Randolph E. Bank[1] Todd F. Dupont[2]

Harry Yserentant[3]

The Hierarchical Basis Multigrid Method

1. Department of Mathematics
University of California at San Diego
La Jolla, California USA 92093

2. Department of Mathematics
University of Chicago
Chicago, Illinois USA 60637

3. Fachbereich Mathematik der Universität Dortmund
D-4600 Dortmund 50
Federal Republic of Germany

Preprint SC-87-2 (April 1987)

Konrad-Zuse-Zentrum für Informationstechnik;
Heilbronner Straße 10; D-1000 Berlin 31

Contents

Abstract

We derive and analyze the hierarchical basis-multigrid method for solving discretizations of self-adjoint, elliptic boundary value problems using piecewise linear triangular finite elements. The method is analyzed as a block symmetric Gauß-Seidel iteration with inner iterations, but it is strongly related to 2-level methods, to the standard multigrid V-cycle, and to earlier Jacobi-like hierarchical basis methods. The method is very robust, and has a nearly optimal convergence rate and work estimate. It is especially well suited to difficult problems with rough solutions, discretized using highly nonuniform, adaptively refined meshes.

Keywords: hierarchical basis, multigrid, finite elements, adaptive mesh refinement, preconditioned conjugate gradients, symmetric Gauß-Seidel.

Subject Classification: AMS(MOS): 65F10, 65F35, 65N20, 65N30

1. Introduction

In this work we describe and analyse the hierarchical basis multigrid method for solving selfadjoint, positive definite, elliptic boundary value problems. This method is related to standard multigrid methods [5,9], to the 2–level scheme of [1,3,6,11], and the hierarchical basis method in [14,15,16,17].

The method can be formulated as a standard multi-grid V–cycle [9], except that a smaller than normal subset of unknowns are updated during the smoothing phase at a given level. In particular, each unknown on the given finest level is uniquely associated with exactly one level, not several, and is updated only at that level. This formulation of the hierarchical basis method is especially useful when considering questions of implementation, since these aspects of multigrid methods are now well understood.

Although the multigrid–like viewpoint gives the appearance of a recursively defined algorithm, the hierarchical basis method can be mathematically formulated as a standard block iteration, albeit using the somewhat nonstandard hierarchical basis. In this respect, it resembles the 2–level scheme. The algebraic theory of block iteration, and in particular, the block symmetric Gauß-Seidel iteration considered here, is relatively straightforward. One interesting feature is that we allow for "inner" iterations to solve linear systems involving the diagonal blocks. In any event, a fairly complete algebraic analysis can be developed using only the assumptions that the matrix is symmetric and positive definite.

While we ultimately return to and use the properties of the finite element subspaces, in order to make our final estimates, the assumptions we need are all very weak and are almost always satisfied in practice. In particular, we assume shape regularity (e.g. a small angle condition) for each element but do *not* assume quasiuniformity of the global mesh. Our estimates involve only *local* ellipticity; we use no global regularity for the solution beyond the minimal H^1–regularity required for the standard weak formulation. Finally, we use local properties of piecewise polynomials.

Within this framework, we are able to show the hierarchical basis methods used as preconditioners have generalized condition numbers which grow like j^2, where j is the number of levels. This is slightly suboptimal in comparison with standard multigrid methods, where the condition numbers are uniformly bounded, and it introduces a logarithmic–like factor into the overall work estimate.

The hierarchical basis method requires $O(n)$ operations per iteration, where n is the number of unknowns on the finest level. This is the same as for standard multigrid methods. However, because of their recursive nature, one often requires geometric growth in the dimensions of the subspaces to

insure that the work on coarse levels will not dominate the overall work per cycle. Because our hierarchical basis method is just a symmetric block Gauß-Seidel iteration, its work estimate remains valid for any allowable distribution of unknowns among the levels; it relies only on the usual and normal sparsity of the global stiffness matrix for the nodal basis, (i.e. $O(1)$ nonzeroes per row).

The overall complexity of the hierarchical basis multigrid method, used as a preconditioner for the conjugate gradient iteration, is thus $O(nj|\log\epsilon|)$ operations required to reduce the initial error by the factor ϵ. In the realm of smooth model problems on rectangular regions, solved using a sequence of uniform and uniformly refined meshes, it is fair to say that the hierarchical basis method is just another pretty face in a big crowd of very good optimal or nearly optimal methods [1,9]. On the other hand, for geometrically complex regions, involving highly nonuniform and adaptively refined meshes, and/or problems with rough coefficients and solutions, there are far fewer good candidates for a simple but effective iterative method [7,8,14,15,16,17]. It is in this regime that the hierarchical basis multigrid method looks very attractive. Since its theory is based on only weak assumptions, its performance and work estimate remain essentially stable over an extremely broad range of problems. We remark that the theory developed here can be extended in straightforward fashion to other finite elements (e.g. quadrilaterals as well as triangles), other types of refinements procedures (e.g. [12]) and to higher degree polynomial spaces. We also mention that the method works well for many strongly indefinite and highly nonsymmetric problems (e.g. singular perturbation problems) but our theory does not cover such cases. Thus, as a general purpose, robust, elliptic solver, the hierarchical basis multigrid method has a lot to recommend it, and we believe its future is bright.

The remainder of this paper is organized as follows: In Section 2, we discuss the finite element discretization, introduce the nodal and hierarchical bases and derive the linear algebra problem to be solved. In Section 3, we present and analyze the block symmetric Gauß-Seidel iteration for a general symmetric, positive definite matrix. We analyze three possibilities for solving linear systems involving the diagonal blocks: direct solution, and point Gauß-Seidel and point symmetric Gauß-Seidel inner iterations. The theory developed here is strictly algebraic; the generalized condition numbers for the three cases are estimated in terms of a few constants, which ultimately are bounded, in Section 4, using the properties of the finite element space. In Section 5, we outline the precise relationship between the hierarchical basis method and the standard multigrid V-cycle, and present some numerical illustrations.

2. The Finite Element Discretization

We assume that $\Omega \subseteq \mathbf{R}^2$ is a bounded polygonal domain. As a model problem we consider the differential equation

$$-\sum_{i,j=1}^{2} D_j(a_{ij}D_i u) = f \tag{2.1}$$

on Ω with Dirichlet boundary conditions

$$u = 0 \tag{2.2}$$

on the boundary piece Γ and natural boundary conditions

$$\sum_{i=1}^{2}(\sum_{j=1}^{2} a_{ij}\vec{n}|_j)D_i u = 0 \tag{2.3}$$

on the remaining part $\partial\Omega\backslash\Gamma$ of the boundary of Ω. $\vec{n}$ denotes the outward unit normal vector. The appropriate solution space of this boundary value problem is

$$H(\Omega) = \{u \in W^{1,2}(\Omega) \mid u = 0 \text{ on } \Gamma\} \tag{2.4}$$

where Γ is assumed to be composed of some or all edges of the polygonal domain Ω and the zero boundary conditions have to be understood in the sense of the trace operator. The seminorm

$$|u|^2_{1,2;\Omega} = \sum_{i=1}^{2} \int_\Omega |(D_i u)(x)|^2 dx \tag{2.5}$$

is a norm on $H(\Omega)$. The weak formulation of our boundary value problem is to find a function $u \in H(\Omega)$ satisfying

$$B(u,v) = f^*(v), \quad v \in H(\Omega), \tag{2.6}$$

where f^* is a given bounded linear functional on $H(\Omega)$ and the bilinear form B is defined by

$$B(u,v) = \int_\Omega \sum_{i,j=1}^{2} a_{ij}D_i u D_j v dx\,. \tag{2.7}$$

We assume that the a_{ij} are measurable and bounded functions satisfying

$$a_{ij} = a_{ji}, \qquad i,j = 1,2 \tag{2.8}$$

and

$$\delta \sum_{i=1}^{2} \eta_i^2 \leq \sum_{i,j=1}^{2} a_{ij}(x)\eta_i\eta_j \leq M \sum_{i=1}^{2} \eta_i^2 \tag{2.9}$$

for almost all $x \in \Omega$ and all $\eta \in \mathbf{R}^2$. δ and M are positive constants independent of x and η.

By (2.8) and (2.9) B is a symmetric bounded and coercive bilinear form on $H(\Omega)$.

$$\|u\|^2 = B(u,u) \tag{2.10}$$

defines a norm on $H(\Omega)$, the energy norm. This norm is equivalent to the norm (2.5). Since $H(\Omega)$ is a Hilbert space under this norm, the Riesz representation theorem guarantees that the boundary value problem (2.6) has a unique solution. In this paper we consider only the weak formulation (2.6) and not the classical formulation given above.

By a triangulation $\mathcal{T}$ of the polygonal domain Ω we mean a set of triangles such that the union of these triangles is $\bar{\Omega}$ and such that the intersection of two triangles of $\mathcal{T}$ either consists of a common side or a common vertex of both triangles or is empty. Here we start with an intentionally coarse initial triangulation $\mathcal{T}_1$ of Ω. For every triangle $T \in \mathcal{T}_1$ let $\delta(T)$ and $M(T)$ be positive constants with

$$\delta(T) \sum_{i=1}^{2} \eta_i^2 \leq \sum_{i,j=1}^{2} a_{ij}(x)\eta_i\eta_j \leq M(T) \sum_{i=1}^{2} \eta_i^2 \tag{2.11}$$

for almost all $x \in T$ and all $\eta \in \mathbf{R}^2$. Let

$$\frac{M(T)}{\delta(T)} \leq \sigma, \qquad T \in \mathcal{T}_1. \tag{2.12}$$

This constant σ and not the global ratio M/δ with the constants M and δ from (2.9) enters into our estimates.

