

# THE HIERARCHICAL BASIS MULTIGRID METHOD FOR CONVECTION-DIFFUSION EQUATIONS

RANDOLPH E. BANK\* AND MOHAMED BENBOURENANE †

**Abstract.** We make a theoretical study of the application of a standard hierarchical basis multigrid iteration to the convection diffusion equation, discretized using an upwind finite element discretizations. We show behavior that in some respects is similar to the symmetric positive definite case, but in other respects is markedly different. In particular, we find the rate of convergence depends significantly on parameters which measure the strength of the upwinding, and the size of the convection term. Numerical calculations illustrating some of these effects are given.

**Key words.** Finite Element Methods, Upwinding, Convection Diffusion Equations.

**AMS subject classifications.** 65N05, 65N10, 65N20

**1. Introduction.** In this paper, we consider the numerical solution of the model convection diffusion equation

$$(1) \quad \begin{aligned} -\Delta u + \beta \cdot \nabla u &= f && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega \end{aligned}$$

where  $\Omega$  is a (polygonal) region in  $\mathcal{R}^2$ , and, for simplicity,  $\beta$  is a constant velocity vector. We are mainly interested in the case  $|\beta| \gg 1$ .

A weak formulation of (1) is: find  $u \in \mathcal{H}_0^1(\Omega)$  such that

$$(2) \quad a(u, v) = (f, v)$$

for all  $v \in \mathcal{H}_0^1(\Omega)$ . Here  $\mathcal{H}_0^1(\Omega)$  is the usual subspace of the Sobolev space  $\mathcal{H}^1(\Omega)$  whose elements satisfy the homogeneous Dirichlet boundary conditions, and

$$\begin{aligned} a(u, v) &= \int_{\Omega} \nabla u \cdot (\nabla v + \beta v) dx \\ (f, v) &= \int_{\Omega} f v dx \end{aligned}$$

To discretize (2), we first construct a triangulation  $\mathcal{T}$  of  $\Omega$ , consisting of shape-regular triangles characterized by a small parameter  $h$ . As is the usual case with hierarchical basis multigrid methods, quasiuniformity of the mesh is not essential to the proofs, so these results will typically be developed in terms of a local mesh size  $h_t$ , denoting the size of a triangle  $t \in \mathcal{T}$ . The fine mesh  $\mathcal{T}$  can be constructed in the usual hierarchical fashion, beginning with a course mesh  $\mathcal{T}_1$  and then inductively creating refined meshes  $\mathcal{T}_j$ ,  $2 \leq j \leq k$ , (with  $\mathcal{T} \equiv \mathcal{T}_k$ ), by taking each triangle in  $\mathcal{T}_{j-1}$  and creating 4 triangles in  $\mathcal{T}_j$  by pairwise connecting the midpoints. Nonuniform refinements of the type described in [5], or implemented in [2], could be allowed, since these types of local refinement cause only trivial changes in our convergence analysis compared with case of uniform refinement.

---

\* Department of Mathematics, University of California at San Diego, La Jolla, CA 92093, USA. The work of this author was supported by the Office of Naval Research under contract N00014-89J-1440.

† Department of Mathematics, University of California at San Diego, La Jolla, CA 92093, USA. The work of this author was supported by the Office of Naval Research under contract N00014-89J-1440.

Let  $\mathcal{M}_j \subset \mathcal{H}_0^1(\Omega)$  be a space of continuous piecewise linear polynomials associated with the triangulation  $\mathcal{T}_j$ ,  $1 \leq j \leq k$ , with  $\mathcal{M} \equiv \mathcal{M}_k$ . We will make use of the hierarchical decomposition

$$\mathcal{M}_j = V_1 \oplus V_2 \oplus \dots \oplus V_j$$

$1 \leq j \leq k$ , where  $V_1 \equiv \mathcal{M}_1$ , and

$$V_j = \{\phi \in \mathcal{M}_j | \phi(x) = 0 \text{ for } x \text{ a vertex of } \mathcal{T}_{j-1}\}$$

for  $j > 1$ . This is the usual hierarchical decomposition used in the hierarchical basis multigrid method [5].

The hierarchical basis multigrid method extends in straightforward fashion to higher degree polynomial spaces, using, for example, the development given in [4] for the symmetric, positive definite case. Essentially, if  $\mathcal{P}_j$  is the space of piecewise polynomial of degree  $j > 1$ , we can make the decomposition  $\mathcal{P}_j = \mathcal{M} \oplus \mathcal{W}$ ; the subspace  $\mathcal{W}$  contains the higher degree polynomials and is often characterized in terms of nodal basis functions. This space  $\mathcal{W}$  simply becomes a special “finest” level  $k + 1$  in a hierarchical basis multigrid method. The remaining  $k$  levels are composed of piecewise linear subspaces in the usual fashion.

The standard Galerkin discretization of (2) is: find  $u_h \in \mathcal{M}$  such that

$$(3) \quad a(u_h, v) = (f, v)$$

for all  $v \in \mathcal{M}$ . The standard Galerkin discretization is known to give rise to unstable oscillations in the solution until the product  $|\beta|h$  becomes sufficiently small. When the convection term is large, this restriction often precludes the use of the method for computationally realistic meshes.

A standard approach to remedy this situation is the use of *upwind* finite element discretizations such as artificial diffusion and streamline diffusion [10]. Such methods modify the bilinear form  $a(\cdot, \cdot)$  through the addition of (mesh dependent) stabilization terms. For the artificial diffusion method, we have

$$(4) \quad a_d(u, v) = \int_{\Omega} \nabla u \cdot \{(1 + \delta_t)\nabla v + \beta v\} dx$$

where

$$(5) \quad \delta_t = c_0 h_t |\beta| (1 + h_t |\beta|)$$

The inclusion of the  $h_t^2 |\beta|^2$  term is not necessary for the stability of the method, but it turns out to be useful in our convergence analysis of the hierarchical basis multigrid method. In this method, a mesh dependent local multiple of  $-\Delta u$  is added to the convection diffusion equation (1). Because this represents a modification of the original partial differential equation, this scheme is at most first order, regardless of the degree of polynomial approximation.

The streamline diffusion method is a Petrov-Galerkin method in which the test space is taken to be of the form

$$\phi + \frac{\delta_t}{|\beta|} \phi_n$$

where  $n = \beta/|\beta|$  is a unit vector,  $\phi_n = n \cdot \nabla \phi$ . and  $\delta_t$  is given by (5). The use of a different test space changes both the left and right hand sides of (2), so the streamline

diffusion methods allows for the possibility of higher order convergence [10]. In this study, we are mainly concerned with solving the linear system, and thus view the streamline diffusion method as a modification of the bilinear form for the standard Galerkin method. This leads to the same matrix but different right hand side in the linear system (and also to a first order discretization). However, since the analysis of the convergence rate does not depend on the right hand side but only on the matrix, all of our convergence results apply to the Petrov-Galerkin formulation as well. For the streamline diffusion method, the bilinear form is:

$$(6) \quad a_s(u, v) = \int_{\Omega} \nabla u \cdot (\nabla v + \beta v) + \delta_t u_n v_n dx$$

This is similar to the artificial diffusion, except that the upwinding is restricted to just the streamline direction.

Let  $\{\phi_i\}_{i=1}^N$  be the usual hierarchical basis for  $\mathcal{M}$ . Then the upwind discretizations lead to linear systems of equations of the form

$$(7) \quad Au = F$$

where

$$A_{ij} = a(\phi_j, \phi_i)$$

and  $a(\cdot, \cdot)$  corresponds to either (4) or (6).

We remark that *computationally* it is undesirable to assemble and solve the set of equations represented in the hierarchical basis, because the matrix is much less sparse (although better conditioned) than the corresponding formulation using the standard nodal basis. In practice, hierarchical basis methods are implemented using the standard nodal basis, in combination with some recursive algorithms that are very similar to the standard multigrid V-cycle [9] [5]. Because these computational aspects are the same as for the symmetric, positive definite case, we will not deal with them here, but merely refer to the appropriate literature [5] [2]. Here we are concerned mainly with obtaining estimates for the rate of convergence for hierarchical basis methods, and the methods we are interested in are *mathematically* equivalent to standard block iterative methods for (7). Thus, for the development of our estimates, we consider (7) as our model problem.

If we order the standard hierarchical basis functions by level, then the matrix  $A$  has the block structure

$$(8) \quad A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & & A_{2k} \\ \vdots & & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix}$$

where  $A_{jj}$  corresponds to the subspace  $V_j$ .

For the symmetric, positive definite case ( $\beta = 0$ ), the matrix  $A$  has a condition number of order  $O(k^2)$ . In this case, many standard block iterative methods (Richardson, Jacobi, or symmetric Gauss-Seidel) have generalized condition numbers of the same order. Although all these methods share similar theoretical convergence properties, practically we have implemented the symmetric Gauss-Seidel version because it converges more rapidly due to a better constant in the  $O(k^2)$  term.

In this work, we extend these theoretical results to the nonsymmetric case. Our goal is to gain insight into the behavior of the method as a function of the size of  $|\beta|h$  of the nonsymmetric term, and the strength of the upwinding, which is governed by the parameter  $c_0$ . This is a bit unusual in comparison with most theoretical studies in this area, where the intent is usually to demonstrate the mesh independence of various quantities.

For nonsymmetric problems, the usual approach is to treat the lower order terms as perturbations of a symmetric, positive definite operator, and thus obtain estimates similar to the symmetric, positive definite case, often with an additional restriction that  $h$  is sufficiently small, or the coarse mesh is fine enough [1] [12] [14] [15] [8]. This approach is not useful for our situation, since it cannot be used to explain the behavior of the method on meshes of practical size for the class of problems we consider here. For these problems, it is in fact the upwind terms which stabilize discretization, and also allows for the rapid convergence of the hierarchical basis method. Thus our goal is not to eliminate the mesh dependence from our estimates, but rather to try to gain insight into how the mesh dependent upwinding parameters affect the convergence. Several empirical investigations of the effect have been carried out for regular multigrid methods [6] [13].

