_{i,j} a_{ij}(x) \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} dx \tag{2.1}$$

We assume that the matrix $(a_{i,j}(x))$, $i,j = 1,2$ is bounded, symmetric and positive definite on $\overline{\Omega}$ and we let

$$(f,v) = \int_\Omega fv dx.$$

As is well known the bilinear form is then symmetric, bounded and H_0^1-elliptic.

By a standard refining procedure we obtain a sequence of triangulations τ_k, $k = 1,2,...,\ell$. The refinement is done by introducing new nodes in the midpoints of the edges of the triangles of the previous triangulation and connecting them with the opposite vertices. This refinement can be done locally, i.e., it does not have to take place in all triangles.

With any triangulation τ_k we associate the corresponding finite element space V_k of piecewise polynomial functions that are continuous in $\overline{\Omega}$. As basis functions for V_k we choose the standard nodal basis functions, $\{\varphi_i^{(k)}\}_{i=1}^{n_k}$, where n_k is the number of nodepoints in τ_k. Therefore, if N_k is the set of nodes in τ_k and x_j runs over these nodes then

$$\varphi_i^{(k)}(x_j) = \delta_{i,j}, \qquad \text{the Kronecker function.}$$

Once having a basis the following sequence of stiffness matrices can be computed

$$A^{(k)} = \left(a(\varphi_i^{(k)}, \varphi_j^{(k)})\right)_{i,j=1}^{n_k}, \quad k = 1, 2, ..., \ell.$$

By construction we have $N_k \supset N_{k-1}$. Hence at the k-th level the partitioning $N_k \backslash N_{k-1}$ (new nodes) and N_{k-1} (old nodes) of the nodes in N_k can be used. Corresponding to this ordering $A^{(k)}$ takes the following block two by two form

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{bmatrix},$$

where

$$\begin{aligned} A_{11}^{(k)} &= \{a(\varphi_i^{(k)}, \varphi_j^{(k)})\}i,j : x_i, x_j \in N_k \backslash N_{k-1} \\ A_{12}^{(k)} &= \{a(\varphi_i^{(k)}, \varphi_j^{(k)})\}i,j : x_i \in N_k \backslash N_{k-1}, x_j \in N_{k-1} \\ A_{22}^{(k)} &= \{a(\varphi_i^{(k)}, \varphi_j^{(k)})\}i,j : x_i, x_j \in N_{k-1}. \end{aligned} \tag{2.2}$$

As $V_{k-1} \subset V_k$ we may alternatively use the so-called hierarchical two-level basis functions (see [9], [8] and [2])

$$\{\varphi_i^{(k)}, \quad x_i \in N_k \backslash N_{k-1} \text{ and } \varphi_i^{(k-1)}, x_i \in N_{k-1}\}.$$

Then any function $v \in V_k$ can be expanded by using either of these two basis, i.e., we have

$$v(x) = \sum_{i=1}^{n_k} v_i \varphi_i^{(k)},$$

where $v_i = v(x_i)$ and

$$v(x) = \sum_{i: x_i \in N_k \backslash N_{k-1}} \widehat{v}_i \varphi_i^{(k)} + \sum_{i: x_i \in N_{k-1}} v_i \varphi_i^{(k-1)}.$$

This expression defines a mapping $J = J_k$, which transforms any coefficient vector $\widehat{\mathbf{v}} = \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \widehat{\mathbf{v}}_2 \end{bmatrix}$ of the representation of a function $v \in V_k$ with respect to the hierarchical two-level basis to the coefficient vector $\mathbf{v}$ of the representation of v in the nodal basis of V_k, i.e.,

$$\mathbf{v} = J\widehat{\mathbf{v}}.$$

In matrix notation J has the form

$$J = \begin{bmatrix} I & J_{12} \\ 0 & I \end{bmatrix}. \tag{2.3}$$

We assume here that the nodes from $N_k \backslash N_{k-1}$ are ordered first and then the nodes from N_{k-1}.

We may compute also the two-level stiffness matrices $\overline{A}^{(k)}$ for the hierarchical two-level basis functions and let $\overline{A}^{(k)}$ be partitioned in the same manner as $A^{(k)}$ into the two by two block structure $\{\overline{A}_{i,j}^{(k)}\}, i,j = 1,2$. For $A^{(k)}$ and $\overline{A}^{(k)}$ we have the following exact factorizations

$$\overline{A}^{(k)} = \begin{bmatrix} \overline{A}_{11}^{(k)} & 0 \\ \overline{A}_{21}^{(k)} & \overline{S}^{(k)} \end{bmatrix} \begin{bmatrix} I & \overline{A}_{11}^{(k)-1}\overline{A}_{12}^{(k)} \\ 0 & I \end{bmatrix},$$