To produce a sufficiently accurate solution we refine $\mathcal{T}_1$ several times, giving a family of nested triangulations $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \ldots$. A triangle of $\mathcal{T}_{k+1}$ is either a triangle of the triangulation $\mathcal{T}_k$ to be refined or is generated by subdividing a triangle of $\mathcal{T}_k$ into four congruent subtriangles as shown in Fig. 1a or into two triangles as shown in Fig. 1b. The two triangles in Fig. 1b are obtained by connecting a given vertex of the original triangle with the midpoint of the opposite side. The situation described by Fig. 1a we call a *regular refinement*, and the resulting triangles as well as the triangles of the initial triangulation are *regular triangles*. A refinement as in Fig. 1b is an *irregular refinement* and results in two *irregular triangles*.

The irregular refinement is potentially dangerous because interior angles of the resulting irregular elements might be reduced. Therefore we add the rule that irregular triangles may not be further refined. This rule insures

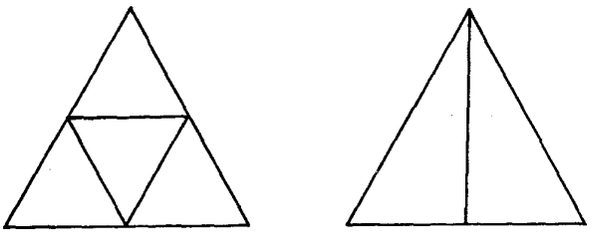

Figure 1 a,b

that every triangle of any triangulation $\mathcal{T}_k$ is geometrically similar to a triangle of the initial triangulation $\mathcal{T}_1$ or an irregular refinement of a triangle of $\mathcal{T}_1$.

The triangles of $\mathcal{T}_1$ are called *level 1 elements*, and the regular and irregular triangles created by the refinement of level $k-1$ triangles are called *level k elements*. The vertices of the mesh $\mathcal{T}_1$ are the *level 1 vertices* or *nodes* and those vertices created by the refinement of a level $k-1$ element are *level k vertices*. It is important to recognize that not all elements in $\mathcal{T}_{k-1}$ need to be refined in creating $\mathcal{T}_k$. In particular, the mesh $\mathcal{T}_k$ may contain unrefined elements from all lower levels, and thus it may be a highly nonuniform mesh.

Algorithms for adaptively generating meshes $\mathcal{T}_k$ satisfying the rules given here are described in [5] and have been implemented in the finite element package PLTMG [2]. The details are unimportant for our considerations but one should note that the levels introduced here do not necessarily reflect the dynamic refinement process.

Corresponding to the triangulations $\mathcal{T}_k$ we have finite element spaces S_k. S_k consists of all functions which are continous on $\bar{\Omega}$ and linear on the triangles $T \in \mathcal{T}_k$ and which vanish on the boundary piece Γ. Clearly S_k is a subspace of S_l for $l \geq k$.

For each space S_k there are two sets of basis functions which play important roles in our discussion: the *nodal basis* $\hat{\psi}_i^{(k)}$, $i = 1, \ldots, n_k$, and the *hierarchical basis* ψ_i, $i = 1, \ldots, n_k$. The nodal basis is the standard basis used in practical computation. The basis function $\hat{\psi}_i^{(k)} \in S_k$ is defined by

$$\hat{\psi}_i^{(k)}(x_l) = \delta_{il} \tag{2.13}$$

where the x_l runs over the vertices of the triangles of $\mathcal{T}_k$ not lying on Γ. The hierarchical basis, on the other hand is defined as follows:

1. The hierarchical basis of S_1 is $\psi_i := \hat{\psi}_i^{(1)}$, $i = 1, \ldots, n_1$.

2. For $k = 2, 3, \ldots$ the hierarchical basis of S_k consists of the hierarchical basis functions

$$\psi_i, \quad i = 1, \ldots, n_{k-1}$$

of S_{k-1} and the nodal basis functions

$$\psi_i := \hat{\psi}_i^{(k)}, \quad i = n_{k-1}+1, \ldots, n_k.$$

In other words the hierarchical basis for S_k is built from that of S_{k-1} by adding the nodal basis functions of S_k associated with the level k nodes not lying on Γ, namely with x_i, $i = n_{k-1}+1, \ldots, n_k$. The hierarchical basis of S_k induces a natural partitioning of the finite element space. Let

$$\mathcal{V}_k = \text{ span } \{\psi_i \mid x_i \text{ is a level } k \text{ vertex}\} \tag{2.14}$$

Then one has the decomposition

$$S_k = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_k \tag{2.15}$$

For convenience we fix a finite element space $S = S_j$. Every function $u \in S$ can uniquely be written as

$$u = \sum_{i=1}^{j} u_i, \qquad u_i \in \mathcal{V}_i \tag{2.16}$$

We now define the interpolation operators

$$J_k : S \to S_k, \quad k = 1, \ldots, j \tag{2.17}$$

by

$$J_k u = \sum_{i=1}^{k} u_i \tag{2.18}$$

$J_k u$ is the uniquely given function of S_k interpolating $u \in S$ at the vertices of the triangles of $\mathcal{T}_k$. The finite element space S_k is the range of J_k and for $k = 2, \ldots, j$ the space $\mathcal{V}_k$ is the range of $J_k - J_{k-1}$.

Our final aim is to solve the linear system

$$\hat{A}\hat{x} = \hat{b} \tag{2.19}$$

corresponding to the boundary value problem (2.6) and the discrete solution space S represented in terms of the nodal basis. The components of the solution vector $\hat{x}$ are the values of the discrete solution at the nodal points. To solve (2.19) we implicitly switch to the hierarchical basis formulation

$$Ax = b \tag{2.20}$$

of the system (2.19). This system is solved by a preconditioned conjugate gradient method. As preconditioners we use relaxation procedures associated with the blocking of the matrix A induced by the splitting (2.15) of the

finite element space $S_j = S$. A complete description of the preconditioning procedures is given in the next section.

Algorithmically the preconditioners are realized *exactly* as V-cycle multigrid methods with Gauß-Seidel smoothers, except that only unknowns corresponding to $\mathcal{V}_k$ (and not to S_k) are smoothed at level k. *Therefore the amount of work per iteration step is proportional to the number of unknowns*, but unlike standard multigrid methods, it is unnecessary to assume geometrically increasing dimensions of the subspaces S_k to achieve this bound. This feature makes the method especially attractive for adaptively refined grids.

In the next two sections we show that the growth of the condition number of the preconditioned matrices is bounded by $O(j^2)$, with j the number of levels. This is slightly suboptimal and leads to an $O(jn)$ algorithm. Practically this represents logarithmic growth in the number of iterations as a function of n, *but contrary to usual multigrid methods this estimate does not require the usual strong elliptic regularity assumptions or the quasiuniformity of the family of triangulations.*

3. The Block Gauß-Seidel Iteration

Let A be arbitrary symmetric positive definite $(n \times n)$-matrix written in block form as

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1j} \\ \vdots & \ddots & \vdots \\ A_{j1} & \cdots & A_{jj} \end{bmatrix} \tag{3.1}$$

with the A_{ii} square matrices, remaining fixed in this section. We use the decomposition

$$A = L + D + L^T \tag{3.2}$$

of A into its block lower triangular part

$$L = \begin{bmatrix} 0 & \cdot & \cdots & 0 \\ A_{21} & 0 & & \cdot \\ \vdots & \ddots & \ddots & \vdots \\ A_{j1} & \cdots & A_{j,j-1} & 0 \end{bmatrix} \tag{3.3}$$

its symmetric, positive definite block diagonal part

$$D = \begin{bmatrix} A_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_{jj} \end{bmatrix} \tag{3.4}$$

and its corresponding upper triangular part L^T.