The remainder of the paper is organized as follows: In sections 2-3, we analyze the two level case. Here we are able to obtain very precise results, quite comparable to those obtained for one space dimensional model problems using Fourier analysis [3] [8]. Since the results are essentially the same for the cases of artificial diffusion and streamline diffusion, we analyze only the latter case. In section 4, we analyze the linear algebraic aspects of the block symmetric Gauss-Seidel iteration for matrices of the form (8) for the case of  $k$  levels. Then in sections 5 and 6, we make estimates for the artificial diffusion and streamline diffusion discretizations. Here our results for artificial diffusion are stronger than for streamline diffusion, although we think this is at least partly due to our use of less sharp estimates in the latter case. In section 7 we present some numerical illustrations, while in section 8 we make some concluding remarks. The reader not interested in technical details of the analysis can skip sections 2, 5, and 6 without a significant loss in continuity.

**2. Some Preliminary Results.** In this section we prove some technical results which are necessary for the analysis in subsequent sections. We begin with a simple lemma from linear algebra.

LEMMA 2.1. *Let  $v \in \mathcal{R}^n$ ,  $w \in \mathcal{R}^m$  and  $C = vw^t \in \mathcal{R}^{n \times m}$ . Let  $A \in \mathcal{R}^{n \times n}$  and  $B \in \mathcal{R}^{m \times m}$  be symmetric, positive semi-definite with  $v \in \text{Range}(A)$ , and  $w \in \text{Range}(B)$ . Then there exists a positive constant  $\gamma$  such that for every  $x \in \mathcal{R}^n$  and every  $y \in \mathcal{R}^m$*

$$(9) \quad |x^t C y| \leq \gamma (x^t A x)^{\frac{1}{2}} (y^t B y)^{\frac{1}{2}}$$

where

$$(10) \quad \gamma = \sqrt{v^t A^+ v} \sqrt{w^t B^+ w}$$

and  $A^+$  is the (generalized) inverse of  $A$  restricted to  $\text{Range}(A)$ .

*Proof.* Let  $x \in \mathcal{R}^n$ ,  $y \in \mathcal{R}^m$ .

If  $x \in \text{Kernel}(A)$  or  $y \in \text{Kernel}(B)$ , then  $x^t C y = 0$  and (9) is trivially satisfied. Thus we may assume without loss of generality that  $x \in \text{Range}(A)$  and  $y \in \text{Range}(B)$ . Hence

$$\begin{aligned}
\gamma &= \sup_{\substack{x^t Ax=1 \\ y^t By=1}} x^t Cy \\
&= \sup_{x^t Ax=1} x^t v \sup_{y^t By=1} y^t w
\end{aligned}$$

Using Lagrange multipliers, it is easy to see that

$$\begin{aligned}
\sup_{x^t Ax=1} x^t v &= \sqrt{v^t A^+ v} \\
\sup_{y^t By=1} y^t w &= \sqrt{w^t B^+ w}
\end{aligned}$$

and the lemma follows.  $\square$

LEMMA 2.2. *Let  $V$  and  $W$  be two subspaces of a given vector space  $\mathcal{S}$ , and let  $(\cdot, \cdot)_1$  and  $(\cdot, \cdot)_2$  be two inner products defined on  $\mathcal{S}$ , with induced norms  $\|\cdot\|_i$ ,  $1 \leq i \leq 2$ . Assume for all  $v \in V$  and for all  $w \in W$  that the strengthened Cauchy inequalities*

$$|(v, w)_i| \leq \gamma_i \|v\|_i \|w\|_i$$

hold with  $\gamma_i < 1$ ,  $1 \leq i \leq 2$ . If we define a third inner product by

$$\begin{aligned}
(v, w) &= \rho (v, w)_1 + \eta (v, w)_2 \\
\rho &> 0 \\
\eta &> 0
\end{aligned}$$

with induced norm  $\|\cdot\|$ , then

$$|(v, w)| \leq \gamma \|v\| \|w\|$$

where

$$\gamma = \max(\gamma_1, \gamma_2)$$

*Proof.* The proof is similar to that given in [4], and is a relatively simple algebraic manipulation.

$$\begin{aligned}
|(v, w)| &= |\rho (v, w)_1 + \eta (v, w)_2| \\
&\leq \gamma \{\rho \|v\|_1 \|w\|_1 + \eta \|v\|_2 \|w\|_2\} \\
&\leq \gamma \{\rho \|v\|_1^2 + \eta \|v\|_2^2\}^{\frac{1}{2}} \{\rho \|w\|_1^2 + \eta \|w\|_2^2\}^{\frac{1}{2}} \\
&= \gamma \|v\| \|w\|
\end{aligned}$$

$\square$

We now must be a bit more specific in our choices of  $V$  and  $W$  and  $\mathcal{S}$ .

LEMMA 2.3. *Let  $\mathcal{S}$  be the space of  $\mathcal{C}^0$  piecewise linear polynomials associated with the triangulation  $\mathcal{T}$ . Let  $\mathcal{S} = V \oplus W$  be the decomposition of  $\mathcal{S}$  in terms of the hierarchical basis. Let*

$$b(v, w) = \sum_{t \in \mathcal{T}} b(v, w)_t$$

be an inner product defined in  $\mathcal{S}$ , with induced norm

$$\begin{aligned}\|u\|^2 &= \sum_{t \in \mathcal{T}} b(u, u)_t \\ &= \sum_{t \in \mathcal{T}} \|u\|_t^2\end{aligned}$$

Suppose for each  $t \in \mathcal{T}$  there exists  $0 \leq \gamma_t < 1$  such that for all  $v \in V$  and for all  $w \in W$

$$|b(v, w)_t| \leq \gamma_t \|v\|_t \|w\|_t$$

Then

$$|b(v, w)| \leq \gamma \|v\| \|w\|$$

with

$$\gamma = \max_{t \in \mathcal{T}} \gamma_t$$

*Proof.* The proof here is again based on that given in [4].

$$\begin{aligned}|b(u, v)| &\leq \sum_{t \in \mathcal{T}} |b(v, w)_t| \\ &\leq \gamma \sum_{t \in \mathcal{T}} \|v\|_t \|w\|_t \\ &\leq \gamma \left\{ \sum_{t \in \mathcal{T}} \|v\|_t^2 \right\}^{\frac{1}{2}} \left\{ \sum_{t \in \mathcal{T}} \|w\|_t^2 \right\}^{\frac{1}{2}} \\ &= \gamma \|v\| \|w\|\end{aligned}$$

□

We now consider a coarse grid triangle  $t$ . We denote the vertices of  $t$  by  $p_i$ ,  $1 \leq i \leq 3$ , and the midpoints of  $t$  by  $m_i$ ,  $1 \leq i \leq 3$ . Triangle  $t$  has four son elements, denoted by  $s_i$ ,  $0 \leq i \leq 3$ , which are triangles in the fine mesh  $\mathcal{T}$ . This is illustrated in figure (1).

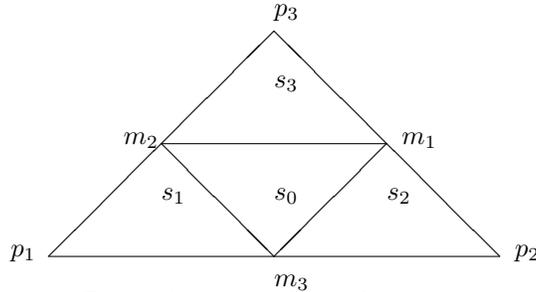


FIG. 1. A coarse grid triangle  $t$

The restriction of the subspaces  $V$  to  $t$  is the space of linear functions on  $t$ , and the hierarchical basis functions for  $V$  are just the nodal basis associated with the

vertices. The restriction of  $W$  to  $t$  is the three dimensional space spanned by fine grid nodal basis functions associated with the midpoints of  $t$ .

LEMMA 2.4. Let  $b_n(\cdot, \cdot)$  be a bilinear form defined on  $\mathcal{S}_t$ , given by

$$\begin{aligned} b_n(v, w) &= \int_t \mathbf{n} \cdot \nabla v \mathbf{n} \cdot \nabla w dx \\ &= \int_t v_n w_n dx \end{aligned}$$

where

$$\mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$$

is a unit vector. Let  $\|\cdot\|$  denote the corresponding seminorm. Then for all  $v \in V_t$  and for all  $w \in W_t$ , we have the strengthened Cauchy inequality

$$(11) \quad |b_n(v, w)| \leq \gamma_t \|v\| \|w\|$$

where  $\gamma \leq \sqrt{3}/2$ .