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & S^{(k)} \end{bmatrix} \begin{bmatrix} I & A_{11}^{(k)-1}A_{12}^{(k)} \\ 0 & I \end{bmatrix},$$

where $\overline{S}^{(k)}$ and $S^{(k)}$ are the Schur complements

$$\begin{aligned} \overline{S}^{(k)} &= \overline{A}_{22}^{(k)} - \overline{A}_{21}^{(k)}\overline{A}_{11}^{(k)-1}\overline{A}_{12}^{(k)}, \\ S^{(k)} &= A_{22}^{(k)} - A_{21}^{(k)}A_{11}^{(k)-1}A_{12}^{(k)}. \end{aligned} \tag{2.4}$$

Note that these are positive definite. Based on the identity

$$\overline{A}^{(k)} = J^t A^{(k)} J$$

a straightforward computation (see [8]) shows that

$$\overline{A}_{12}^{(k)} = A_{12}^{(k)} + A_{11}^{(k)}J_{12}, \quad \overline{A}_{21}^{(k)} = A_{21}^{(k)} + J_{12}^t A_{11}^{(k)}$$

and

$$\overline{S}^{(k)} = S^{(k)}. \tag{2.5}$$

Note further that $\overline{A}_{22}^{(k)} = A^{(k-1)}$ and that $\overline{A}_{11}^{(k)} = A_{11}^{(k)}$. Consider now the following strengthened C.-B.-S. inequality:

There exists $\gamma = \gamma_k$, $0 < \gamma < 1$ such that

$$a(u,v) \le \gamma\big(a(u,u)\big)^{\frac{1}{2}}\big(a(v,v)\big)^{\frac{1}{2}} \text{ all } u \in V_k \backslash V_{k-1}, \text{ all } v \in V_{k-1}. \tag{2.6}$$

Here γ is the constant in the hierarchical two-level method studied in [1], [4].

With the hierarchical two-level basis functions, (2.6) takes the form

$$\mathbf{v}_1^t \overline{A}_{12}^{(k)} \mathbf{v}_2 \le \gamma(\mathbf{v}_1^t A_{11}^{(k)} \mathbf{v}_1)^{\frac{1}{2}} (\mathbf{v}_2^t A^{(k-1)} \mathbf{v}_2)^{\frac{1}{2}} \text{ all } \mathbf{v}_1 \in \mathbb{R}^{n_k - n_{k-1}}, \mathbf{v}_2 \in \mathbb{R}^{n_{k-1}} \tag{2.7}$$

The constant γ can be computed locally from each local finite element matrix (see [7] and [1], for instance). Therefore, γ is independent of the jumps in the diffusion coefficients $a_{i,j}$ in (2.1). In [4] and [1] it has further been shown that the spectral condition number of $A_{11}^{(k)}$ is bounded above by a constant independent of the level k of the mesh, if we construct the finer mesh triangles so that they are congruent with the original, parent triangle.

The following results will be used for the analysis of the method.

Lemma 2.1. (see [1], [2] and [8]).
Let γ be the constant in the strengthened C.-B.-S. inequality. Then

a) $1-\gamma^2 \leq \mathbf{v}_2^t S^{(k)} \mathbf{v}_2 / \mathbf{v}_2^t A^{(k-1)} \mathbf{v}_2 \leq 1$, all $\mathbf{v}_2 \in \mathbb{R}^{n_{k-1}}$

b) $\widehat{\mathbf{v}}_1^t A_{11}^{(k)} \widehat{\mathbf{v}}_1 \leq \frac{1}{1-\gamma^2} \widehat{\mathbf{v}}^t \overline{A}^{(k)} \widehat{\mathbf{v}}$, all $\widehat{\mathbf{v}} \in \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \widehat{\mathbf{v}}_2 \end{bmatrix} \in \mathbb{R}^{n_k}$.

3. THE AUXILIARY MATRIX SEQUENCE

The recursive construction of the preconditioner to the matrix sequence $\{A^{(k)}\}$ shall be based on a to $\{A^{(k)}\}$ spectrally equivalent sequence. We present this first for the fully hierarchical sequence $\widehat{A}^{(k)}$ or for a rth level hierarchical sequence $\widehat{A}^{(k,r)}$, where $\widehat{A}^{(k,2)} = \overline{A}^{(k)}$, the two-level hierarchical matrix and $\widehat{A}^{(k,k)} = \widehat{A}^{(k)}$, the fully hierarchical basisfunction matrix.