We consider the solution of

$$Ax = b \tag{3.5}$$

by the symmetric block Gauß-Seidel iteration

$$x^{(i+1/2)} = x^{(i)} + (L + D)^{-T}(b - Ax^{(i)}) \tag{3.6}$$

$$x^{(i+1)} = x^{(i+1/2)} + (L + D)^{-1}(b - Ax^{(i+1/2)}) \tag{3.7}$$

In each step (3.6), (3.7) one must solve $2j$ linear systems with the coefficient matrices $A_{jj}, \ldots, A_{11}$ and $A_{11}, \ldots, A_{jj}$. In many applications, such as ours, the cost of direct solution is prohibitively large. Therefore we replace the block diagonal matrix D in (3.6) and (3.7) by another nonsingular, but not necessarily symmetric, block diagonal matrix $\tilde{D}$ obtaining the modified iteration

$$x^{(i+1/2)} = x^{(i)} + (L + \tilde{D})^{-T}(b - Ax^{(i)}) \tag{3.8}$$

$$x^{(i+1)} = x^{(i+1/2)} + (L + \tilde{D})^{-1}(b - Ax^{(i+1/2)}) \tag{3.9}$$

We begin our analysis with:

Theorem 3.1 *Let the symmetric matrix*

$$X = \tilde{D} + \tilde{D}^T - D \tag{3.10}$$

be positive definite. Then the iteration (3.8), (3.9) *can be written as*

$$x^{(i+1)} = x^{(i)} + B^{-1}(b - Ax^{(i)}) \tag{3.11}$$

with the symmetric and positive definite matrix

$$B = (L + \tilde{D})^T X^{-1}(L + \tilde{D}) \tag{3.12}$$

Proof: Let

$$r^{(i)} = b - Ax^{(i)}, \quad r^{(i+1/2)} = b - Ax^{(i+1/2)}$$

The definition of $x^{(i+1/2)}$ leads to

$$\begin{aligned} r^{(i+1/2)} &= r^{(i)} - A(x^{(i+1/2)} - x^{(i)}) \\ &= r^{(i)} - (L + D + L^T)(x^{(i+1/2)} - x^{(i)}) \\ &= r^{(i)} - (L + \tilde{D})^T(x^{(i+1/2)} - x^{(i)}) + (\tilde{D}^T - D - L)(x^{(i+1/2)} - x^{(i)}) \\ &= (\tilde{D}^T - D - L)(x^{(i+1/2)} - x^{(i)}). \end{aligned}$$

Using the definitions of $x^{(i+1)}$ and $x^{(i+1/2)}$ we obtain

$$\begin{aligned} &(L + \tilde{D})(x^{(i+1)} - x^{(i)}) \\ &= (L + \tilde{D})(x^{(i+1)} - x^{(i+1/2)}) + (L + \tilde{D})(x^{(i+1/2)} - x^{(i)}) \\ &= r^{(i+1/2)} + (L + \tilde{D})(x^{(i+1/2)} - x^{(i)}) \\ &= (\tilde{D}^T - D - L)(x^{(i+1/2)} - x^{(i)}) + (L + \tilde{D})(x^{(i+1/2)} - x^{(i)}) \\ &= (\tilde{D} + \tilde{D}^T - D)(x^{(i+1/2)} - x^{(i)}) \\ &= (\tilde{D} + \tilde{D}^T - D)(L + \tilde{D})^{-T}(b - Ax^{(i)}) \end{aligned}$$

As $\tilde{D} + \tilde{D}^T - D = X$ is positive definite this is the proposition. ∎

The matrix B^{-1} can be regarded as an approximate inverse of A. The efficiency of the positive definite matrix B as a preconditioner for A is largely described by the generalized condition number

$$\kappa = \frac{\mu_2}{\mu_1} \tag{3.13}$$

where μ_1 and μ_2 are defined by

$$\frac{1}{\mu_1} = \max_{x \neq 0} \frac{(x, Bx)}{(x, Ax)} \tag{3.14}$$

and

$$\mu_2 = \max_{x \neq 0} \frac{(x, Ax)}{(x, Bx)} \tag{3.15}$$

The brackets denote the Euclidean inner product.

We consider three different choices for $\tilde{D}$. The first one is

$$\text{Case 1:} \quad \tilde{D} = D \tag{3.16}$$

which corresponds to the original block Gauß-Seidel method (3.6), (3.7). The other two choices are described in terms of the decomposition

$$D = l + d + l^T \tag{3.17}$$

of the block diagonal matrix (3.4). Here

$$d = \begin{bmatrix} A_{11} & \cdot & \cdots & 0 \\ \cdot & \text{diag } (A_{22}) & & \cdot \\ \vdots & & \ddots & \vdots \\ 0 & \cdot & \cdots & \text{diag } (A_{jj}) \end{bmatrix} \tag{3.18}$$

is a pointwise diagonal matrix except for the $(1,1)$–block and l is a strictly lower–triangular matrix. Now we can define

$$\text{Case 2:} \qquad \tilde{D} = l + d \tag{3.19}$$

$$\text{Case 3:} \qquad \tilde{D} = (l + d)^T d^{-1} (l + d) \tag{3.20}$$

Case 2 corresponds to using one Gauß-Seidel step for all diagonal blocks except for the first one where the corresponding system is solved exactly. In Case 3 each of the diagonal blocks with exception of the first one is treated by one symmetric Gauß-Seidel step; compare Theorem 3.1. In all three cases the matrix (3.10) is positive definite: We have for

$$\text{Case 1:} \quad X = D \tag{3.21}$$

$$\text{Case 2:} \quad X = d \tag{3.22}$$

$$\text{Case 3:} \quad X = D + 2l^T d^{-1} l \tag{3.23}$$

Therefore Theorem 3.1 applies.

For all three cases the constant μ_2 in (3.15) is explicitly known:

Theorem 3.2 *For all three cases defined above*

$$\mu_2 = \max_{x \neq 0} \frac{(x, Ax)}{(x, Bx)} = 1 \tag{3.24}$$

Proof: The proof that μ_2 is bounded by 1 requires only the hypothesis of Theorem 3.1. Let $Y = \tilde{D} - D$. Because $\tilde{D} = X - Y^T$ we have

$$B = (L + X - Y^T)^T X^{-1}(L + X - Y^T)$$

Using the identities $X^T = X$ and $D = X - Y - Y^T$ we obtain

$$\begin{aligned} B &= ((L - Y^T)^T + X)X^{-1}((L - Y^T) + X) \\ &= L + (X - Y - Y^T) + L^T + (L - Y^T)^T X^{-1}(L - Y^T) \\ &= A + F \end{aligned}$$

where F is the symmetric, positive semidefinite matrix

$$F = (L + D - \tilde{D}^T)^T X^{-1}(L + D - \tilde{D}^T)$$

An immediate consequence of this representation is

$$\frac{(x, Ax)}{(x, Bx)} = \frac{(x, Ax)}{(x, Ax) + (x, Fx)} \leq \frac{(x, Ax)}{(x, Ax)} = 1$$

for all vectors $x \neq 0$ or $\mu_2 \leq 1$.

Therefore the proof is completed by showing that for each choice of $\tilde{D}$ there exists a vector $x^* \neq 0$ satisfying

$$Fx^* = 0$$

or equivalently

$$(L + D - \tilde{D}^T)x^* = 0$$

We have for

$$\begin{aligned} &\text{Case 1:} \quad L + D - \tilde{D}^T = L \\ &\text{Case 2:} \quad L + D - \tilde{D}^T = L + l \\ &\text{Case 3:} \quad L + D - \tilde{D}^T = L - l^T d^{-1} l \end{aligned}$$

The last column $x^* = (0, \ldots, 0, 1)^T$ of the identity matrix satisfies

$$Lx^* = 0, \quad lx^* = 0$$

and therefore in all three cases $Fx^* = 0$. ∎

By Theorem 3.2 for all three choices of $\tilde{D}$, the generalized condition number (3.13) is given by

$$\kappa = \max_{x \neq 0} \frac{(x, Bx)}{(x, Ax)} \tag{3.25}$$

In the following we use the Euclidean norm

$$|x| = (x, x)^{1/2} \tag{3.26}$$

its associated matrix norm

$$|M| = \max_{x \neq 0} \frac{|Mx|}{|x|} \tag{3.27}$$

which is the spectral norm, the energy norm

$$\|x\| = (x, Ax)^{1/2} \tag{3.28}$$

and its associated matrix norm

$$\|M\| = \max_{x \neq 0} \frac{\|Mx\|}{\|x\|} \tag{3.29}$$

which is given by

$$\|M\| = |A^{1/2} M A^{-1/2}| \tag{3.30}$$

The rate of convergence of a stationary iterative method

$$x^{(i+1)} = x^{(i)} + M^{-1}(b - Ax^{(i)}) \tag{3.31}$$

with respect to the energy norm is

$$\|E - M^{-1}A\| = |E - A^{1/2} M^{-1} A^{1/2}| \tag{3.32}$$

E denotes the identiy matrix.

Theorem 3.3 *For all three choices of the block diagonal matrix $\tilde{D}$ the modified symmetric block Gauß-Seidel iteration* (3.8), (3.9) *has the convergence rate*

$$1 - \frac{1}{\kappa}$$

Proof: By definition μ_1 is the smallest and μ_2 the largest eigenvalue of the matrix $B^{-1/2} A B^{-1/2}$ which is similar to the symmetric matrix $A^{1/2} B^{-1} A^{1/2}$. Using Theorem 3.2, (3.14) and (3.25), we obtain

$$|E - A^{1/2} B^{-1} A^{1/2}| = 1 - \frac{1}{\kappa}$$

■

We note that the convergence rates of the modified forward block Gauß-Seidel iteration

$$x^{(i+1)} = x^{(i)} + (L + \tilde{D})^{-1}(b - Ax^{(i)}) \tag{3.33}$$

and the corresponding backward iteration

$$x^{(i+1)} = x^{(i)} + (L + \tilde{D})^{-T}(b - Ax^{(i)}) \tag{3.34}$$

can be expressed in terms of the generalized condition number (3.13). This is shown following [13].

Theorem 3.4 *For all three choices of $\tilde{D}$ the rate of convergence of the iterations* (3.33), (3.34) *with respect to the energy norm is*

$$(1 - \frac{1}{\kappa})^{1/2}$$

Thus two steps of the iterations (3.33) or (3.34) have approximately the same effect as one step of the symmetrized iteration (3.8), (3.9).