*Proof.* Without loss of generality, we can consider  $t$  to be the reference triangle with vertices  $p_1 = (0, 0)$ ,  $p_2 = (1, 0)$ , and  $p_3 = (0, 1)$ . To see this, we first show that inequality (11) cannot depend of the diameter  $h_t$  of  $t$ . Consider the rescaling  $t \rightarrow \hat{t}$  defined by

$$\hat{x} = \frac{x - p_1}{h_t}$$

This maps  $t$  to a triangle with similar angles and orientation, but with one vertex at the origin and diameter one. It is easy to see that

$$\int_t v_n w_n dx = \int_{\hat{t}} \hat{v}_n \hat{w}_n d\hat{x}$$

showing that  $\gamma$  will be invariant with respect to scalings and translations. We now map  $\hat{t}$  to the reference element  $\tilde{t}$  by the linear mapping

$$J\tilde{x} = \hat{x}$$

where  $J$  is a  $2 \times 2$  matrix.  $J$  is nonsingular and has bounded inverse due to our assumption of shape regularity for the elements in  $\mathcal{T}$ . Let

$$\begin{aligned} \tilde{\mathbf{n}} &= J^{-1} \mathbf{n} \|J^{-1} \mathbf{n}\|_{\ell_2}^{-1} \\ \alpha &= |\text{Det } J| \|J^{-1} \mathbf{n}\|_{\ell_2}^2 \end{aligned}$$

Then

$$\int_{\hat{t}} \hat{v}_n \hat{w}_n d\hat{x} = \alpha \int_{\tilde{t}} \tilde{v}_{\tilde{\mathbf{n}}} \tilde{w}_{\tilde{\mathbf{n}}} d\tilde{x}$$

showing that on the reference element the inequality will have the same form but for the new unit direction  $\tilde{\mathbf{n}}$ . However, since we are seeking a bound independent of direction, we may disregard this effect. Thus, to simplify notation in the remainder

of the proof, we will assume that  $t$  is the reference element. The element matrix corresponding to the hierarchical basis is block  $2 \times 2$  with  $3 \times 3$  blocks of the form

$$\begin{aligned} M &= \frac{1}{2} \begin{bmatrix} A & -C \\ -C^t & B \end{bmatrix} \\ A &= vv^t \\ C &= vv^t \\ B &= 2vv^t + D \end{aligned}$$

with  $v$  the 3 vector given by

$$v = \begin{bmatrix} -(n_1 + n_2) \\ n_1 \\ n_2 \end{bmatrix}$$

and  $D$  is the diagonal matrix

$$D = \begin{bmatrix} n_1^2 + n_2^2 - (n_1 + n_2)^2 & & \\ & (n_1 + n_2)^2 + n_2^2 - n_1^2 & \\ & & (n_1 + n_2)^2 + n_1^2 - n_2^2 \end{bmatrix}$$

In linear algebraic terms, the problem of computing  $\gamma$  is exactly the estimate (9) in Lemma 2.1. We note that since  $A$  has rank one,  $v^t A^+ v = 1$ , so that  $\gamma^2 = v^t B^+ v$ . While  $B$  is nonsingular for most directions  $\mathbf{n}$ , it will have rank two for  $\mathbf{n} = (1, 0)^t$ ,  $\mathbf{n} = (0, 1)^t$ , and  $\mathbf{n} = (1, -1)^t/\sqrt{2}$ . However, even in these cases  $v \in \text{Range}(B)$ , so the hypothesis of the lemma is satisfied.

In the case where  $B$  is nonsingular, one may compute

$$\begin{aligned} B^{-1}v &= \frac{1}{2v^t D^{-1}v + 1} D^{-1}v \\ &= \frac{-1}{4n_1 n_2 (n_1 + n_2)} \begin{bmatrix} (n_1 + n_2)^2 \\ n_1^2 \\ n_2^2 \end{bmatrix} \end{aligned}$$

and

$$v^t B^{-1}v = \frac{3}{4}$$

In the case  $B$  is singular

$$B^+v = \frac{1}{2v^t v} v$$

and

$$v^t B^+v = \frac{1}{2}$$

Thus we have  $\gamma = \sqrt{3}/2$ .  $\square$

Interestingly, the estimate  $\gamma = \sqrt{3}/2$  was also obtained by Maitre and Musy [11] as the worst case when  $\gamma$  was studied as a function of the element geometry for the Dirichlet integral. This occurred for the degenerate triangle. The fact that

our analysis uses only one directional derivative, while the Dirichlet integral uses two directions which become linearly dependent as the matrix  $J$  becomes singular, explains the similarity.

We can now prove our first main result.

**THEOREM 2.5.** *Let  $\mathbf{n} = \beta/|\beta|$ , and let  $\delta_t$  be given by (5). Consider the bilinear form*

$$b(v, w) = \sum_{t \in \mathcal{T}} \int_t \nabla v \cdot \nabla w + \delta_t v_n w_n dx$$

with corresponding norm  $\|v\|^2 = b(v, v)$ . Then for all  $v \in V$  and for all  $w \in W$ , we have the strengthened Cauchy inequality

$$|b(v, w)| \leq \gamma \|v\| \|w\|$$

where  $\gamma$  is as in lemma 2.4.

*Proof.* Using lemma 2.3, we may reduce the estimate to the corresponding estimate for a single element. Defining  $\hat{\mathbf{n}} = \beta^\perp/|\beta|$ , we may write the bilinear form  $b(v, w)_t$  as

$$\begin{aligned} b(v, w)_t &= \int_t v_{\hat{\mathbf{n}}} w_{\hat{\mathbf{n}}} dx + (1 + \delta_t) \int_t v_n w_n dx \\ &= b_1(v, w) + (1 + \delta_t) b_2(v, w) \end{aligned}$$

Strengthened Cauchy inequalities for both  $b_1(v, w)$  and  $b_2(v, w)$  can be found using lemma 2.4. These can then be combined using lemma 2.3 to make an estimate for  $b(v, w)_t$ , completing the proof.  $\square$

We emphasize that, despite the fact that the inner product and the corresponding natural norm are both mesh dependent, the bound in the strengthened Cauchy inequality does not depend on the mesh, on the magnitude of  $\beta$  or its direction, or the size of the upwinding parameter  $c_0$ .

Our next lemma is concerned with estimating the size of the convection term.

**LEMMA 2.6.** *Let  $b(\cdot, \cdot)$  and  $b(\cdot, \cdot)_t$  be defined as in theorem 2.5, and let  $\|v\|_t^2 = b(v, v)_t$ . Then for all  $v \in V_t$  and all  $w \in W_t$ , we have the estimate*

$$|(\beta \cdot \nabla v, w)_t| \leq \nu_t \|v\|_t \|w\|_t$$

where

$$\nu_t \leq \frac{Ch_t |\beta|}{\sqrt{1 + \delta_t}}$$

where  $C$  depends only on the shape regularity of  $t$ .

*Proof.* The proof follows from the estimates

$$\begin{aligned} \|\beta \cdot \nabla v\|_{\mathcal{L}^2(t)} &\leq \frac{|\beta|}{\sqrt{1 + \delta_t}} \|v\|_t \\ \|w\|_{\mathcal{L}^2(t)} &\leq Ch_t \|\nabla \cdot w\|_{\mathcal{L}^2(t)} \\ &\leq Ch_t \|w\|_t \end{aligned}$$

The second inequality can be verified by making a scaling and translation mapping of  $t$  to an element of diameter  $h = 1$ . This establishes the dependence on  $h_t$ .

Since a nonzero function  $w \in W_t$  is necessarily oscillatory, a subsequent mapping to the reference element establishes the result. We also note that the  $3 \times 3$  matrix corresponding to the inner product  $(\beta \cdot \nabla v, w)_t$  has rank one, and exact evaluation of  $\nu_t$  could be made using lemma 2.1.  $\square$

The importance of this result is that one may control the size of  $\nu_t$  by controlling the size of the coefficient  $c_0$  in the upwinding term  $\delta_t$ . In particular, for  $c_0$  sufficiently large, one can make  $\nu_t$  arbitrarily small, independent of the size of  $h_t|\beta|$ .

**THEOREM 2.7.** *Let  $b(v, w)$  be defined as in theorem 2.5 Then for all  $v \in V$  and all  $w \in W$ , we have the estimates*

$$\begin{aligned} |(\beta \cdot \nabla v, w)| &\leq \nu \|v\| \|w\| \\ |(\beta \cdot \nabla w, v)| &\leq \nu \|v\| \|w\| \end{aligned}$$

where

$$\nu = \max_{t \in \mathcal{T}} \nu_t$$

*Proof.* The first inequality is proved using lemma 2.6 and an argument entirely analogous to that used in proving lemma 2.3. The second inequality follows by integration by parts (noting  $\beta$  is constant and homogeneous Dirichlet boundary conditions) yielding

$$(\beta \cdot \nabla w, v) = -(\beta \cdot \nabla v, w)$$

$\square$

**3. A Two Level Iteration.** In this section, we will consider the solution of the linear system: find  $u \in \mathcal{M}_2$  such that

$$(12) \quad a_s(u, v) = f(v)$$

for all  $v \in \mathcal{M}_2$ . Here  $f$  is a linear functional on  $\mathcal{M}_2$  and  $a_s(v, w)$  is defined in (6)

We will consider the hierarchical decomposition of  $\mathcal{M}_2$  as  $\mathcal{M}_2 = V \oplus W$ . Our two level iteration will be a block Gauss-Seidel iteration. Let the iterates  $u_k \in \mathcal{M}_2$  be decomposed as  $u_k = v_k + w_k$ , where  $v_k \in V$  and  $w_k \in W$ . Given  $u_0$ , the remaining iterates are defined by

$$(13) \quad a_s(v_{k+1}, \phi) = f(\phi) - a_s(w_k, \phi)$$

for all  $\phi \in V$  and

$$(14) \quad a_s(w_{k+1}, \phi) = f(\phi) - a_s(v_{k+1}, \phi)$$

for all  $\phi \in W$ . Note that block Gauss-Seidel and symmetric block Gauss-Seidel amount to essentially the same algorithm when only two levels are used.

Let  $e_k \in \mathcal{M}_2$  denote the error at the  $k$ th step of (13)-(14). The error can be decomposed as  $e_k = \rho_k + \eta_k$ , where  $\rho_k \in V$  and  $\eta_k \in W$ . Then the error  $e_k$  propagates as

$$(15) \quad a_s(\rho_{k+1} + \eta_k, \phi) = 0$$

for all  $\phi \in V$  and

$$(16) \quad a_s(\rho_{k+1} + \eta_{k+1}, \phi) = 0$$

for all  $\phi \in W$ .