The rth level hierarchical set of basis functions is defined by

$$\widehat{V}^{(k,r)} = \{\varphi_i^{(k)}, x_i \in N_k \backslash N_{k-1}\} \cup \{\varphi_i^{(k-1)}, x_i \in N_{k-1} \backslash N_{k-2}\} \cup \ldots$$
$$\cup \{\varphi_i^{(k-r+1)}, x_i \in N_{k-r+1}\}, r = 2, \ldots, k.$$

Similarly, the vectors can be represented by a direct sum,

$$\widehat{\mathbf{v}}^{(k,r)} = \widehat{\mathbf{v}}_1^{(k)} \oplus \widehat{\mathbf{v}}_1^{(k-1)} \oplus \ldots \oplus \widehat{\mathbf{v}}_1^{(k-r+2)} \oplus \mathbf{v}^{(k-r+1)}, r = 2, \ldots, k. \tag{3.1}$$

Let $\widehat{A}^{(k,r)}$ be the matrix corresponding to the rth level set of basisfunctions. Then, by the definitions of matrices $\overline{A}_{i,j}^{(k)}$ in section 2, we have

$$\widehat{A}^{(k,r)} = \begin{bmatrix} A_{11}^{(k)} & \overline{A}_{12}^{(k)} & & \\ \overline{A}_{21}^{(k)} & A_{11}^{(k-1)} & \overline{A}_{12}^{(k-1)} & \\ & \overline{A}_{21}^{(k-1)} & \ddots & \\ & & & A^{(k-r+1)} \end{bmatrix}, r = 2, \ldots, k \tag{3.2}$$

Note that this matrix is *not* block tridiagonal in general.

Similarly, let $\widehat{B}^{(k,r)}$ be defined as $\widehat{A}^{(k,r)}$, but where $A_{11}^{(k-s)}$ has been approximated with $B_{11}^{(k-s)}, s = 0, ..., r-2$ so that

$$\widehat{\mathbf{v}}_1^{(k-s)^T} A_{11}^{(k-s)} \widehat{\mathbf{v}}_1^{(k-s)} \leq \widehat{\mathbf{v}}_1^{(k-s)^T} B_{11}^{(k-s)} \widehat{\mathbf{v}}_1^{(k-s)} \leq (1+b_s) \widehat{\mathbf{v}}_1^{(k-s)^T} A_{11}^{(k-s)} \widehat{\mathbf{v}}_1^{(k-s)}, \quad (3.3)$$

all $\widehat{\mathbf{v}}_1^{(k-s)}$, and where $b_s \geq 0, s = 0, ..., r-2$. Then

$$\widehat{B}^{(k,r)} = \begin{bmatrix} B_{11}^{(k)} & \overline{A}_{12}^{(k)} & & \\ \overline{A}_{21}^{(k)} & B_{11}^{(k-1)} & \overline{A}_{12}^{(k-1)} & \\ & \overline{A}_{21}^{(k-1)} & \ddots & \\ & & & A^{(k-r+1)} \end{bmatrix}, r = 2, ..., k \quad (3.4)$$

Let $\widehat{B}^{(k)}$ be the corresponding fully hierarchical matrix, and $\overline{B}^{(k)} = \widehat{B}^{(k,2)}$, i.e.

$$\overline{B}^{(k)} = \begin{bmatrix} B_{11}^{(k)} & \overline{A}_{12}^{(k)} \\ \overline{A}_{21}^{(k)} & A^{(k-1)} \end{bmatrix}.$$

Under some further assumptions on the size of b_s, we shall show that $\widehat{A}^{(k)}$ and $\widehat{B}^{(k)}$ are spectrally equivalent,

To this end we need the following lemma.

Lemma 3.1. Let $\widehat{\mathbf{v}}^{(k,r)}$ be represented by (3.1). Then

$$\widehat{\mathbf{v}}^{(k,r)^T} \widehat{A}^{(k,r)} \widehat{\mathbf{v}}^{(k,r)} \geq (1-\gamma^2)^s \widehat{\mathbf{v}}^{(k-s,2)^T} \widehat{A}^{(k-s,2)} \widehat{\mathbf{v}}^{(k-s,2)}, \quad (3.5)$$

all $\widehat{\mathbf{v}}^{(k,r)}, s = 1, ..., r-1$.

Proof. (2.4) shows that

$$\widehat{\mathbf{v}}^{(k,2)^T} \widehat{A}^{(k,2)} \widehat{\mathbf{v}}^{(k,2)} \geq \mathbf{v}^{(k-1)^T} S^{(k)} \mathbf{v}^{(k-1)}.$$

Lemma 2.1a now shows that

$$\begin{aligned} &\widehat{\mathbf{v}}^{(k,2)^T} \widehat{A}^{(k,2)} \widehat{\mathbf{v}}^{(k,2)} \geq (1-\gamma^2) \mathbf{v}^{(k-1)^T} A^{(k-1)} \mathbf{v}^{(k-1)} \\ &= (1-\gamma^2) \widehat{\mathbf{v}}^{(k-1)^T} \overline{A}^{(k-1)} \widehat{\mathbf{v}}^{(k-1)} = (1-\gamma^2) \widehat{\mathbf{v}}^{(k-1,2)^T} \widehat{A}^{(k-1,2)} \widehat{\mathbf{v}}^{(k-1,2)}. \end{aligned}$$

By recursion, we find

$$\widehat{\mathbf{v}}^{(k,2)^T} \widehat{A}^{(k,2)} \widehat{\mathbf{v}}^{(k,2)} \geq (1-\gamma^2)^s \widehat{\mathbf{v}}^{(k-s,2)^T} \widehat{A}^{(k-s,2)} \widehat{\mathbf{v}}^{(k-s,2)}$$

with equality if and only if (3.5) holds.