Proof: The convergence rate of the first method is

$$|E - A^{1/2}(L + \tilde{D})^{-1}A^{1/2}|$$

and the convergence rate of the second iteration

$$|E - A^{1/2}(L + \tilde{D})^{-T}A^{1/2}|$$

As the spectral norm of a matrix and its transpose coincide both norms are equal and given by the squareroot of the largest eigenvalue of the matrix

$$[E - A^{1/2}(L + \tilde{D})^{-T}A^{1/2}][E - A^{1/2}(L + \tilde{D})^{-1}A^{1/2}] = E - A^{1/2}B^{-1}A^{1/2}$$

As we have stated in the proof of Theorem 3.3 this eigenvalue is

$$1 - \frac{1}{\kappa}$$

■

We are left with the problem of bounding the constant (3.14). This cannot be done without using further properties of the matrix A. In the remainder of this section we develop some estimates that can be applied to our finite element equations.

We define the following constants:

$$\alpha_0^2 = \max_{x \neq 0} \frac{(x, (L + D)^T D^{-1}(L + D)x)}{(x, Ax)} \tag{3.35}$$

$$\alpha_1^2 = \max_{x \neq 0} \frac{(x, Dx)}{(x, Ax)} \tag{3.36}$$

$$\beta^2 = \max_{x \neq 0} \frac{(x, l^T d^{-1} l x)}{(x, dx)} \tag{3.37}$$

$$\gamma_1^2 = \max_{x \neq 0} \frac{(x, dx)}{(x, Dx)} \tag{3.38}$$

$$\gamma_2^2 = \max_{x \neq 0} \frac{(x, Dx)}{(x, dx)} \tag{3.39}$$

The constants α_0 and α_1 are invariant under any block diagonal scaling of A and β, γ_1 and γ_2 under any diagonal scaling of A.

Lemma 3.5 *For the block Gauß-Seidel method (3.6), (3.7) (Case 1),*

$$\frac{1}{\mu_1} = \alpha_0^2 \tag{3.40}$$

Proof: The matrix (3.12) of Theorem 3.1 is given by

$$B = (L + D)^T D^{-1}(L + D)$$

■

By Theorem 3.2 and Lemma 3.5

$$\kappa_1 = \alpha_0^2 \tag{3.41}$$

is the generalized condition number corresponding to Case 1.

$1/\alpha_1^2$ is the smallest eigenvalue of the symmetric matrix

$$D^{-1/2} A D^{-1/2} \tag{3.42}$$

This corresponds to the original matrix A preconditioned by its block diagonal part D. Note that the maximum eigenvalue

$$\alpha_2^2 = \max_{x \neq 0} \frac{(x, Ax)}{(x, Dx)} \tag{3.43}$$

of the matrix (3.42) does not enter into our estimates.

$1/\gamma_1^2$ is the minimum and γ_2^2 the maximum eigenvalue of the scaled block diagonal part

$$d^{-1/2} D d^{-1/2} \tag{3.44}$$

of the matrix A.

Using the spectral norm the constants can be written as follows:

$$\alpha_0 = |D^{-1/2}(L + D)A^{-1/2}| \tag{3.45}$$
$$\alpha_1 = |D^{1/2}A^{-1/2}| \tag{3.46}$$
$$\beta = |d^{-1/2} l d^{-1/2}| \tag{3.47}$$
$$\gamma_1 = |d^{1/2}D^{-1/2}| \tag{3.48}$$
$$\gamma_2 = |D^{1/2}d^{-1/2}| \tag{3.49}$$

As the spectral norm of a matrix and of its transpose coincide in addition we have

$$\beta = |d^{-1/2} l^T d^{-1/2}| \tag{3.50}$$
$$\gamma_1 = |D^{-1/2}d^{1/2}| \tag{3.51}$$
$$\gamma_2 = |d^{-1/2}D^{1/2}| \tag{3.52}$$

Lemma 3.6 *For Case 2 the constant (3.14) satisfies*

$$\frac{1}{\mu_1} \leq (\gamma_2 \alpha_0 + \beta \gamma_1 \alpha_1)^2 \tag{3.53}$$

Proof: The matrix B is given by

$$B = (L + l + d)^T d^{-1}(L + l + d)$$

and therefore we have

$$\frac{1}{\mu_1} = |d^{-1/2}(L + l + d)A^{-1/2}|^2$$

Using (3.45), (3.46), (3.48), (3.50) and (3.52) we obtain

$$\begin{aligned}
&|d^{-1/2}(L + l + d)A^{-1/2}| \\
&= |d^{-1/2}(L + D)A^{-1/2} - d^{-1/2}l^T A^{-1/2}| \\
&\leq |d^{-1/2}D^{1/2}|\,|D^{-1/2}(L + D)A^{-1/2}| + \\
&\quad + |d^{-1/2}l^T d^{-1/2}|\,|d^{1/2}D^{-1/2}|\,|D^{1/2}A^{-1/2}| \\
&\leq \gamma_2\alpha_0 + \beta\gamma_1\alpha_1
\end{aligned}$$

■

Lemma 3.7 *For Case 3 the constant (3.14) satisfies*

$$\frac{1}{\mu_1} \leq (\alpha_0 + \frac{1}{\sqrt{2}}\beta\gamma_1\alpha_1)^2 \tag{3.54}$$

Proof: For this calculation only, let

$$Z = l^T d^{-1} l$$

Then the matrix B has the representation

$$B = (L + D + Z)^T (D + 2Z)^{-1}(L + D + Z),$$

and therefore we have

$$\begin{aligned}
(\frac{1}{\mu_1})^{1/2} &= |(D + 2Z)^{-1/2}(L + D + Z)A^{-1/2}| \\
&\leq |(D + 2Z)^{-1/2}D^{1/2}|\,|D^{-1/2}(L + D)A^{-1/2}| + \\
&\quad + |(D + 2Z)^{-1/2}ZD^{-1/2}|\,|D^{1/2}A^{-1/2}|
\end{aligned}$$

Using the positive semidefinite symmetric matrix

$$Q = D^{-1/2}ZD^{-1/2}$$

one obtains

$$\begin{aligned}
[(D + 2Z)^{-1/2}D^{1/2}]^T[(D + 2Z)^{-1/2}D^{1/2}] &= D^{1/2}(D + 2Z)^{-1}D^{1/2} \\
&= (E + 2Q)^{-1}
\end{aligned}$$

and

$$\begin{aligned}&[(D+2Z)^{-1/2}ZD^{-1/2}]^T[(D+2Z)^{-1}ZD^{-1/2}]\\&= D^{-1/2}Z(D+2Z)^{-1}ZD^{-1/2}\\&= Q(E+2Q)^{-1}Q\end{aligned}$$

If Λ denotes the set of eigenvalues of Q,

$$\frac{1}{1+2\lambda}, \quad \lambda \in \Lambda,$$

is the set of eigenvalues of $(E+2Q)^{-1}$. As Q is a positive semidefinite and singular matrix this means

$$|(D+2Z)^{-1/2}D^{1/2}| = 1.$$

The set of eigenvalues of the matrix $Q(E+2Q)^{-1}Q$ is

$$\frac{\lambda^2}{1+2\lambda}, \quad \lambda \in \Lambda.$$

Since

$$\frac{\lambda^2}{1+2\lambda} \leq \frac{\lambda}{2}, \quad \lambda \geq 0,$$

we get

$$|(D+2Z)^{-1/2}ZD^{-1/2}|^2 \leq \frac{1}{2}|D^{-1/2}ZD^{-1/2}|$$

Inserted above we obtain the estimate

$$(\frac{1}{\mu_1})^{1/2} \leq |D^{-1/2}(L+D)A^{-1/2}| + \frac{1}{\sqrt{2}}|D^{-1/2}ZD^{-1/2}|^{1/2}|D^{1/2}A^{-1/2}|$$

Now, using (3.51), (3.50), (3.47), (3.48)

$$\begin{aligned}|D^{-1/2}ZD^{-1/2}| &= |D^{-1/2}l^Td^{-1}lD^{-1/2}|\\&\leq |D^{-1/2}d^{1/2}|\,|d^{-1/2}l^Td^{-1/2}|\,|d^{-1/2}ld^{-1/2}|\,|d^{1/2}D^{-1/2}|\\&\leq \gamma_1\beta\beta\gamma_1\end{aligned}$$

Combining this with (3.45) and (3.46) completes the proof. ■

Theorem 3.2 and Lemma 3.6 and 3.7 lead immediately to the following estimates for the generalized condition number (3.13):

$$\text{Case 2:} \quad \kappa_2 \leq (\gamma_2\alpha_0 + \beta\gamma_1\alpha_1)^2 \tag{3.55}$$

$$\text{Case 3:} \quad \kappa_3 \leq (\alpha_0 + \frac{1}{\sqrt{2}}\beta\gamma_1\alpha_1)^2 \tag{3.56}$$

We conjecture

$$\kappa_1 \leq \kappa_3 \leq \kappa_2 \tag{3.57}$$

Because of $\gamma_2 \geq 1$ our estimates reflect this supposed behaviour.