We define the symmetric, positive definite bilinear form  $b_s(v, w)$  by

$$(17) \quad b_s(v, w) = \int_{\Omega} \nabla \cdot v \nabla \cdot w + \delta_t v_n w_n dx$$

This is just the symmetric part of  $a_s(v, w)$

$$b_s(v, w) = \frac{a_s(v, w) + a_s(w, v)}{2}$$

The corresponding energy norm  $||| \cdot |||$  is defined by

$$|||u|||^2 = b_s(u, u)$$

It is reasonably straightforward to analyze the convergence of (13)-(14). Taking  $\phi = \rho_{k+1}$  in (15), and applying theorems 2.5 and 2.7 we have

$$|||\rho_{k+1}||| \leq (\gamma + \nu) |||\eta_k|||$$

Similarly, taking  $\phi = \eta_{k+1}$  in (16),

$$|||\eta_{k+1}||| \leq (\gamma + \nu) |||\rho_{k+1}|||$$

with  $\gamma$  and  $\nu$  defined as in theorems 2.5 and 2.7, respectively.

Combining these estimates we have

$$\begin{aligned} |||\rho_{k+1}||| &\leq (\gamma + \nu)^2 |||\rho_k||| \\ |||\eta_{k+1}||| &\leq (\gamma + \nu)^2 |||\eta_k||| \end{aligned}$$

or

$$(18) \quad \{ |||\rho_{k+1}|||^2 + |||\eta_{k+1}|||^2 \}^{\frac{1}{2}} \leq (\gamma + \nu)^2 \{ |||\rho_k|||^2 + |||\eta_k|||^2 \}^{\frac{1}{2}}$$

The norm used in (18) is mesh dependent, and also depends on the hierarchical decomposition. Using theorem 2.5, we may establish the norm comparability

$$(1 - \gamma) \{ |||v|||^2 + |||w|||^2 \} \leq |||u|||^2 \leq (1 + \gamma) \{ |||v|||^2 + |||w|||^2 \}$$

for all  $u = v + w \in \mathcal{M}_2$ . This implies

$$(19) \quad |||e_k||| \leq \sqrt{\frac{1 + \gamma}{1 - \gamma}} (\gamma + \nu)^{2k} |||e_0|||$$

This estimate now is independent of the hierarchical decomposition, but is still mesh dependent. We note that  $\gamma < 1$  independent of the mesh,  $\beta$ , and the size of the upwinding constant  $c_0$ . The constant  $\nu$  can be made arbitrarily small by a sufficiently large value for  $c_0$ , which is independent of  $h_t|\beta|$ . Thus, for sufficiently strong upwinding, we may force convergence of the two level scheme. Despite the fact that the energy norm is mesh dependent, our bound on the convergence rate does not depend on the mesh.

**4. The block symmetric Gauss-Seidel iteration.** In this section we make a general analysis of the behavior of the block symmetric Gauss-Seidel iteration for the non-symmetric case. We emphasize that the hierarchical basis multigrid method is mathematically equivalent to such an iteration [5], if the block structure of the matrix corresponds to the hierarchical decomposition

$$\mathcal{M}_k = V_1 \oplus V_2 \oplus \dots \oplus V_k$$

Let  $A = D + L + U$  be a block  $k \times k$  matrix, with  $D$ ,  $L$ , and  $U$  being block diagonal, lower triangular, and upper triangular, respectively. We assume that the symmetric parts of  $A$  and  $D$  are positive definite, as they are for the case we have in mind. We define

$$(20) \quad \begin{aligned} \hat{D} &= (D + D^t)/2 \\ L &= T + E \\ U &= T^t - E^t \\ S &= (D - D^t)/2 \\ B &= (A + A^t)/2 = \hat{D} + T + T^t \end{aligned}$$

and

$$R = \begin{bmatrix} 0 & & & & \\ & S_2 & & & \\ & & \ddots & & \\ & & & S_{k-1} & \\ & & & & 0 \end{bmatrix}$$

A natural norm to use in analyzing the method is the norm induced by  $\hat{D}$ . Thus we are led to define  $\hat{A} = \hat{D}^{-1/2} A \hat{D}^{-1/2}$ , and similarly  $\hat{B}$ ,  $\hat{T}$ ,  $\hat{L}$ ,  $\hat{U}$ ,  $\hat{S}$ ,  $\hat{E}$ , and  $\hat{R}$ .

We consider the solution of the system of equations

$$(21) \quad Au = F$$

by the symmetric Gauss-Seidel iteration

$$(22) \quad \begin{aligned} (D + L)(u_{j+\frac{1}{2}} - u_j) &= F - Au_j \\ (D + U)(u_{j+1} - u_{j+\frac{1}{2}}) &= F - Au_{j+\frac{1}{2}} \end{aligned}$$

where  $u_0$  is given.

We define the error  $\varepsilon_j$  by

$$(23) \quad \varepsilon_j = u_j - u$$

Then a straightforward calculation shows that the error propagates as

$$(24) \quad \begin{aligned} \varepsilon_{j+1} &= \{(D + L)D^{-1}(D + U)\}^{-1}LD^{-1}U\varepsilon_j \\ &= (I + D^{-1}U)^{-1}(I + D^{-1}L)^{-1}(D^{-1}L)(D^{-1}U)\varepsilon_j \end{aligned}$$

From (24) we see that

$$(25) \quad \hat{D}^{1/2}\varepsilon_j = Q(G_L G_U)^j Q^{-1} \hat{D}^{1/2}\varepsilon_0$$

where

$$\begin{aligned}
(26) \quad Q &= \hat{D}^{1/2}(I + D^{-1}U)^{-1}\hat{D}^{-1/2} \\
G_L &= \hat{D}^{1/2}(I + D^{-1}L)^{-1}(D^{-1}L)\hat{D}^{-1/2} \\
&= (I + \hat{S} + \hat{L})^{-1}\hat{L} \\
G_U &= \hat{D}^{1/2}(I + D^{-1}U)^{-1}(D^{-1}U)\hat{D}^{-1/2}
\end{aligned}$$

Our analysis fundamentally consists of estimating the norm  $\|G_L\|_{\ell_2}$  (the estimates for  $\|G_U\|_{\ell_2}$  are identical, and the proofs are quite analogous). We begin with

LEMMA 4.1. *Let  $\hat{L}$ ,  $\hat{S}$  and  $\hat{R}$  be the block  $k \times k$  matrices defined above. Then*

$$(27) \quad \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t (I + \hat{S} + \hat{L})(I + \hat{S} + \hat{L})^t x} \leq \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t (I + \hat{R} + \hat{L})(I + \hat{R} + \hat{L})^t x}$$

*Proof.* First, notice that since  $\hat{L}$  is block lower triangular, the numerator  $x^t \hat{L} \hat{L}^t x$  is independent of the first component,  $x_1$ , of  $x$ . Since  $\hat{S}_1$  is skew,  $(I_{n_1} + \hat{S}_1^t)$  is nonsingular. Thus, without loss of generality, we can replace the first component  $x$  by  $(I_{n_1} + \hat{S}_1^t)^{-1}x_1$ , which is equivalent to replacing the 1 – 1 block in the matrix  $(I + \hat{S}^t)$  by  $I_{n_1}$ . Thus

$$(28) \quad \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t (I + \hat{S} + \hat{L})(I + \hat{S} + \hat{L})^t x} = \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t (I + \bar{R} + \hat{L})(I + \bar{R} + \hat{L})^t x}$$

where

$$\bar{R} = \begin{bmatrix} 0 & & & & \\ & S_2 & & & \\ & & \ddots & & \\ & & & S_{k-1} & \\ & & & & S_k \end{bmatrix}$$

Next, by direct calculation, and noting the last block column and first block row of  $\hat{L}$  are zero, we find

$$\begin{aligned}
x^t (I + \bar{R} + \hat{L})(I + \bar{R} + \hat{L})^t x &= x^t (I + \hat{R} + \hat{L})(I + \hat{R} + \hat{L})^t x + x_k^t \hat{S}_k \hat{S}_k^t x_k \\
&\geq x^t (I + \hat{R} + \hat{L})(I + \hat{R} + \hat{L})^t x
\end{aligned}$$

completing the proof.  $\square$

LEMMA 4.2. *Let  $\hat{U}$ ,  $\hat{S}$  and  $\hat{R}$  be the block  $k \times k$  matrices defined above. Then*

$$(29) \quad \sup_{x \neq 0} \frac{x^t \hat{U} \hat{U}^t x}{x^t (I + \hat{S} + \hat{U})(I + \hat{S} + \hat{U})^t x} \leq \sup_{x \neq 0} \frac{x^t \hat{U} \hat{U}^t x}{x^t (I + \hat{R} + \hat{U})(I + \hat{R} + \hat{U})^t x}$$

*Proof.* The proof is analogous to lemma 4.1.  $\square$

We remark that, in the case of only two levels,  $\hat{R} = 0$ ; this is one of the features which simplifies the analysis in that case. Using the fact that  $\hat{R}$  is skew, we have

$$\begin{aligned}
x^t (I + \hat{R} + \hat{L})(I + \hat{R} + \hat{L})^t x &= x^t (\hat{B} + \hat{L} \hat{L}^t + 2\hat{E} + 2\hat{R} \hat{L}^t + \hat{R} \hat{R}^t) x \\
&\geq x^t (\hat{B} + \hat{L} \hat{L}^t + 2\hat{E} + 2\hat{R} \hat{L}^t) x
\end{aligned}$$

Thus we have

$$(30) \quad \|G_L\|_{\ell_2}^2 \leq \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t (\hat{B} + \hat{L} \hat{L}^t + 2\hat{E} + 2\hat{R} \hat{L}^t) x}$$

The remainder of our analysis will focus on technical estimates for the various terms on the right hand side of (30). In particular, note that since  $\hat{B}$  is symmetric, positive definite

$$\sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t (\hat{B} + \hat{L} \hat{L}^t) x} < 1$$

Obtaining a more precise estimate for this term is very similar to the symmetric, positive definite case [5]. The remaining terms,  $2x^t (\hat{E} + \hat{R} \hat{L}^t) x$ , involve matrices whose norm can be made small by choosing a sufficiently large value of the upwinding parameter  $c_0$ .