Since

$$\widehat{\mathbf{v}}^{(k,2)^T}\widehat{A}^{(k,2)}\widehat{\mathbf{v}}^{(k,2)}=\widehat{\mathbf{v}}^{(k,r)^T}\widehat{A}^{(k,r)}\widehat{\mathbf{v}}^{(k,r)}$$

if $\widehat{\mathbf{v}}^{(k,r)}$ is the corresponding transformation of $\widehat{\mathbf{v}}^{(k,2)}$, the proof is complete. ■

Theorem 3.1. Let $\widehat{A}^{(k,r)}$ and $\widehat{B}^{(k,r)}$ be defined by (3.2) and (3.4), respectively and let (3.3) hold with

$$b_s \le q^s b_0 \tag{3.6}$$

for some positive b_0 and $q < 1-\gamma^2$. Then

$$0 \le \widehat{\mathbf{v}}^{(k,r)^T}(\widehat{B}^{(k,r)}-\widehat{A}^{(k,r)})\widehat{\mathbf{v}}^{(k,r)} \le \frac{b_0}{1-\gamma^2-q}\widehat{\mathbf{v}}^{(k,r)^T}\widehat{A}^{(k,r)}\widehat{\mathbf{v}}^{(k,r)},\ \text{all } \widehat{\mathbf{v}}^{(k,r)}. \tag{3.7}$$

Proof. Using (3.3) we find

$$\begin{aligned}&\widehat{\mathbf{v}}^{(k,2)^T}(\widehat{B}^{(k,2)}-\widehat{A}^{(k,2)})\widehat{\mathbf{v}}^{(k,2)}=\widehat{\mathbf{v}}_1^{(k)^T}(B_{11}^{(k)}-A_{11}^{(k)})\widehat{\mathbf{v}}_1^{(k)}\\ &\le b_0\widehat{\mathbf{v}}_1^{(k)^T}A_{11}^{(k)}\widehat{\mathbf{v}}_1^{(k)}.\end{aligned}$$

Lemma 2.1b now shows that

$$\widehat{\mathbf{v}}^{(k,2)^T}(\widehat{B}^{(k,2)}-\widehat{A}^{(k,2)})\widehat{\mathbf{v}}^{(k,2)} \le \frac{b_0}{1-\gamma^2}\widehat{\mathbf{v}}^{(k,2)^T}\widehat{A}^{(k,2)}\widehat{\mathbf{v}}^{(k,2)},\ \text{all } \widehat{\mathbf{v}}^{(k,2)}.$$

Similarly, by (3.2), (3.3) and (3.4),

$$\widehat{\mathbf{v}}^{(k,r)^T}(\widehat{B}^{(k,r)}-\widehat{A}^{(k,r)})\widehat{\mathbf{v}}^{(k,r)}=\sum_{s=0}^{r-2}\widehat{\mathbf{v}}_1^{(k-s)^T}(B_{11}^{(k-s)}-A_{11}^{(k-s)})\widehat{\mathbf{v}}_1^{(k-s)} \tag{3.8}$$

so

$$\begin{aligned}\widehat{\mathbf{v}}^{(k,r)^T}(\widehat{B}^{(k,r)}-\widehat{A}^{(k,r)})\widehat{\mathbf{v}}^{(k,r)} &\le \sum_{s=0}^{r-2} b_s\widehat{\mathbf{v}}_1^{(k-s)^T}A_{11}^{(k-s)}\widehat{\mathbf{v}}_1^{(k-s)}\\ &\le \sum_{s=0}^{r-2} b_s/(1-\gamma^2)\widehat{\mathbf{v}}^{(k-s,2)^T}\widehat{A}^{(k-s,2)}\widehat{\mathbf{v}}^{(k-s,2)}\\ &\le \sum_{s=0}^{r-2} b_s/(1-\gamma^2)^{s+1}\widehat{\mathbf{v}}^{(k,r)^T}\widehat{A}^{(k,r)}\widehat{\mathbf{v}}^{(k,r)},\end{aligned}$$

where we have used Lemma 3.1. (3.6) now shows that

$$\begin{aligned}&\widehat{\mathbf{v}}^{(k,r)^T}(\widehat{B}^{(k,r)}-\widehat{A}^{(k,r)})\widehat{\mathbf{v}}^{(k,r)} \le b_0\sum_{s=0}^{r-2}\left(\frac{q}{1-\gamma^2}\right)^s(1-\gamma^2)^{-1}\widehat{\mathbf{v}}^{(k,r)^T}\widehat{A}^{(k,r)}\widehat{\mathbf{v}}^{(k,r)}\\ &\le \frac{b_0}{1-\gamma^2-q}\widehat{\mathbf{v}}^{(k,r)^T}\widehat{A}^{(k,r)}\widehat{\mathbf{v}}^{(k,r)}.\end{aligned}$$