Our next task is to obtain simple bounds for the constants α_0 and α_1. For this purpose let

$$I_k = \begin{bmatrix} E_{11} & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & E_{kk} & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} \tag{3.58}$$

be the block–diagonal matrix the first k block–diagonal entries of which are the identity matrices E_{ii} and all other blocks are zero. In our application the matrices I_k, $k = 1, \ldots, j$, can be interpreted as the finite element interpolation operators of Section 2.

Lemma 3.8 *For all n–dimensional vectors x,*

$$|D^{-1/2}(L+D)A^{-1/2}x|^2 \leq \sum_{k=1}^{j} |A^{1/2}I_kA^{-1/2}x|^2 \tag{3.59}$$

and

$$|D^{1/2}A^{-1/2}x|^2 \leq 4\sum_{k=1}^{j} |A^{1/2}I_kA^{-1/2}x|^2 \tag{3.60}$$

Proof: The first proposition holds if and only if one has for any fixed vector x

$$|x|^2 \leq \sum_{k=1}^{j} |A^{1/2}I_k(L+D)^{-1}D^{1/2}x|^2 \tag{3.61}$$

Let $y = (L+D)^{-1}D^{1/2}x$ and define the matrices

$$P_1 = I_1, \quad P_k = I_k - I_{k-1}, \quad k = 2, \ldots, j\ .$$

Because of the block structure of the matrices L and D we have

$$D^{1/2}I_1y = P_1x$$

Since

$$I_1DI_1 = I_1AI_1$$

we obtain

$$|P_1x| = |A^{1/2}I_1y| \tag{3.62}$$

For estimating $|P_kx|$ for $k = 2, \ldots, j$ we introduce the decomposition

$$I_ky = v + w,$$

where

$$\begin{aligned} v &= I_k(L+D)^{-1}D^{1/2}P_k x \\ w &= I_k(L+D)^{-1}D^{1/2}I_{k-1}x \end{aligned}$$

Since $I_k(L+D)^{-1}D^{1/2}P_k$ is zero except for its $k-$th diagonal block, we obtain from a simple computation

$$v = P_k v = P_k D^{-1/2} P_k x, \tag{3.63}$$

and because

$$P_k A P_k = P_k D P_k,$$

as a first consequence

$$|P_k x| = |A^{1/2} v|. \tag{3.64}$$

Now, since

$$P_k A I_k = P_k(L+D),$$

we get

$$\begin{aligned} P_k A w &= P_k A I_k (L+D)^{-1} D^{1/2} I_{k-1} x \\ &= P_k D^{1/2} I_{k-1} x \\ &= 0 \end{aligned}$$

Using (3.63)

$$(v, Aw) = (P_k v, Aw) = (v, P_k A w) = 0$$

and hence

$$|A^{1/2} I_k y|^2 = |A^{1/2}(v+w)|^2 = |A^{1/2} v|^2 + |A^{1/2} w|^2.$$

Thus from (3.64)

$$|P_k x| \le |A^{1/2} I_k y|. \tag{3.65}$$

Combining (3.62) and (3.65) we finally obtain

$$|x|^2 = \sum_{k=1}^{j} |P_k x|^2 \le \sum_{k=1}^{j} |A^{1/2} I_k y|^2$$

which is (3.61) and completes the proof of (3.59). The proof of the second estimate is less tricky. Using the triangle and Cauchy-Schwarz inequalities for all vectors x we obtain

$$\begin{aligned} |D^{1/2} x|^2 &= (x, Dx) \\ &= \sum_{k=1}^{j} (P_k x, A P_k x) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{j} |A^{1/2} P_k x|^2 \\
&= |A^{1/2} I_1 x|^2 + \sum_{k=2}^{j} |A^{1/2}(I_k - I_{k-1})x|^2 \\
&\leq |A^{1/2} I_1 x|^2 + \sum_{k=2}^{j} (|A^{1/2} I_k x| + |A^{1/2} I_{k-1} x|)^2 \\
&\leq 4 \sum_{k=1}^{j} |A^{1/2} I_k x|^2
\end{aligned}
$$

which implies (3.60). ∎

An immediate consequence of Lemma 3.8 and the representations (3.45) and (3.46) are the bounds

$$\alpha_0^2 \leq \sum_{k=1}^{j} \|I_k\|^2 \tag{3.66}$$

$$\alpha_1^2 \leq 4 \sum_{k=1}^{j} \|I_k\|^2 \tag{3.67}$$

for the constants (3.35) and (3.36) in terms of the energy norms of the matrices I_k. We remark that since

$$
\begin{aligned}
(x, (L+\tilde{D})^T X^{-1} (L+\tilde{D}) x) &\geq |I_1 X^{-1/2} (L+\tilde{D}) x|^2 \\
&= |I_1 D^{1/2} x|^2 \\
&= (I_1 x, A I_1 x),
\end{aligned}
$$

for all three choices of $\tilde{D}$ we obtain the lower bounds

$$\kappa \geq \|I_1\|^2 \tag{3.68}$$

for the condition numbers $\kappa = \kappa_1, \kappa_2, \kappa_3$.

In the next lemma we derive simple estimates for β and γ_2 of (3.47) and (3.49), respectively.

Lemma 3.9 *Suppose the A_{kk}, $k = 2, \ldots, j$, have at most $m+1$ nonzeros per row. Then*

$$\beta = |d^{-1/2} l d^{-1/2}| \leq m \tag{3.69}$$

and

$$\gamma_2 = |D^{1/2} d^{-1/2}| \leq \sqrt{m+1} \tag{3.70}$$

Proof: By the Cauchy-Schwarz inequality the entries q_{ij} of any symmetric positive definite matrix satisfy

$$|q_{ij}| \leq \sqrt{q_{ii} q_{jj}}$$

Therefore all entries of the scaled matrix

$$d^{-1/2} D d^{-1/2}$$

are bounded by 1. Remembering that the $(1,1)$–block of this matrix is an identity matrix, simple Gershgorin estimates show that the eigenvalues of the matrices

$$(d^{-1/2} l d^{-1/2})^T (d^{-1/2} l d^{-1/2})$$

and

$$(D^{1/2} d^{-1/2})^T (D^{1/2} d^{-1/2})$$

are bounded by m^2 and $m+1$, respectively. ∎

Finally, we note that the estimates of this section (with a small and obvious modification of Lemma 3.9) remain unchanged if one replaces the matrix d from (3.18) by the diagonal of D.

In the next section we derive estimates for

$$\|I_k\|, \quad k = 1, \ldots, j, \tag{3.71}$$

and for the constant γ_1 of (3.48). For these estimates we must return to the finite element discretization and use the properties that it implies for A.

4. Condition Number Estimates

In this section let A be the discretization matrix corresponding to the model problem (2.20) of Section 2. Assume that the blocking (3.1) of A is induced by the partition (2.15) of the finite element space into different refinement levels. Within this framework, we derive bounds for the growth of the condition numbers (3.13) as a function of the number of levels j.

First we observe that the matrices (3.58) correspond to the interpolation operator (2.17): If $u \in S$ is the function represented by the hierarchical basis coefficient vector x, the function $J_k u$ is represented by $I_k x$. For finding bounds for the norms of the interpolation operators J_k we utilize one of the key results of [14]. This result is restricted to two space dimensions. One obtains more rapid growth rates for three-dimensional problems; compare the remarks in the introduction of [14] concerning this topic.

Lemma 4.1 *There exists a constant K_1 such that*

$$|J_k u|^2_{1,2;T} \leq K_1(j-k+1)\ |u|^2_{1,2;T} \tag{4.1}$$

for $k = 1, \ldots, j$, all triangles T of the initial triangulation $\mathcal{T}_1$ and all functions $u \in S$. K_1 depends only on a lower bound for the interior angles of the triangles in $\mathcal{T}_1$ but not on k or j.

Proof: The proposition is an immediate consequence of Lemma 2.2 of [14]. ∎

Remark: By Lemma 2.3 of [14] one gets the estimate

$$\|J_k u\|^2_{0,2;T} \leq K_1^*(j-k+1)\ \{\|u\|^2_{0,2;T} + (2^{-k}H)^2\ |u|^2_{1,2;T}\} \tag{4.2}$$

for the L_2–norm of $J_k u$. H denotes the diameter of the triangle $T \in \mathcal{T}_1$.

Lemma 4.2 *For $k = 1, \ldots, j$ and all functions $u \in S$ one has*

$$\|J_k u\|^2 \leq K_1 \sigma(j-k+1)\|u\|^2 \tag{4.3}$$

where K_1 is the constant from Lemma 4.1 and σ the constant (2.12) describing the local ellipticity of the boundary value problem.

Proof: By (2.11), (2.12) and Lemma 4.1 one has

$$\begin{aligned}\|J_k u\|^2 &\leq \sum_{T\in\mathcal{T}_1} M(T)|J_k u|^2_{1,2;T} \\ &\leq \sum_{T\in\mathcal{T}_1} M(T)K_1(j-k+1)|u|^2_{1,2;T} \\ &\leq K_1\sigma(j-k+1)\sum_{T\in\mathcal{T}_1}\delta(T)|u|^2_{1,2;T} \\ &\leq K_1\sigma(j-k+1)\|u\|^2\end{aligned}$$

■

As a consequence of Lemma 4.2 and the estimates (3.66) and (3.67) one obtains the bounds

$$\alpha_0^2 \leq \frac{1}{2}K_1\sigma j(j+1) \tag{4.4}$$

$$\alpha_1^2 \leq 2K_1\sigma j(j+1) \tag{4.5}$$

for the constants (3.35) and(3.36).