**5. Estimates For The Artificial-Diffusion Method.** In this section we make some estimates to be used in bounding the terms in (30). Let

$$\mathcal{M}_\ell = V_1 \oplus V_2 \oplus \dots \oplus V_\ell$$

for  $1 \leq \ell \leq k$ , and let

$$\mathcal{N}_\ell = V_{\ell+1} \oplus V_{\ell+2} \oplus \dots \oplus V_k$$

for  $1 \leq \ell \leq k-1$ , with  $\mathcal{N}_k = \emptyset$ . Then

$$\mathcal{M}_k = \mathcal{M}_\ell \oplus \mathcal{N}_\ell$$

for  $1 \leq \ell \leq k$ .

We will consider the solution of the linear system: find  $u \in \mathcal{M}_k$  such that

$$(31) \quad a_d(u, v) = f(v)$$

for all  $v \in \mathcal{M}_k$ . Here  $f$  is a linear functional on  $\mathcal{M}_k$  and  $a_d(v, w)$  is the bilinear form given by (4). We define the bilinear form  $b_d(\cdot, \cdot)$  by

$$\begin{aligned} b_d(v, w) &= \frac{a_d(v, w) + a_d(w, v)}{2} \\ &= \int_{\Omega} (1 + \delta) \nabla v \cdot \nabla w \, dx \end{aligned}$$

and set

$$|||v|||^2 = b_d(v, v)$$

We begin with

LEMMA 5.1. *let  $v_i \in V_i$  and  $v_j \in V_j$ . Let  $1 \leq i \leq k$  and  $1 \leq j \leq k$ , excluding the case  $i = j = 1$ . Then*

$$(32) \quad |(\beta \cdot \nabla v_i, v_j)| \leq \frac{c|\beta| \min(h_i, h_j)}{1 + \delta} |||v_i||| |||v_j|||$$

*Proof.* First suppose that  $j \geq i$ . In particular, this implies  $j > 1$ . The proof follows from the elementary estimates

$$(33) \quad \|\beta \cdot \nabla v_i\|_{\mathcal{L}^2} \leq \frac{c|\beta|}{\sqrt{1+\delta}} \|v_i\|$$

$$(34) \quad \|v_j\|_{\mathcal{L}^2} \leq \frac{ch_j}{\sqrt{1+\delta}} \|v_j\|$$

Inequality (34) follows from the necessarily oscillatory behavior of any nonzero function in  $V_j$ , and is established by an element-by-element analysis using the triangles of level  $j-1$ .

If  $j < i$ , then intergration by parts gives

$$(\beta \cdot \nabla v_i, v_j) = -(\beta \cdot \nabla v_j, v_i)$$

and we may repeat the argument with  $i$  and  $j$  interchanged.  $\square$

LEMMA 5.2. *let  $v_i \in V_i$ ,  $i > 1$  and  $z \in \mathcal{M}_k$ . Then*

$$(35) \quad |(\beta \cdot \nabla z, v_i)| \leq \frac{c|\beta|h_i}{1+\delta} \|z\| \|v_i\|$$

*Proof.* The proof is analogous to Lemma 5.1.  $\square$

We now prove a strengthened Cauchy inequality for the bilinear form  $b_d(\cdot, \cdot)$ .

LEMMA 5.3. *Let  $v_\ell \in \mathcal{M}_\ell$  and  $w_\ell \in \mathcal{N}_\ell$ ,  $1 \leq \ell \leq k-1$ . Then there exists a positive constant  $\gamma_\ell$ ,  $1 \leq \ell \leq k-1$  such that*

$$(36) \quad |b_d(v_\ell, w_\ell)| \leq \gamma_\ell \|v_\ell\| \|w_\ell\|$$

and

$$(37) \quad \gamma_\ell \leq 1 - \frac{c}{k-\ell}$$

where  $c$  is a positive constant depending only on shape regularity of the elements in  $\mathcal{T}_k$ .

*Proof.* It is sufficient to prove (36) triangle by triangle, for each triangle  $t \in \mathcal{T}_\ell$ . For  $t \in \mathcal{T}_\ell$ , we note that  $v_\ell$  restricted to  $t$  is just a linear polynomial and  $w_\ell$  is a piecewise linear polynomial which is zero at the vertices of  $t$ . In particular, note that the constant function on  $t$  is contained in the space  $V_{\ell t}$  and not  $W_{\ell t}$ . Because  $b_d(c, \chi)_t = 0$  for any constant  $c$  and  $\chi \in V_{\ell t} \oplus W_{\ell t}$ , without loss in generality, we may assume that  $v_\ell = 0$  at one vertex of  $t$ , in order to exclude the constant function.

For such a function  $v_\ell + w_\ell$ , we have the embedding

$$\begin{aligned} \|v_\ell - w_\ell\|_{\mathcal{L}^\infty(t)} &\leq c \left| \log \frac{h_k}{h_\ell} \right|^{1/2} \|v_\ell - w_\ell\|_{\mathcal{H}^1(t)} \\ &\leq c(k-\ell)^{1/2} \|v_\ell - w_\ell\|_{\mathcal{H}^1(t)} \end{aligned}$$

and the norm comparability

$$(38) \quad \frac{c}{\sqrt{1+\delta}} \|v_\ell - w_\ell\|_t \leq \|v_\ell - w_\ell\|_{\mathcal{H}^1(t)} \leq \frac{c'}{\sqrt{1+\delta}} \|v_\ell - w_\ell\|_t$$

Now, a straightforward calculation shows

$$\gamma_\ell \leq \max_{t \in \mathcal{T}_\ell} \gamma_{\ell t}$$

where

$$\gamma_{\ell t} = \sup_{\substack{\|v_\ell\|_t = 1 \\ \|w_\ell\|_t = 1}} b_d(v_\ell, w_\ell)_t$$

Finally, to estimate  $\gamma_{\ell t}$ , we have

$$\begin{aligned} \gamma_{\ell t} &= \sup_{\substack{\|v_\ell\|_t = 1 \\ \|w_\ell\|_t = 1}} b_d(v_\ell, w_\ell)_t \\ &= \sup \left\{ 1 - \frac{1}{2} \|v_\ell - w_\ell\|_t^2 \right\} \\ &\leq \sup \left\{ 1 - \frac{c(1+\delta)}{k-\ell} \|v_\ell - w_\ell\|_{\mathcal{L}^\infty(t)}^2 \right\} \end{aligned}$$

Let  $p_j$ ,  $1 \leq j \leq 3$  denote the vertices of  $t$ , and let  $v_\ell(p_3) = 0$ . Since  $w_\ell(p_j) = 0$  for  $1 \leq j \leq 3$ , we have

$$\begin{aligned} \|v_\ell - w_\ell\|_{\mathcal{L}^\infty(t)} &\geq \max\{|v_\ell(p_1)|, |v_\ell(p_2)|\} \\ &\geq \frac{c}{\sqrt{1+\delta}} \end{aligned}$$

and finally,

$$\gamma_{\ell t} \leq 1 - \frac{c}{k-\ell}$$

□

We now make an estimate for the hierarchical basis method in the symmetric, positive definite case.

LEMMA 5.4. *Let  $T$ ,  $B$ ,  $\hat{D}$ ,  $\hat{T}$  and  $\hat{B}$  be the block  $k \times k$  matrices defined above. Then*

$$(39) \quad \sup_{x \neq 0} \frac{x^t \hat{T} \hat{T}^t x}{x^t \hat{B} x} \leq ck^2$$

*Proof.* To prove (39), we begin by noting

$$\begin{aligned} \sup_{x \neq 0} \frac{x^t \hat{T} \hat{T}^t x}{x^t \hat{B} x} &= \sup_{x \neq 0} \frac{x^t T \hat{D}^{-1} T^t x}{x^t B x} \\ &= \sup_{x \neq 0} \frac{y^t \hat{D} y}{x^t B x} \end{aligned}$$

where  $\hat{D}y = T^t x$ .

Let  $w = w_1 + w_2 + \dots + w_k$  denote the piecewise linear function corresponding to the vector  $x^t = (x_1^t, x_2^t, \dots, x_k^t)$  and  $v = v_1 + v_2 + \dots + v_k$  denote the piecewise linear function corresponding to the vector  $y^t = (y_1^t, y_2^t, \dots, y_k^t)$ . Furthermore, we define  $z_i = w_{i+1} + w_{i+2} + \dots + w_k$ . Then  $\hat{D}y = T^t x$  corresponds to

$$(40) \quad b_d(v_i, \chi) = b_d(z_i, \chi)$$

for  $1 \leq i \leq k-1$ ,  $v_k = 0$  and  $\chi \in V_i$ . Taking  $\chi = v_i$ , we have

$$(41) \quad |||v_i||| \leq |||z_i|||$$

Thus

$$\begin{aligned} y^t \hat{D}y &= \sum_{i=1}^k |||v_i|||^2 \\ &\leq \sum_{i=1}^k |||z_i|||^2 \\ &\leq \sum_{i=1}^k \frac{1}{1 - \gamma_i} |||w|||^2 \\ &\leq c k^2 |||w|||^2 \\ &= c k^2 x^t Bx \end{aligned}$$

□

We now generalize Lemma 5.4 to the nonsymmetric case.