The left side inequality follows by (3.8) using (3.3). ∎

Remark 3.1. Since there exists a transformation $J^{(k,r)}$ (on upper block triangular form with unit matrix diagonal blocks and with $J^{(k,2)} = J$ in (2.3)) taking $\widehat{\mathbf{v}}^{(k,r)}$ into $\mathbf{v}^{(k)}$, the spectral equivalence is valid also between the standard basis function matrices $A^{(k,r)}$ and $B^{(k,r)}$ where $A^{(k,r)}$ and $B^{(k,r)}$ are defined as $\widehat{A}^{(k,r)}$ and $\widehat{B}^{(k,r)}$, respectively but with $A_{12}^{(k-s)}$ and $A_{21}^{(k-s)}$ replacing $\overline{A}_{12}^{(k-s)}$ and $\overline{A}_{21}^{(k-s)}$, respectively and where the last matrix block is $A_{22}^{(k-r+2)}$ instead of $A^{(k-r+1)}$.

4. THE ALGEBRAIC MULTILEVEL ITERATIVE METHOD AND ITS OPTIMALITY

We shall now construct a sequence of preconditioners $\{M^{(k)}\}$ to $\{A^{(k)}\}$ using the spectrally equivalent sequence $\{B^{(k)}\}$. This will be done as in the authors first report [2]. Let then $B^{(k)} = J^{-T}\overline{B}^{(k)}J^{-1}$ and

$$M^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ \overline{A}_{12}^{(k)} & \widetilde{B}^{(k)} \end{bmatrix} \begin{bmatrix} I & B_{11}^{(k)^{-1}}\overline{A}_{12}^{(k)} \\ 0 & I \end{bmatrix} \tag{4.1}$$

where $\widetilde{B}^{(k)}$ is defined by

$$\widetilde{B}^{(k)^{-1}} = [I - P_\nu(M^{(k-1)^{-1}}B^{(k-1)})]B^{(k-1)^{-1}} \tag{4.2}$$

and P_ν is defined as a scaled and normalized Chebyshev polynomial

$$P_\nu(t) = [T_\nu(\frac{1+\alpha-2t}{1-\alpha}) + 1]/[T_\nu(\frac{1+\alpha}{1-\alpha}) + 1]$$

where

$$T_\nu(x) = \frac{1}{2}[(x + (x^2-1)^{1/2})^\nu + (x - (x^2-1)^{1/2})^\nu].$$

The parameter α is a lower eigenvalue bound of $M^{(k-1)^{-1}}B^{(k-1)}$ and shall be determined later on. We let

$$Q_{\nu-1}(t) = \big(1 - P_\nu(t)\big)/t.$$

Then $Q_{\nu-1}$ is a polynomial of degree $\nu - 1$,

$$Q_{\nu-1}(t) = q_0 + q_1 t + ... + q_{\nu-1}t^{\nu-1}. \tag{4.3}$$

If $\nu = 1$ then $P_\nu(t) = 1 - t$ and

$$\widetilde{B}^{(k)} = M^{(k-1)}.$$

This latter is the method studied previously by the second author [8].
Note that P_ν is nonnegative in the interval $[\alpha, 1]$, $0 < \alpha < 1$ and has the smallest local maximum of all polynomials of degree ν in this interval.
As can be seen from the definition (4.1), (4.2) of the preconditioner it is defined recursively from one level to the previous coarser level. At every level when applying this preconditioner we need to solve two linear systems with $B_{11}^{(k)}$, and a linear system with $\widetilde{B}^{(k-1)}$, or multiplying a vector with $B_{21}^{(k)^{-1}}$ and $\widetilde{B}^{(k-1)^{-1}}$ respectively.

For the latter computation we use (4.3), i.e. for a given vector $\mathbf{v}_2$ we find

$$\mathbf{y}_2 \equiv \widetilde{B}^{(k)-1}\mathbf{v}_2 = Q_{\nu-1}(M^{(k-1)^{-1}}B^{(k-1)})M^{(k-1)^{-1}}\mathbf{v}_2$$

or

$$\mathbf{y}_2 = [q_0 I + q_1 M^{(k-1)^{-1}}B^{(k-1)} + ... + q_{\nu-1}(M^{(k-1)^{-1}}B^{(k-1)})^{\nu-1}]M^{(k-1)^{-1}}\mathbf{v}_2 \quad (4$$

which can be implemented recursively by the well known Horners algorithm (see [2]).