We remark that estimates for α_0 and α_1 can also be derived from strengthed Cauchy-Schwarz inequalities. The inequalities appropriate here are

$$|B(v_k,w_k)| \leq \epsilon_k\|v_k\|\,\|w_k\|,\quad k=1,\ldots,j-1,$$
$$0\leq\epsilon_k\leq 1-\frac{c}{j-k}<1,$$

which hold for all $v_k\in\mathcal{V}_1\oplus\mathcal{V}_2\oplus\cdots\oplus\mathcal{V}_k$, $w_k\in\mathcal{V}_{k+1}\oplus\cdots\oplus\mathcal{V}_j$. These can be established using techniques similar to those in [3,14]. The estimates

$$\|J_k\|^2\leq\frac{1}{1-\epsilon_k}$$

are then immediate from the decomposition

$$u=v_k+w_k,\quad J_k u=v_k,$$

and establish Lemma 4.2.

As second step we have to derive a bound for the constant γ_1 given by (3.38).

Lemma 4.3 *Let $\psi_i=\hat{\psi}_i^{(k)}$, $i=n_{k-1}+1,\ldots,n_k$, be the nodal basis of $\mathcal{V}_k$ with the associated level k nodes x_i, $i=n_{k-1}+1,\ldots,n_k$, as introduced in Section 2. Then there exists a constant K_2 such that*

$$\sum_{\substack{i=n_{k-1}+1\\ x_i\in T}}^{n_k}|\psi_i|^2_{1,2;T}|v(x_i)|^2\leq K_2|v|^2_{1,2;T} \tag{4.6}$$

for all functions $v\in\mathcal{V}_k$ and all triangles $T\in\mathcal{T}_1$. This constant depends only on a lower bound for the interior angles of the triangles in the initial triangulation $\mathcal{T}_1$ but not on the choice of $k=2,\ldots,j$.

Proof: First we observe that it is sufficient to prove the proposition for the level $k-1$ triangles T of $\mathcal{T}_{k-1}$ containing a level k node instead for the triangles of $\mathcal{T}_1$. Here one has to distinguish the two types of triangles shown in Fig. 2a,b. In the first case T contains three level k nodes and in the

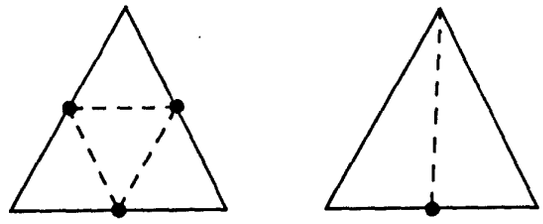

Figure 2 a,b

second case one. The functions $v \in \mathcal{V}_k$ are piecewise linear on the dashed triangles in Fig. 2a,b. They vanish at the vertices of T. Therefore

$$\left(\sum_{\substack{i=n_{k-1}+1 \\ x_i \in T}}^{n_k} |\psi_i|^2_{1,2;T} |v(x_i)|^2 \right)^{1/2}$$

and

$$|v|_{1,2;T}$$

are norms on the space of all functions $v \in \mathcal{V}_k$ restricted to T. These norms are uniformly equivalent as one can see by transforming T to a fixed reference triangle. In the second case one has

$$\sum_{\substack{i=n_{k-1}+1 \\ x_i \in T}}^{n_k} |\psi_i|^2_{1,2;T} |v(x_i)|^2 = |v|^2_{1,2;T} \ .$$

■

Remark: Similarly one can prove the estimate

$$\sum_{\substack{i=n_{k-1}+1 \\ x_i \in T}}^{n_k} \|\psi_i\|^2_{0,2;T} |v(x_i)|^2 \leq K_2^* \|v\|^2_{0,2;T} \tag{4.7}$$

with the seminorms (2.5) replaced by the L_2-norm.

Lemma 4.4 *For $k = 2, \ldots, j$ and all functions $v \in \mathcal{V}_k$ one has*

$$\sum_{i=n_{k-1}+1}^{n_k} \|\psi_i\|^2 |v(x_i)|^2 \leq K_2 \sigma \|v\|^2 \tag{4.8}$$

where K_2 is the constant from Lemma 4.3 *and σ the constant* (2.12).

Proof: By (2.11), (2.12) and Lemma 4.1 one has

$$\begin{aligned}
&\sum_{x_i} \|\psi_i\|^2 |v(x_i)|^2 \\
\leq\ &\sum_{T\in\mathcal{T}_1} M(T) \sum_{x_i\in T} |\psi_i|^2_{1,2;T} |v(x_i)|^2 \\
\leq\ &\sum_{T\in\mathcal{T}_1} M(T) K_2 |v|^2_{1,2;T} \\
\leq\ &K_2\sigma \sum_{T\in\mathcal{T}_1} \delta(T) |v|^2_{1,2;T} \\
\leq\ &K_2\sigma \|v\|^2
\end{aligned}$$

■

Because the diagonal blocks of D can be treated separately and the (1,1)-block of $d^{-1/2}Dd^{-1/2}$ is an identity matrix, Lemma 4.3 gives the bound

$$\gamma_1^2 \leq K_2\sigma \tag{4.9}$$

for the constant (3.38). Here we see why the (1,1)–block of the matrix (3.18) cannot be replaced by its diagonal: Otherwise the lowest eigenvalue of the scaled level 1 discretization matrix would enter into our estimates, and γ_1 would no longer be independent of the global structure of the problem.

By construction for $k = 2,\ldots,j$ every level k node has at most four level k neighbours. Therefore by Lemma 3.9 one gets the bounds

$$\beta \leq 4 \tag{4.10}$$

$$\gamma_2 \leq \sqrt{5} \tag{4.11}$$

for the constants (3.37) and (3.39). For many cases these estimates will not be very sharp. If, for example, the scaled matrix (3.44) is weakly diagonally dominant one has $\beta \leq 1$ and $\gamma_2 \leq \sqrt{2}$.

We can now state our final theorem on the condition numbers κ_1, κ_2 and κ_3 corresponding to the three iterative methods:

Theorem 4.5 *There exist positive constants C_1, C_2 and C_3, which depend only on the lower bound for the interior angles of the triangles in the initial triangulation, such that*

$$\kappa_1 \leq C_1\sigma j^2 \tag{4.12}$$

$$\kappa_2 \leq C_2\sigma^2 j^2 \tag{4.13}$$

$$\kappa_3 \leq C_3\sigma^2 j^2 \tag{4.14}$$

Proof: By (3.41) and (4.4) one obtains

$$\kappa_1 \leq \frac{1}{2} K_1 \sigma j(j+1).$$

(3.55), (4.11), (4.4), (4.10), (4.9) and (4.5) lead to

$$\kappa_2 \leq \left(\sqrt{\frac{5}{2}} + 4\sqrt{2}\sqrt{K_2 \sigma} \right)^2 K_1 \sigma j(j+1),$$

and by (3.56), (4.4), (4.10), (4.9) and (4.5) one obtains

$$\kappa_3 \leq \left(\frac{1}{\sqrt{2}} + 4\sqrt{K_2 \sigma} \right)^2 K_1 \sigma j(j+1)$$

Using $\sigma \geq 1$ the proposition follows. ∎

Through examples given in Section 5 of [14] one can see that $\|I_1\|^2 = \|J_1\|^2$ usually grows like $O(j)$. As a consequence of this fact and (3.68), the condition numbers κ_1, κ_2 and κ_3 cannot be excepted to be bounded uniformly in the number j of levels.

We remark that for the spectral condition number κ_0 of the scaled hierarchical basis discretization matrix

$$d^{-1/2} A d^{-1/2} \tag{4.15}$$

our estimate of the same type as in Theorem 4.5 holds: Using the constant (3.43) one gets

$$\kappa_0 \leq (\gamma_1 \alpha_1 \gamma_2 \alpha_2)^2, \tag{4.16}$$

and using a bound like

$$\alpha_2^2 \leq K_3 \sigma \ , \tag{4.17}$$

which has been derived in [14,15], we obtain $\kappa_0 \leq C\sigma^3 j^2$ what can be improved to

$$\kappa_0 \leq C_0 \sigma^2 j^2 \ . \tag{4.18}$$

We now turn to work estimates for our procedures. By Theorem 3.3 the modified symmetric block Gauß-Seidel iterations here denoted as Case 2 and Case 3 have the convergence rate

$$1 - \frac{1}{\kappa} \tag{4.19}$$

with respect to the energy norm. After

$$k \leq \kappa |\log \epsilon| + 1 \tag{4.20}$$

iteration steps the energy norm of the initial error is reduced by the factor $0 < \epsilon < 1$. The estimates for κ of Theorem 4.5 lead to the upper bounds

$$1 - \frac{1}{C\sigma^2 j^2} \tag{4.21}$$

for the convergence rate and to the upper bounds

$$C\sigma^2 j^2 |\log \epsilon| + 1 \tag{4.22}$$

for the number of iteration steps necessary to reach the given accuracy. Using Theorem 3.4 a similar result is found for the forward and backward modified block Gauß-Seidel methods. We recommend our Gauß-Seidel iterations as preconditioners for the conjugate gradient method. k conjugate gradient steps reduce the energy norm of the initial error by at least the factor

$$\frac{2q^k}{1+q^{2k}} \tag{4.23}$$

where

$$q = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \ ; \tag{4.24}$$

see [1,10]. After

$$k \leq \frac{1}{2}\sqrt{\kappa} |\log(\frac{\epsilon}{2})| + 1 \tag{4.25}$$

steps the energy norm of the initial error has been reduced by at least the factor ϵ. In our application not more than

$$k \leq \frac{1}{2}\sqrt{C}\sigma j |\log(\frac{\epsilon}{2})| + 1 \tag{4.26}$$

steps are needed to reach this accuracy. As each iteration step requires $O(n)$ operations a total of $O(jn|\log \epsilon|)$ operations are required to reduce the energy norm of the initial error by the factor ϵ.