LEMMA 5.5. *Let  $L, B, E, R, \hat{D}, \hat{L}, \hat{B}, \hat{E}$ , and  $\hat{R}$  be the block  $k \times k$  matrices defined above. Define  $\theta$  as*

$$(42) \quad \theta = \frac{\sqrt{k} h_1 |\beta|}{1 + \delta}$$

Then

$$(43) \quad \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t \hat{B} x} \leq c(k + \theta)^2$$

$$(44) \quad \sup_{x \neq 0} \frac{x^t \hat{E} \hat{E}^t x}{x^t \hat{B} x} \leq c\theta^2$$

$$(45) \quad \sup_{x \neq 0} \frac{x^t \hat{R} \hat{R}^t x}{x^t \hat{B} x} \leq c\theta^2$$

*Proof.* The proof is analogous to that for Lemma 5.4, and we adopt the same notation as in that proof. We begin with a proof of (44). Following the proof for Lemma 5.4, we have

$$\sup_{x \neq 0} \frac{x^t \hat{E} \hat{E}^t x}{x^t \hat{B} x} = \sup_{x \neq 0} \frac{y^t \hat{D}y}{x^t Bx}$$

where  $\hat{D}y = E^t x$ . This corresponds to

$$b_d(v_i, \chi) = (\beta \cdot \nabla z_i, \chi)$$

for  $1 \leq i \leq k-1$ ,  $v_k = 0$  and  $\chi \in V_i$ . Taking  $\chi = v_i$ , we have, from Lemma 5.2,

$$\|v_i\| \leq \frac{ch_i|\beta|}{1+\delta} \|z_i\|$$

Thus

$$\begin{aligned} y^t \hat{D}y &= \sum_{i=1}^k \|v_i\|^2 \\ &\leq \sum_{i=1}^k \left( \frac{ch_i|\beta|}{1+\delta} \right)^2 \|z_i\|^2 \\ &\leq \sum_{i=1}^k \left( \frac{ch_i|\beta|}{1+\delta} \right)^2 \frac{1}{1-\gamma_i} \|w\|^2 \\ &\leq c\theta^2 \|w\|^2 \\ &= c\theta^2 x^t Bx \end{aligned}$$

The proof of (43) is now a simple application of Lemma 5.4 and (44). For the proof of (45) we have

$$\sup_{x \neq 0} \frac{x^t \hat{R} \hat{R}^t x}{x^t \hat{B}x} = \sup_{x \neq 0} \frac{y^t \hat{D}y}{x^t Bx}$$

where  $\hat{D}y = R^t x$ . This corresponds to

$$b_d(v_i, \chi) = (\beta \cdot \nabla w_i, \chi)$$

for  $2 \leq i \leq k-1$ . Taking  $\chi = v_i$ , we have, from Lemma 5.1,

$$\|v_i\| \leq \frac{ch_i|\beta|}{1+\delta} \|w_i\|$$

The remainder of the proof is basically the same as for (44). However, in this argument, it was important that the diagonal block  $S_1$  is excluded (by virtue of Lemmas 4.1 and 4.2), since Lemma 5.1 does not apply to the case  $i = j = 1$ .  $\square$  We remark that the estimates (43)-(45) exhibit a dependence on the number of levels. In particular, the numerator  $c|\beta|h_1 = O(|\beta|h_k 2^k)$ , while  $\delta = c_0 O(|\beta|h_k + |\beta|^2 h_k^2)$ . To make such terms small requires either increasing the upwinding parameter  $c_0$  in a level-dependent fashion (somewhat unnatural, since larger values of  $c_0$  would correspond to more refined meshes, where presumably less upwinding should be necessary for good approximation), or alternatively, placing a restriction on  $h_1$ , e.g. requiring  $|\beta|h_1$  to be small in comparison with one. This amounts to the requirement that the coarsest mesh be sufficiently fine, and is typical of restrictions arising in the analysis of regular multigrid methods for nonsymmetric problems.

LEMMA 5.6. *Let  $B$  and  $\hat{D}$  be the block  $k \times k$  matrices defined above. Then*

$$(46) \quad \frac{x^t Bx}{x^t \hat{D}x} \leq ck$$

$$(47) \quad \frac{x^t \hat{D}x}{x^t Bx} \leq ck^2$$

*Proof.* Let  $w = w_1 + w_2 + \cdots + w_k$  denote the piecewise linear function corresponding to the vector  $x^t = (x_1^t, x_2^t, \cdots, x_k^t)$ . The proof of (46) is straightforward from the following estimate

$$\begin{aligned} x^t B x &= \|w\|^2 \\ &\leq c k \sum_{i=1}^k \|w_i\|^2 \\ &= c k x^t \hat{D} x \end{aligned}$$

The estimate in (47) is similar to estimate (39). Let  $z_i = w_1 + w_2 + \cdots + w_{i-1}$ . Then

$$\begin{aligned} x^t \hat{D} x &= \sum_{i=1}^k \|w_i\|^2 \\ &\leq \frac{1}{1 - \gamma_{k-1}} \sum_{i=1}^k \|z_i\|^2 \\ &\leq \frac{1}{1 - \gamma_{k-1}} \sum_{i=1}^k \left( \frac{1}{1 - \gamma_i} \right) \|w\|^2 \\ &\leq c k^2 x^t B x \end{aligned}$$

□

We can now make our final estimate.

**THEOREM 5.7.** *Let  $G_L$  be defined as in (26). Then if  $\theta$  given by (42) is sufficiently small,*

$$(48) \quad \|G_L\|_{\ell_2}^2 \leq \frac{c_1(k + \theta)^2}{1 + c_1(k + \theta)^2 - c_2\theta(k + \theta)}$$

where  $c_1$  and  $c_2$  are positive constants depending only on the geometry of the elements.

*Proof.* Use Lemmas 5.4, 5.5 and 5.6, to bound the various terms appearing on the right hand side of (30). □

We pause to make several remarks. First, a similar estimate to (48) holds for  $G_U$ ,

$$\|G_U\|_{\ell_2}^2 \leq \frac{c_1(k + \theta)^2}{1 + c_1(k + \theta)^2 - c_2\theta(k + \theta)}$$

Both estimates reduce to that of the symmetric problem when  $\theta = 0$ .

Second, to guarantee convergence using (48) we must require

$$(49) \quad c_2\theta(k + \theta) < 1$$

We can satisfy (49) by forcing  $h_1|\beta|$  to be small (e.g., making the coarsest mesh fine enough) or by making the upwinding parameter  $c_0$ , hence  $\delta$ , sufficiently large. In either case there is a mesh dependent term  $\sqrt{k}2^k$  which must be dominated. As we do not know if (48) is sharp, we cannot characterize for certain the dependence of the rate of convergence on the number of levels, but one should expect at least a weak dependence for problems of practical interest.

**6. Estimates For The Streamline-Diffusion Method.** In this section, we estimate convergence rates for the streamline diffusion method. Since most of the proofs of the results are similar to the proofs of the estimates made in section 5, we will only state the modifications to be made.

We define  $V_\ell$ ,  $\mathcal{M}_\ell$  and  $\mathcal{N}_\ell$  as in section 5 and we consider the solution of the equation

$$(50) \quad a_s(u, v) = f(v)$$

where  $a_s(v, w)$  is the bilinear form defined by (6), and  $f(v)$  is defined as in (31). Similarly, the symmetric part of  $a_s(v, w)$  is given by

$$\begin{aligned} b_s(v, w) &= \frac{a_s(v, w) + a_s(w, v)}{2} \\ &= \int_{\Omega} \nabla v \cdot \nabla w + \delta_t v_n w_n \, dx \end{aligned}$$

The differences in the results between the artificial diffusion case and the streamline diffusion case are due mainly to differences in the norm comparability (38). Unfortunately, in the streamline diffusion case the embedding is not so strong as (38). It is replaced by the following weaker norm comparability

$$(51) \quad \frac{c}{\sqrt{1+\delta}} \| \|v_\ell - w_\ell\| \|_t \leq \| \|v_\ell - w_\ell\| \|_{\mathcal{H}^1(t)} \leq c' \| \|v_\ell - w_\ell\| \|_t$$

which leads to slightly weaker results.

We first begin with a lemma similar to lemma 5.1

LEMMA 6.1. *let  $v_i \in V_i$  and  $v_j \in V_j$ . Let  $1 \leq i \leq k$  and  $1 \leq j \leq k$ , excluding the case  $i = j = 1$ . Then*

$$(52) \quad |(\beta \cdot \nabla v_i, v_j)| \leq \frac{c|\beta| \min(h_i, h_j)}{\sqrt{1+\delta}} \| \|v_i\| \| \| \|v_j\| \|$$

*Proof.* The proof is exactly the same as lemma 5.1, except for estimate (34), which becomes

$$(53) \quad \| \|v_j\| \|_{\mathcal{L}^2} \leq ch_j \| \|v_j\| \|$$

□

LEMMA 6.2. *let  $v_i \in V_i$ ,  $i > 1$  and  $z \in \mathcal{M}_k$ . Then*

$$(54) \quad |(\beta \cdot \nabla z, v_i)| \leq \frac{c|\beta|h_i}{\sqrt{1+\delta}} \| \|z\| \| \| \|v_i\| \|$$

*Proof.* The proof is analogous to Lemma 6.1. □ We now prove a strengthened Cauchy inequality (similar to the one in lemma 5.3) for the bilinear form  $b_s(\cdot, \cdot)$ .

LEMMA 6.3. *Let  $v_\ell \in \mathcal{M}_\ell$  and  $w_\ell \in \mathcal{N}_\ell$ ,  $1 \leq \ell \leq k-1$ . Then there exists a positive constant  $\gamma_\ell$ ,  $1 \leq \ell \leq k-1$  such that*

$$(55) \quad |b_s(v_\ell, w_\ell)| \leq \gamma_\ell \| \|v_\ell\| \| \| \|w_\ell\| \|$$

and

$$(56) \quad \gamma_\ell \leq 1 - \frac{c}{(1+\delta)(k-\ell)}$$

where  $c$  is a positive constant depending only on shape regularity of the elements in  $\mathcal{T}_k$ .

*Proof.* The proof is analogous to lemma 5.3, where the norm comparability (38) is replaced by the norm comparability (51).  $\square$  In the symmetric positive definite case, the hierarchical basis multigrid method yields the following estimate

LEMMA 6.4. *Let  $T, B, \hat{D}, \hat{T}$  and  $\hat{B}$  be the block  $k \times k$  matrices defined above. Then*

$$(57) \quad \sup_{x \neq 0} \frac{x^t \hat{T} \hat{T}^t x}{x^t \hat{B} x} \leq c(1 + \delta)k^2$$

*Proof.* The proof is analogous to lemma 5.4.  $\square$

We now generalize Lemma 6.4 to the nonsymmetric case.