The matrix $B_{11}^{(k)}$ can be constructed in many ways. Here we use the method presented in the previous publication of the authors [3], namely we let

$$B_{11}^{(k-s)^{-1}} = [I - (\tilde{M}_{11}^{(k-s)})^{2\beta_s}]A_{11}^{(k-s)^{-1}},$$

where $\beta_s = m(s+1)$, $m \geq 1$ an integer.
Further

$$\tilde{M}_{11}^{(k-s)} = I - C_{11}^{(k-s)^{-1}}A_{11}^{(k-s)}$$

and, when $A_{11}^{(k-s)}$ is an M-matrix, we let

$$C_{11}^{(k-s)} = \text{ diag }(A_{11}^{(k-s)}).$$

Note that, similar to the case for $\tilde{B}^{(k-s)}$, the factor $I - \tilde{M}_{11}^{(k-s)^{2\beta_s}}$ contains the factor $I - \tilde{M}_{11}^{(k-s)} = C_{11}^{(k-s)^{-1}}A_{11}^{(k-s)}$, so the inverse of $A_{11}^{(k-s)^{-1}}$ is cancelled out.
Since $\rho(\tilde{M}_{11}^{(k-s)}) < 1$, we have

$$\mathbf{v}_1^t B_{11}^{(k-s)}\mathbf{v}_1 \geq \mathbf{v}_1^t A_{11}^{(k-s)}\mathbf{v}_1, \text{ all } \mathbf{v}_1. \qquad (4.5)$$

Further, since $A_{11}^{(k-s)}$ is diagonally dominant,

$$\rho(\tilde{M}_{11}^{(k-s)}) \leq \widehat{q} < 1,$$

for some $\widehat{q}$ which does not depend on k (see [1]), we find

$$\mathbf{v}_1^t B_{11}^{(k-s)} \mathbf{v}_1 \leq \frac{1}{1-\widehat{q}^{2\beta_s}} \mathbf{v}_1^t A_{11}^{(k-s)} \mathbf{v}_1 \tag{4.6}$$

or, by (4.5) and (4.6),

$$\mathbf{v}_1^t A_{11}^{(k-s)} \mathbf{v}_1 \leq \mathbf{v}_1^t B_{11}^{(k-s)} \mathbf{v}_1 \leq (1+\frac{\widehat{q}^{2\beta_s}}{1-\widehat{q}^{2\beta_s}}) \mathbf{v}_1^t A_{11}^{(k-s)} \mathbf{v}_1,$$

so (3.3) is valid with

$$b_s = \widehat{q}^{2\beta_s}/(1-\widehat{q}^{2\beta_s}).$$

Hence, if

$$\widehat{q}^{2m} < 1-\gamma^2 \tag{4.7}$$

then (3.6) is satisfied with $b_0 = \frac{q}{1-q}, q = \widehat{q}^{2m}$.

Note that there exists always an m, independent of the level k, such that (4.7) is satisfied.

The computation of $B_{11}^{(k-s)^{-1}} \mathbf{v}$ for some vector $\mathbf{v}$ can be implemented in the same way as the computation of $\mathbf{y}_2$ above (for details, see [1]). The algorithm AMLI to implement one preconditioning step can also be found in [3], together with the formulas for the polynomial coefficients q_i in (4.3), for various polynomial degrees ν. Likewise the relative spectral condition numbers follow readily from the previous results.

Based on the results in [2] and the fact that $B^{(k)}$ and $A^{(k)}$ are spectrally equivalent the following main result follows.

Theorem 4.1. The algebraic multilevel preconditioners $M^{(k)}$ based on the auxiliary matrix sequence $\{B^{(k)}\}$ and matrix blocks $B_{11}^{(k)}$ satisfying (3.3), (3.6), are spectrally equivalent to the hierarchical basis stiffness matrices $\widehat{A}^{(k)}$ if $\nu^{-1} < \sqrt{1-\gamma^2}$ and then the following inequalities hold,

$$\widehat{\mathbf{v}}^T \widehat{A}^{(k)} \widehat{\mathbf{v}} \leq \widehat{\mathbf{v}}^T M^{(k)} \widehat{\mathbf{v}} \leq [1+\frac{b_0}{1-\gamma^2-q}] \lambda_k \widehat{\mathbf{v}}^T \widehat{A}^{(k)} \widehat{\mathbf{v}}, \text{ all } \widehat{\mathbf{v}} \in \mathbf{R}^{n_k},$$

where λ_k is the largest eigenvalue of $B^{(k)^{-1}} M^{(k)}$ and satisfies

$$\lambda_k \leq \frac{1}{\alpha},$$

and $\alpha \in (0,1)$, the parameter in the definition of the polynomials $P_\nu(t)$, (4.2), is the smallest positive root of the equation

$$\sqrt{1-\gamma^2} = \frac{(1+\sqrt{t})^\nu + (1-\sqrt{t})^\nu}{2\sum_{s=1}^{\nu}(1+\sqrt{t})^{\nu-s}(1-\sqrt{t})^{s-1}} \tag{4.8}$$