5. Implementation and Numerical Results

We begin this section with a brief discussion of the hierarchical basis method and its relationship to multigrid methods. For convenience we will limit the presentation to the case of $j = 2$ levels. As we noted in Section 2 we wish to solve the linear system (2.19), corresponding to the nodal basis, which we now write in block form as

$$\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} \tag{5.1}$$

The system (5.1) is related to the hierarchical basis system

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{5.2}$$

via a nonsingular matrix S which transforms the representation of a finite element function with respect to the hierarchical basis into its representation with respect to a nodal basis. S has the block structure

$$S = \begin{bmatrix} E & 0 \\ R & E \end{bmatrix} \tag{5.3}$$

The off-diagonal entries of S are zero except for at most two nonzeroes in

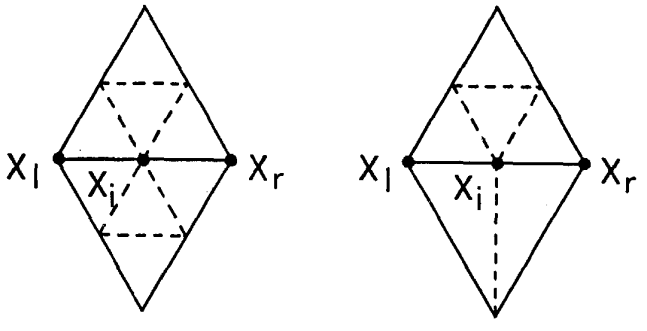

Figure 3 a,b

each row of R, whose values are $\frac{1}{2}$. In particular, if vertex x_i was created during the refinement process as the midpoint of the edge with endpoints x_l and x_r then $S_{il} = S_{ir} = \frac{1}{2}$; see Fig. 3a,b. Using (5.3) we relate the matrices A and $\hat{A}$ by

$$A = S^T \hat{A} S, \tag{5.4}$$

the right hand sides b and $\hat{b}$ by

$$b = S^T \hat{b} \tag{5.5}$$

and the solutions $\hat{x}$ and x by

$$\hat{x} = Sx. \tag{5.6}$$

Blockwise this gives

$$\begin{aligned} A_{11} &= \hat{A}_{11} + R^T \hat{A}_{21} + \hat{A}_{12} R + R^T \hat{A}_{22} R \\ A_{12} &= \hat{A}_{12} + R^T \hat{A}_{22} \\ A_{21} &= \hat{A}_{21} + \hat{A}_{22} R \\ A_{22} &= \hat{A}_{22} \end{aligned}$$

and

$$\begin{aligned} b_1 &= \hat{b}_1 + R^T \hat{b}_2 \\ b_2 &= \hat{b}_2 \\ \hat{x}_1 &= x_1 \\ \hat{x}_2 &= R x_1 + x_2 \end{aligned}$$

A single iteration cycle of the hierarchical basis method, Case 3, giving

$$x = B^{-1} b, \quad \hat{x} = S B^{-1} S^T \hat{b}, \tag{5.7}$$

B as in (3.12), is reformulated in the current notation as:

Procedure HB/MG

1. *smooth* $A_{22} y_2 = b_2$
   using symmetric Gauß-Seidel; denote approximate solution by y_2
2. *form* $r_1 = b_1 - A_{12} y_2$ *as*
   (a) $\hat{r}_1 = \hat{b}_1 - \hat{A}_{12} y_2$
   (b) $\hat{r}_2 = \hat{b}_2 - \hat{A}_{22} y_2$
   (c) $r_1 = \hat{r}_1 + R^T \hat{r}_2$
3. *solve* $A_{11} y_1 = r_1$
   set $x_1 = y_1, \hat{x}_1 = x_1$
4. *form* $r_2 = b_2 - A_{21} x_1 - A_{22} y_2 = \hat{r}_2$ *as*
   (a) $\hat{t}_2 = R x_1 + y_2$

(b) $\hat{r}_2 = \hat{b}_2 - \hat{A}_{21}\hat{x}_1 - \hat{A}_{22}\hat{t}_2$

5. *smooth* $A_{22}(x_2 - y_2) = r_2$
   using symmetric Gauß-Seidel; denote approximate solution by $x_2 - y_2$
   set $\hat{x}_2 = \hat{t}_2 + (x_2 - y_2)$.

Procedure HB/MG is just a standard multigrid V-cycle with $j = 2$ levels. Steps (1) and (5) are the traditional multigrid smoothing steps, except that only the level-2 unknowns (y_2 and x_2) are smoothed, rather than all unknowns. Step (3) is the standard multigrid coarse grid correction, done by direct elimination for $j = 2$ (or inductively for $j > 2$). In step (2), the vector $(\hat{r}_1^T, \hat{r}_2^T)$ is just the fine grid residual after smoothing, while forming r_1 is the standard finite element version of the fine–to–coarse grid residual transfer. Step (4) updates the fine grid solution as $(\hat{x}_1^T, \hat{t}_2^T)$ using the standard finite element coarse–to–fine grid interpolation.

At this point, we briefly reflect on the mathematic aspects of the method developed in Sections 2–4. It seems clear that the only essential difference between Procedure HB/MG and standard multigrid methods is the restriction on which points are smoothed in steps (1) and (5). Since only some of the fine grid points are smoothed, HB/MG should converge more slowly than a comparable V-cycle in which all grid points are smoothed. The logarithmic growth in condition number is thus forshadowed.

In order to carry out Procedure HB/MG, one needs the fine grid submatrices $A_{22} = \hat{A}_{22}$, $\hat{A}_{21}$ and $\hat{A}_{12} = \hat{A}_{21}^T$, and the coarse grid matrix A_{11}. R is not stored since the action Rx is always computed in the usual multigrid fashion. $\hat{A}_{22}$ has at most 5 nonzeroes per row corresponding to a vertex x_i as in Fig. 3a and at most 3 nonzeroes per row corresponding to a vertex x_i as in 3b. The corresponding rows of $\hat{A}_{21}$ contain at most 2 and 3 nonzeroes, respectively. Thus the matrix storage is at most 7 nonzeroes per vertex of level 2, and because of the symmetry of $\hat{A}_{22}$ a further reduction is possible.

For any number of levels the amount of work and storage per level (except for the initial level) is a small constant multiple of the number of vertices on this level. As the initial level is fixed, the overall cost of the algorithm becomes asymptotically proportional to the number of unknowns, regardless the distribution of the vertices to the different levels.

As a numerical illustration, we consider the solution of Laplace's equation on a circle of radius one centered at the origin; the domain has a crack along the positive x axis. Homogeneous Dirichlet (Neumann) boundary conditions are imposed on the top (bottom) of the crack, allowing a singular solution with leading term

$$u = r^{1/4}\sin(\theta/4) \tag{5.8}$$

Equation (5.8) is imposed as a Dirichlet boundary condition on the remainder of the boundary, making (5.8) the exact solution.

The problem was solved using aposteriori error estimates and adaptive local mesh refinement, starting from a uniform coarse mesh with 8 elements and 10 vertices (Fig. 4.a) and ending with a nonuniform mesh including 28 levels of refinement and 2560 vertices (Fig. 4.b). The calculation was made

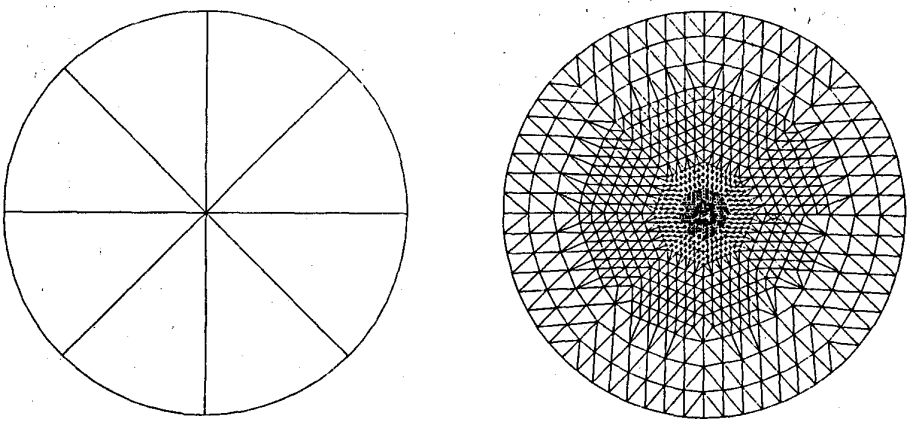

Figure 4 a,b

using an updated version of the package PLTMG [2], in which the hierarchical basis multigrid method, Case 3 from Section 3, was implemented. This implementation used ORTHOMIN acceleration, similar to minimum residual-conjugate gradient acceleration [10]. The distribution of vertices among the refinement levels is given in Table 1. These are approximately

level	1	2	3	4	5	6	7	8	9	10
vertices	10	17	58	212	328	219	204	173	135	141
level	11	12	13	14	15	16	17	18	19	20
vertices	149	139	118	96	110	95	80	42	33	31
level	21	22	23	24	25	26	27	28		
vertices	31	29	29	27	23	19	9	3		

Table 1

the dimensions of the spaces $\mathcal{V}_k$ of Section 3, except for the inclusion of Dirichlet boundary vertices. The meshes are nearly optimal for this problem and they illustrate the fact that local mesh refinement tends to generate subspaces which do not increase geometrically in dimension.