LEMMA 6.5. *Let  $L, B, E, R, \hat{D}, \hat{L}, \hat{B}, \hat{E}$ , and  $\hat{R}$  be the block  $k \times k$  matrices defined above. Define  $\theta$  as*

$$(58) \quad \theta = \frac{\sqrt{k} h_1 |\beta|}{\sqrt{1 + \delta}}$$

Then

$$(59) \quad \sup_{x \neq 0} \frac{x^t \hat{L} \hat{L}^t x}{x^t \hat{B} x} \leq c(1 + \delta)(k + \theta)^2$$

$$(60) \quad \sup_{x \neq 0} \frac{x^t \hat{E} \hat{E}^t x}{x^t \hat{B} x} \leq c(1 + \delta)\theta^2$$

$$(61) \quad \sup_{x \neq 0} \frac{x^t \hat{R} \hat{R}^t x}{x^t \hat{B} x} \leq c(1 + \delta)\theta^2$$

*Proof.* The proof is analogous to Lemma 5.5.  $\square$  Similarly to the artificial diffusion case, the estimates (59)-(61) exhibit a dependence on the number of levels. However, the terms in (60) and (61) do not depend any longer on the upwinding parameter  $c_0$  and the term in the right hand side of (59) actually increases if we increase the upwinding parameter. But requiring  $|\beta|h_1$  to be small, is still a sufficient condition to make such terms small.

LEMMA 6.6. *Let  $B$  and  $\hat{D}$  be the block  $k \times k$  matrices defined above. Then*

$$(62) \quad \frac{x^t B x}{x^t \hat{D} x} \leq c k$$

$$(63) \quad \frac{x^t \hat{D} x}{x^t B x} \leq c(1 + \delta)^2 k^2$$

*Proof.* The proof is analogous to lemma 5.6.  $\square$

We can now make our final estimate.

THEOREM 6.7. *Let  $G_L$  be defined as in (26). Then if  $\theta$  given by (58) is sufficiently small,*

$$(64) \quad \|G_L\|_{\ell_2}^2 \leq \frac{c_1(1 + \delta)(k + \theta)^2}{1 + c_1(1 + \delta)(k + \theta)^2 - c_2(1 + \delta)\theta(k + \theta) - c_3(1 + \delta)^{3/2}\theta}$$

where  $c_1$ ,  $c_2$  and  $c_3$  are positive constants depending only on the geometry of the elements.

*Proof.* Use Lemmas 6.4, 6.5, and 6.6 to bound the various terms appearing on the right hand side of (30).  $\square$

A similar estimate to (64) holds for  $G_U$ ,

$$\|G_U\|_{\ell_2}^2 \leq \frac{c_1(1+\delta)(k+\theta)^2}{1+c_1(1+\delta)(k+\theta)^2 - c_2(1+\delta)\theta(k+\theta) - c_3(1+\delta)^{3/2}\theta}$$

To guarantee convergence using (64), we must require

$$(65) \quad c_2(1+\delta)\theta(k+\theta) + c_3(1+\delta)^{3/2}\theta < 1$$

As we have already noticed, increasing the upwinding parameter is not going to help in this case, the only way we can satisfy (65) is by forcing  $h_1|\beta|$  to be small (e.g., making the coarsest mesh fine enough).

## 7. Numerical Results.

**7.1. The Two-Level Method.** In this section we conduct numerical tests on the two-level method. We will consider the convection-diffusion equation on the square:

$$(66) \quad \begin{cases} -\Delta u + \beta \nabla u & = 1 & \text{in } \Omega \equiv (0,1) \times (0,1) \\ u & = 0 & \text{on } \partial\Omega \end{cases}$$

The discretization used is based on the streamline diffusion method.

This domain is first triangulated into 8 triangles, to form the level 1 grid. The mesh is then uniformly refined, by subdividing each triangle into four congruent triangles. The refinement is continued until we reach the level *clevel* grid, which will be our coarse grid. By refining one more time we will reach level *flevel*, the fine grid. In the following  $\gamma$  will denote the average rate of convergence of the two-level method. The column denoted “iteration” gives the number of iterations required to achieve the corresponding number of digits of accuracy (given in column 3), where

$$digits = -\log_{10} \left\{ \frac{\| \epsilon \|_1}{\| u_h \|_1} \right\}$$

The constant  $c_0$  is the upwinding parameter. NV is the number of vertices (roughly the size of the problem).

We first test the dependence of the convergence rate of the method on the upwinding parameter  $c_0$ . The results (see tables 1 and 2) show that we can force convergence or accelerate it, by choosing sufficiently large  $c_0$ , which is predicted by (19). Notice that the two level scheme diverges for small values of  $c_0$  (table 1).

Next we fix  $|\beta|$  and we vary its direction to illustrate the dependence of the convergence rate on the direction of  $\beta$ . The results are shown in tables 3 and 4. As predicted, the convergence rate appears to be essentially independent of the direction.

**7.2. Multilevel Method.** In this section some numerical results are given for the multi-level method applied to (66).

We solve (66) using two discretization techniques namely the artificial diffusion method and the streamline diffusion method. Both results seem to confirm our theory. First we study the dependence of the performance of the multi-level method on the

TABLE 1

<b>NV = 289</b>			
$c_0$	<i>iterations</i>	<i>digits</i>	$\gamma$
.02	-	0	-
.03	-	0	-
.04	38	4.07	0.781
.05	22	4.01	0.657
.06	18	4.19	0.585
.07	15	4.13	0.530
.08	13	4.03	0.489
.09	12	4.04	0.460
.10	12	4.34	0.434
.20	10	4.28	0.373
.30	10	4.33	0.368

The rate of convergence as a function of  $c_0$ .  
 $flevel = 4$ ,  $clevel = 3$ ,  $\beta^t = (100, 100)^t$

TABLE 2

<b>NV = 16641</b>			
$c_0$	<i>iterations</i>	<i>digits</i>	$\gamma$
.05	70	4.02	0.876
.07	16	4.00	0.562
.09	11	4.02	0.431
.1	10	4.02	0.396
.11	10	4.17	0.382
.12	10	4.26	0.374
.13	10	4.29	0.372
.14	10	4.29	0.372

The rate of convergence as a function of  $c_0$ .  
 $flevel = 7$ ,  $clevel = 6$ ,  $\beta^t = (1000, 1000)^t$

TABLE 3

<b>NV = 16641</b>			
$i$	<i>iterations</i>	<i>digits</i>	$\gamma$
0	5	4.15	0.148
1	6	4.55	0.174
2	5	4.03	0.156
3	6	4.48	0.179
4	5	4.02	0.157
5	6	4.52	0.176
6	5	4.03	0.156
7	6	4.52	0.176
8	5	4.03	0.156

The rate of convergence as a function of the direction of  $\beta$ .  
 $flevel = 7$ ,  $clevel = 6$ ,  $\beta^t = 100(\cos \theta_i, \sin \theta_i)^t$ ,  $\theta_i = \frac{i\pi}{8}$ ,  $c_0 = 1$

TABLE 4

<b>NV = 289</b>			
$i$	$iterations$	$digits$	$\gamma$
0	11	4.12	0.422
1	11	4.27	0.409
2	10	4.18	0.381
3	11	4.30	0.406
4	11	4.08	0.425
5	11	4.28	0.408
6	10	4.16	0.383
7	11	4.24	0.411
8	12	4.00	0.464

The rate of convergence as a function of the direction of  $\beta$ .  
 $flevel = 4$ ,  $clevel = 3$ ,  $\beta^t = 100(\cos \theta_i, \sin \theta_i)^t$ ,  $\theta_i = \frac{i\pi}{8}$ ,  $c_0 = .15$

TABLE 5

<b>NV = 16641</b>			
$clevel$	$iterations$	$digits$	$\gamma$
1	-	-	-
2	-	-	-
3	104	4.00	0.915
4	43	4.00	0.803
5	27	4.09	0.705
6	23	4.11	0.662

#### Streamline Diffusion

Convergence rate as a function of the number of levels.  
 $flevel = 7$ ,  $\beta^t = (1000, 1000)^t$ ,  $c_0 = .8$

coarse grid mesh. The results show a strong dependence on the size of the coarse grid mesh (see tables 5 and 6).

The influence of the direction of  $\beta$  on the convergence rate is given in tables 7 and 8. In the artificial diffusion case, the independence of the direction of  $\beta$  is well illustrated in table 8.

The dependence of the convergence rate on the upwinding parameter  $c_0$  is seen in the third set of results (see tables 9 and 10). While these results seem to agree with our theory in the case of the artificial diffusion, they out perform the result given by theorem (6.7), which we think is not sharp and can be improved.

**8. Concluding Remarks.** In this section, we make a few concluding remarks. First, we comment on the mesh dependent terms growing like  $2^k$  in the  $k$ -level estimates for both artificial diffusion and streamline diffusion cases. In some sense, this is due to our construction of the hierarchical basis. We discuss this effect in the framework of the streamline diffusion method.