Proof. We need only to show that the matrices $\overline{B}^{(k)}$ satisfy the strengthened C.-B.-S. inequality

$$\widehat{\mathbf{v}}_1^T \overline{A}_{12}^{(k)} \mathbf{v}_2 \le \gamma(\widehat{\mathbf{v}}_1^T B_{11}^{(k)} \widehat{\mathbf{v}}_1)^{\frac{1}{2}} (\mathbf{v}_2^T B^{(k-1)} \mathbf{v}_2)^{\frac{1}{2}},$$
$$\text{all } \widehat{\mathbf{v}}_1 \in \mathbf{R}^{n_k - n_{k-1}}, \mathbf{v}_2 \in \mathbf{R}^{n_{k-1}}.$$

This is readily seen as $\widehat{A}^{(k)}$ satisfies such an inequality, namely,

$$\begin{aligned}\widehat{\mathbf{v}}_1^T \overline{A}_{12}^{(k)} \mathbf{v}_2 &\le \gamma(\widehat{\mathbf{v}}_1^T A_{11}^{(k)} \widehat{\mathbf{v}}_1)^{\frac{1}{2}} (\mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2)^{\frac{1}{2}} \\ &\le \gamma(\widehat{\mathbf{v}}_1^T B_{11}^{(k)} \widehat{\mathbf{v}}_1)^{\frac{1}{2}} (\mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2)^{\frac{1}{2}}, \\ &\qquad \text{by (3.3), } s = 0, \\ &\le \gamma(\widehat{\mathbf{v}}_1^T B_{11}^{(k)} \widehat{\mathbf{v}}_1)^{\frac{1}{2}} (\mathbf{v}_2^T B^{(k-1)} \mathbf{v}_2)^{\frac{1}{2}}, \\ &\qquad \text{by Theorem 3.1.}\end{aligned}$$

The remainder of the proof follows the proof of theorem 3.3 in [2]. ■

5. A PRACTICAL TEST

The test problem was as in [3], that is, $-\Delta u = 0$ in Ω, the model L-shaped region shown on Fig. 6.1 with boundary conditions

$$\begin{aligned}\frac{\partial u}{\partial n} = 0 \text{ on } \Gamma_N = \{(x_1, 0), 0 \le x_1 < 1\} \\ \cup \{(0, x_2), 0 < x_2 < 1\}\end{aligned}$$

and $u = 1$ on $\partial\Omega \backslash \Gamma_N$. That is $u = 1$ in Ω is the exact solution.

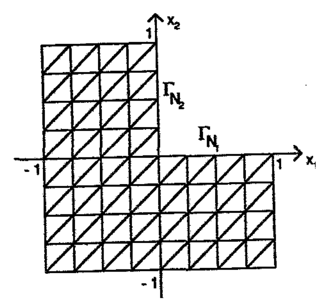

Figure 5.1. Domain Ω

The AMLI-preconditioner M is defined by

$$M^{-1} = \left[I - P_\nu(M^{(\ell)-1}A^{(\ell)})\right] A^{(\ell)-1}$$

and $M^{(\ell)}$ is the matrix defined in (4.1). The blocks $B_{11}^{(k)}$, the approximations to $A_{11}^{(k)}$, are obtained as follows:

$$B_{11}^{(k)-1} = \left[I - \widehat{M}_{11}^{(k)^{2(\ell-k+1)}}\right] A_{11}^{(k)-1},$$

where

$$\widehat{M}_{11}^{(k)} = I - D_{11}^{(k)-1} A_{11}^{(k)}, D_{11}^{(k)} = diag A_{11}^{(k)}.$$

Note that $\rho(\widehat{M}_{11}^{(k)}) \le \sqrt{q_0} < 1$ independent of k. Actually, based on the analysis in Maitre, Musy [7] one can show that $\rho(\widehat{M}_{11}^{(k)}) \le \frac{1}{\sqrt{2}}$, hence $q_0 \le \frac{1}{2}$. Then the required inequality $q_0 < 1-\gamma^2$ is valid since $1-\gamma^2 > \frac{1}{2} \ge q_0$ (however γ^2 approaches $\frac{1}{2}$ with $\ell \to \infty$).

The solution method was the preconditioned conjugate gradient method with preconditioning matrix M further refered to as AMLI-CG method.