In our study, we essentially began with the upwinded nodal basis functions for the finest mesh,

$$(67) \quad \phi_f + \frac{\delta_f}{|\beta|} \beta \cdot \nabla \phi_f$$

TABLE 6

<b>NV = 16641</b>			
<i>clevel</i>	<i>iterations</i>	<i>digits</i>	$\gamma$
1	33	4.11	.750
2	29	4.09	.722
3	24	4.11	0.674
4	16	4.15	0.550
5	9	4.10	0.350
6	4	4.29	0.085

**Artificial Diffusion**

*Convergence rate as a function of the number of levels.*

$$f_{level} = 7, \beta^t = (1000, 1000)^t, c_0 = .8$$

TABLE 7

<b>NV = 16641</b>			
<i>i</i>	<i>iterations</i>	<i>digits</i>	$\gamma$
0	54	4.00	0.843
1	38	4.03	0.783
2	39	4.02	0.788
3	38	4.04	0.782
4	56	4.00	0.848
5	38	4.08	0.780
6	39	4.01	0.789
7	38	4.03	0.783
8	55	4.03	0.844

**Streamline Diffusion**

*The rate of convergence as a function of the direction of  $\beta$ .*

$$f_{level} = 7, c_{level} = 4, \beta^t = 1000(\cos \theta_i, \sin \theta_i)^t, \theta_i = \frac{i\pi}{8}, c_0 = .8$$

TABLE 8

<b>NV = 16641</b>			
<i>i</i>	<i>iterations</i>	<i>digits</i>	$\gamma$
0	16	4.00	0.562
1	17	4.15	0.570
2	16	4.01	0.561
3	16	4.01	0.561
4	16	4.04	0.559
5	16	4.01	0.561
6	16	4.00	0.562
7	17	4.12	0.572
8	17	4.13	0.571

**Artificial Diffusion**

*The rate of convergence as a function of the direction of  $\beta$ .*

$$f_{level} = 7, c_{level} = 4, \beta^t = 1000(\cos \theta_i, \sin \theta_i)^t, \theta_i = \frac{i\pi}{8}, c_0 = .8$$

TABLE 9

<b>NV = 16641</b>			
$c_0$	<i>iterations</i>	<i>digits</i>	$\gamma$
.05	-	0	-
.1	-	0	-
.4	24	4.01	0.680
.5	20	4.08	0.625
.55	19	4.01	0.615
.6	19	4.02	0.614
.7	20	4.11	0.623
.8	20	4.05	0.627

**Streamline Diffusion**

*The rate of convergence as a function of  $c_0$ .*

*$flevel = 7, clevel = 4, \beta^t = (100, 100)^t$*

TABLE 10

<b>NV = 16641</b>			
$c_0$	<i>iterations</i>	<i>digits</i>	$\gamma$
.05	-	0	-
.1	-	0	-
.4	19	4.06	0.609
.5	18	4.02	0.597
.55	18	4.04	0.596
.6	18	4.04	0.596
.7	18	4.05	0.595
.8	18	4.07	0.594

**Artificial Diffusion**

*The rate of convergence as a function of  $c_0$ .*

*$flevel = 7, clevel = 4, \beta^t = (100, 100)^t$*

and then constructed a hierarchical basis using the standard recursive algorithm. This results in course grid basis functions of the form

$$(68) \quad \phi_c + \frac{\delta_f}{|\beta|} \beta \cdot \nabla \phi_c$$

where the  $\phi_c$  are standard nodal basis functions for the coarse grid, as expected. The difficulty arises from the fact that  $\delta_f$  was defined for the fine grid (i.e., using the fine grid  $h_f$ ) rather than the course grid  $h_c = 2^k h_f$ . This means that stabilizing effect of the upwinding is diminished on coarser grids, resulting in the appearance of the growth factors of  $2^k$  in our convergence rate estimates.

In the framework of regular multigrid methods, this effect is well known; see, for example, [6]. The standard remedy to this problem in the case of standard multigrid methods is to use upwinding appropriate for the given level, which essentially means that in the usual case, the strength of the upwinding should grow by a factor of 2 in proceeding to each coarser level. This introduces subtle inconsistencies between the discretizations on different levels, but overall results in a very robust procedure. The inconsistencies are not harmful, since all unknowns are smoothed on the finest level,

so one expects convergence to the discrete solution on the finest level.

For the hierarchical basis method, the situation is not quite analogous. To see this, suppose we construct a hierarchical basis in the following fashion: We begin with coarse grid nodal basis functions

$$(69) \quad \phi_c + \frac{\delta_c}{|\beta|} \beta \cdot \nabla \phi_c$$

with upwinding terms appropriate to the coarse grid. Then, as new vertices are added, we add new nodal basis functions for each level, with upwinding appropriate for that level. The use of such a hierarchical basis in our analysis would certainly improve our estimates, but unfortunately, it would also alter the fine grid discretization, as we show below.

In the case of no unwinding, the standard hierarchical basis is transformed to the nodal basis for the fine mesh using the inverse of the recursive algorithm used in creating the basis in (68). However, with upwind terms present, with different strengths on different levels, the same algorithm when applied to the functions in (69) does not produce the functions (67). Indeed, the resulting functions are of the form  $\phi_f + \psi_f$ , where  $\psi_f$  can be a relatively complicated function, generally not having compact support. In any event, this hierarchical basis leads to a non-standard fine grid basis, and materially alters the discretization on the finest grid. Unlike standard multigrid methods, this inconsistency cannot be ignored, since all unknowns are *not* smoothed on the finest grid.

One possibility is to simply use a hierarchical basis iteration based on basis functions of the form (69) as a preconditioner for a linear system constructed using the basis (68). One hopes this will improve the estimates, and we expect to pursue this line of investigation. However, one practical disadvantage of this approach is that such a preconditioner cannot be constructed from the matrix elements of the fine grid nodal stiffness matrix. In the symmetric positive definite case, and in the case of the algorithms analyzed in this paper, all the necessary matrices are constructed from simple algebraic recurrence relations, starting from the matrix elements of the nodal stiffness matrix for the fine grid. We view this as a significant practical advantage of the hierarchical basis iteration, since matrix assembly is often a costly phase of a finite element calculation. It also makes installation of a hierarchical basis preconditioner in an existing finite element code easier, since some representation of the stiffness matrix for the nodal basis is usually available.

An second issue not addressed in this paper is that of inner iterations. Our analysis assumes that linear systems involving the diagonal blocks  $A_{ii}$  of (8) are all solved exactly. In practice, this is true only of the coarse grid matrix  $A_{11}$ . The remainder are solved by an inner iteration, which corresponds to the smoothing iteration in the standard multigrid method. Generally, these diagonal blocks are well conditioned, and several theoretical studies have analyzed the impact of inner iterations in the symmetric, positive definite case [4] [5]. There it is seen that, while such iterations do change the rate of convergence by changing the sizes of certain constants, the generalized condition number remains  $O(k^2)$ .

We have empirically observed behavior similar to the symmetric positive definite case for the problems studied here. Indeed, the numerical results in section 7 were obtained using one symmetric Gauss-Siedel inner iteration rather than exact solution for all blocks except  $A_{11}$ . We expect that the rates of convergence of these inner iterations will depend on the same parameters (e.g.  $|\beta|h$  and  $c_0$ ) as the outer iteration

studied here. However, we expect, as in the symmetric positive definite case, that well formulated inner iterations will exert only slight influence on the overall rate of convergence.

We note that ordering of equations within a block can have a significant practical impact on the rate of convergence of iterations like symmetric Gauss-Seidel, SSOR, or ILU, and this effect is likely to be more important in the present setting than in the symmetric positive definite case. In our experiments, diagonal blocks other than  $A_{11}$  were ordered using a bandwidth minimization ordering similar to reverse Cuthill-McKee [7]. This choice of ordering, which we also recommend and use in the symmetric positive definite case, is based mainly on empirical observation rather than theory. However, it is probable that orderings based on the magnitudes of the matrix elements as well as the sparsity structure of the  $A_{ii}$  would be better, indicating another area of future investigation. ( $A_{11}$  was ordered using the minimum degree algorithm, since a sparse LDU factorization is computed).

#### REFERENCES

- [1] R. E. BANK, *A comparison of two multi-level iterative methods for nonsymmetric and indefinite elliptic finite element equations*, SIAM J. Numer. Anal., 18 (1981), pp. 724–743.
- [2] ———, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations (Users' Guide 6.0)*, SIAM, Philadelphia, 1990.
- [3] R. E. BANK AND M. BENBOURENANE, *A Fourier analysis of the two-level hierarchical basis multigrid method for convection-diffusion equations*, in Proceedings of the Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations, SIAM, Philadelphia, 1991.
- [4] R. E. BANK AND T. F. DUPONT, *Analysis of a two level scheme for solving finite element equations*, Tech. Report CNA-159, Center for Numerical Analysis, University of Texas at Austin, 1980.
- [5] R. E. BANK, T. F. DUPONT, AND H. YSERENTANT, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427–458.
- [6] P. M. DE ZEEUW AND E. J. VAN ASSELT, *The convergence rate of multi-level algorithms applied to the convection-diffusion equation*, SIAM J. Sci. Stat. Comput., 6 (1985), pp. 492–503.
- [7] A. GEORGE AND J. LIU, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice Hall, Englewood Cliffs, NJ, 1981.
- [8] W. HACKBUSCH, *Multigrid convergence for a singular perturbation problem*, Lin. Alge. Appl., 58 (1984), pp. 125–145.
- [9] ———, *Multigrid Methods and Applications*, Springer-Verlag, Berlin, 1985.
- [10] C. JOHNSON, U. NÄVERT, AND J. PITKARANTA, *Finite element methods for linear hyperbolic problems*, Computer Meth. in Appl. Mech. and Eng., 45 (1984), pp. 285–312.
- [11] J. F. 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, Cologne 1981 (Lecture Notes in Mathematics 960), Springer-Verlag, Heidelberg, 1982, pp. 535–544.
- [12] J. MANDEL, S. MCCORMICK, AND R. E. BANK, *Variational multigrid theory*, in Multigrid Methods, SIAM, Philadelphia, 1987, pp. 131–177.
- [13] E. J. VAN ASSELT, *The multigrid method and artificial viscosity*, in Multigrid Methods: Proceedings, Cologne 1981 (Lecture Notes in Mathematics 960), Springer-Verlag, Heidelberg, 1982, pp. 313–327.
- [14] H. YSERENTANT, *Hierarchical bases of finite element spaces in the discretization of nonsymmetric elliptic boundary value problems*, Computing, 35 (1985), pp. 39–49.
- [15] ———, *On the multi-level splitting of finite element spaces for indefinite elliptic boundary value problems*, SIAM J. Numer. Anal., 23 (1986), pp. 581–595.