We solved the problem

$$A\mathbf{x} = \mathbf{b},\ A = A^{(\ell)}$$

and the initial approximation was chosen as

$$\mathbf{x}_0 = M^{-1}\mathbf{b}.$$

The corresponding initial residual was

$$\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0,$$

and we let $\mathbf{r}$ be the current one.
Setting

$$\Delta_0 = \sqrt{\mathbf{r}_0^T\mathbf{r}_0}, \Delta = \sqrt{\mathbf{r}^T\mathbf{r}},$$

the stopping criterion was chosen

$$\Delta < \varepsilon = 10^{-9}.$$

On the tables 1,2 we report the number of iterations, *iter*, required to satisfy the stopping criterion, with the average reduction factor ρ equal to

$$\left(\frac{\Delta}{\Delta_0}\right)^{\frac{1}{iter}},$$

the cpu times in seconds, and the number of unknowns $N = n_\ell$ for various number of levels $\ell = 3, 4, 5, 6, 7$.

The test clearly indicates that the AMLI-CG method is of optimal order, has a fixed number of iterations and bounded corresponding average reduction factors. Due to the fact that we have chosen very rough approximations $B_{11}^{(k)}$ to $A_{11}^{(k)}$, that is the constant q_0, $(\rho^2(\widehat{M}_{11}^{(k)}) \leq q_0)$, is not small enough, the performance of the method for $\nu = 2$ and $\nu = 3$ is about the same (say, in terms of the number of iterations).

This is clearly seen as for

$$\nu = 2, \lambda = \lambda^{(2)} \leq \left[1 + \frac{C}{1 - \gamma^2 - q_0}\right](1 + \varepsilon_2), \varepsilon_2 < 1$$

and

$$\nu = 3, \lambda = \lambda^{(3)} \left[\leq 1 + \frac{C}{1 - \gamma^2 - q_0}\right](1 + \varepsilon_3), \varepsilon_3 < 1$$

($\lambda^{(\nu)}$ is the extreme eigenvalue of $M^{(\ell)-1}A^{(\ell)}$, see theorem 4.1). Hence we see that the term $\frac{C}{1-\gamma^2-q_0}$ dominates. Also since the constant $\frac{1}{1-\gamma^2-q_0}$ is not uniformly bounded in ℓ we see, by our numerical test, that the restriction $q_0 < 1 - \gamma^2$ is not actually necessary.

Thus since the amount of work per iteration step is the smallest in the case where $\nu = 2$ with about the same rate of convergence, this case is preferable.

The test was run on a Bulgarian computer EC 1037 (about 1.2 Mflops peak performance).

$\nu = 2$

l	iter	ρ	cpu time /sec/	N
3	14	0.201	31.31	176
4	15	0.209	121.72	736
5	15	0.209	431.93	3008
6	15	0.209	1624.61	12160
7	15	0.209	6001.36	48896

$\nu = 3$

l	iter	ρ	cpu time /sec/	N
3	12	0.165	69.40	176
4	13	0.169	339.67	736
5	13	0.168	1433.19	3008
6	13	0.168	5987.46	12160
7	13	0.169	24638.55	48896

Tables 1.-2. Iterative convergence results for solving the test problem with the AMLI-CG method, ν=2, 3.

REFERENCES

1. O. Axelsson, I. Gustafsson, Preconditioning and two-level multigrid methods of arbitrary degree of approximation, Math. Comp. 40 (1983) 219-242.
2. O. Axelsson, P.S. Vassilevski, Algebraic multilevel preconditioning methods, I, Numer. Math. 56 (1989) 157-177.
3. O. Axelsson, P.S. Vassilevski, Algebraic multilevel preconditioning methods, II, SIAM J. Numer. Anal., to appear.
4. R. Bank, T. Dupont, Analysis of a two-level scheme for solving finite element equations, Report CNA-159, Center for Numerical Analysis, The University of Texas at Austin, 1980.
5. D. Braess, The contraction number of a multigrid method for solving the Poisson equation, Numer. Math. 37 (1981) 387-404.
6. D. Braess, The convergence rate of a multigrid method with Gauss-Seidel relaxation for the Poisson equation, in: Multigrid Methods, Proceedings, Köln-Porz, 1981 (W. Hackbusch, U. Trottenberg, eds.) LNM 960, Springer 1982, 368-386.
7. J.-F. Maitre, F. Musy, The contraction of a class of two-level methods; and exact evaluation for some finite element subspaces and model problems, in: Multigrid Methods, Proceedings, Köln-Porz, 1981 (W. Hackbusch, U. Trottenberg, eds.) LNM 960, Springer 1982, 535-544.
8. P.S. Vassilevski, Nearly optimal iterative methods for solving finite element elliptic equations based on the multilevel splitting of the matrix, Report # 1989-09, EORI, University of Wyoming, Laramie, USA, 1989.
9. H. Yserentant, On the multilevel splitting of finite element spaces, Numer. Math. 49 (1986) 379-412.