To illustrate the efficiency of the hierarchical basis multigrid method, we took the discrete solution $\hat{x}$ (in the nodal basis), formed

$$\hat{b} = \hat{A}\hat{x},$$

and resolved the system using $\hat{b}$ as data and initial guess zero. The number of correct digits in the energy norm is given by

$$\text{digits} = -\log(\|\hat{x}^{(i)} - \hat{x}\|/\|\hat{x}\|)$$

The results of the calculation are given in Table 2. Note $\hat{x}$ has about 1.44

cycle	1	2	3	4	5	6	7	8	9	10
digits	.44	.76	1.17	1.54	1.90	2.31	2.74	3.11	3.58	4.33

Table 2

digits of accuracy as a solution to the continuous problem, so only 4 cycles were required to reduce the initial error to the level of discretization error. Within the dynamic framework of an adaptive local mesh refinement procedure, intial guesses substantially better than zero are ususally available, further reducing the number of cycles required at any particular step of the adpative procedure.

From Table 2, we see that the average error reduction per iteration was

$$10^{-.433} \cong .37 = \delta.$$

If we assume

$$\delta \cong \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1},$$

then we can estimate the size of the generalized condition number κ by

$$\kappa \cong \left(\frac{1+\delta}{1-\delta}\right)^2 \cong 4.7.$$

Acknowledgement

We gratefully acknowledge the hospitality of Peter Deuflhard, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Federal Republic of Germany, who made his excellent facilities available to us during preparation of this manuscript.

The first author was supported in part by the United States Office of Naval Research under contract N00014-82-K-0197 and the third author by a Konrad-Zuse-Fellowship.

References

[1] O. Axelsson, V.A. Barker: *Finite Element Solution of Boundary Value Problems: Theory and Computation;* New York: Academic Press 1984

[2] R.E. Bank: *PLTMG User's Guide, Edition 4.0;* Technical Report, Department of Mathematics, University of California at San Diego, 1985

[3] R.E. Bank, T. Dupont.: *Analysis of a Two-Level Scheme for Solving Finite Element Equations;* Report CNA-159, Center for Numerical Analysis, University of Texas at Austin 1980

[4] R.E. Bank, T. Dupont: *An Optimal Order Process for Solving Finite Element Equations;* Math. Comp. 36, 35-51 (1981)

[5] R.E. Bank, A. Sherman, A. Weiser: *Refinement Algorithms and Data Structures for Local Mesh Refinement;* in: Scientific Computing, R. Stepleman et al. eds., Amsterdam: IMACS/ North Holland 1983

[6] D. Braess: *The Contraction Number of a Multigrid Method for Solving the Poisson Equation;* Numer. Math. 37, 387-404 (1981)

[7] J.H. Bramble, J.E. Pasciak, A.H. Schatz: *An Iterative Method for Elliptic Problems and Regions Partioned into Substructures;* Math. Comp. 46, 361-369 (1986)

[8] J.H. Bramble, J.E. Pasciak, A.H. Schatz: *The Construction of Preconditioners for Elliptic Problems by Substructuring. I;* Math. Comp. 47, 103-134 (1986)

[9] W. Hackbusch: *Multigrid Methods and Applications;* Berlin, Heidelberg, New York: Springer 1985

[10] L.A. Hageman, D.M. Young: *Applied Iterative Methods;* New York: Academic Press 1981

[11] J.F. Maitre, 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, W. Hackbusch and U. Trotttenberg eds., Lect. Notes Math. 960, Berlin, Heidelberg, New York: Springer 1982

[12] M.C. Rivara: *Algorithms for Refining Triangular Grids Suitable for Adaptive and Multigrid Techniques;* Int. J. Numer. Methods Eng. 20, 745-756 (1984)

[13] D.M. Young: *Convergence Properties of the Symmetric and Unsymmetric Successive Overrelaxation Method and Related Methods;* Math. Comp. 24, 793-807 (1970)

[14] H. Yserentant: *On the Multi-Level Splitting of Finite Element Spaces;* Numer. Math. 49, 379-412 (1986)

[15] H. Yserentant: *Hierarchical Bases Give Conjugate Gradient Type Methods a Multigrid Speed of Convergence;* Appl. Math. and Comp. 19, 347-358 (1986)

[16] H. Yserentant: *On the Multi-Level Splitting of Finite Element Spaces for Indefinite Elliptic Boundary Value Problems;* SIAM J. Numer. Anal. 23, 581-595 (1986)

[17] H. Yserentant: *Hierarchical Bases of Finite Element Spaces in the Discretization of Nonsymmetric Elliptic Boundary Value Problems;* Computing 35, 39-49 (1985)

APPENDIX: ARBITRARILY MANY INNER ITERATIONS

We use the notations of Section 3.

Theorem 1:

Let $A=L+D+L^T$ be a positive definite and symmetric matrix as in Section 3. Let

$$B = (L+D)^T D^{-1}(L+D) \quad .$$

Set $x^{(\emptyset)}=\emptyset$ and let

$$x^{(i+1)} = x^{(i)}+B^{-1}(b-Ax^{(i)}) \quad .$$

Then

$$(A+Z_i)x^{(i)} = b$$

where Z_i is a positive semidefinite symmetric matrix having the same kernel as L . Z_i satisfies

$$|A^{-1/2}Z_iA^{-1/2}| = \frac{(1 - \frac{1}{1+|Q|})^i}{1 - (1 - \frac{1}{1+|Q|})^i}$$

where Q is given by

$$Q = A^{-1/2}LT^{-1}LA^{-1/2} \quad .$$

Proof:

Using the exact solution $x=A^{-1}b$ one has

$$x^{(i+1)}-x = (E-B^{-1}A)(x^{(i)}-x)$$

and therefore, because of $x^{(0)}=0$,

$$x^{(i)} = (E-(E-B^{-1}A)^{i})x \quad .$$

By Theorem 3.3 one obtains

$$\|(E-B^{-1}A)^{i}\| \leq \|E-B^{-1}A\|^{i} < 1 \ .$$

This means that $E-(E-B^{-1}A)^{i}$ is nonsingular.

$$A(E-(E-B^{-1}A)^{i})^{-1}x^{(i)} = b$$

follows. Now

$$A(E-(E-B^{-1}A)^{i})^{-1} = A+Z_i$$

if and only if

$$\begin{aligned} Z_i &= A(E-B^{-1}A)^{i}(E-(E-B^{-1}A)^{i})^{-1} \\ &= A(A^{-1/2}RA^{1/2})^{i}(E-(A^{-1/2}RA^{1/2})^{i})^{-1} \\ &= A^{1/2}R^{i}(E-R^{i})^{-1}A^{1/2} \end{aligned}$$

where R is given by

$$R = E-A^{1/2}B^{-1}A^{1/2} \quad .$$

Using

$$B = A+L^T D^{-1} L = A^{1/2}(E+Q)A^{1/2} \quad .$$

one gets

$$R = E-(E+Q)^{-1} \quad .$$

Q is a positive semidefinite singular symmetric matrix. Therefore the eigenvalues of Q range from zero to $|Q|$ and the eigenvalues of the symmetric matrix R from zero to

$$|R| = 1 - \frac{1}{1+|Q|} < 1 \quad .$$

The function

$$\lambda \rightarrow \frac{\lambda^i}{1-\lambda^i} \quad , \quad 0 \leq \lambda < 1 \ ,$$

is monotonely increasing. Therefore we can conclude that the eigenvalues of the symmetric matrix

$$A^{-1/2} Z_i A^{-1/2} = R^i (E-R^i)^{-1}$$

range from zero to

$$\frac{|R|^i}{1-|R|^i} = \frac{(1 - \frac{1}{1+|Q|})^i}{1 - (1 - \frac{1}{1+|Q|})^i} \quad .$$

Using the given representations one sees that the symmetric matrices Q , R and $A^{-1/2} Z_i A^{-1/2}$ have the same eigenspaces and kernels. Therefore the kernel of Z_i is nothing else than the kernel of $L^T D^{-1} L$, which is the kernel of L . ■

Now we can generalize Case 3 by allowing for an arbitrary number of symmetric Gauß-Seidel steps for solving the inner equations.

Theorem 2:

If one solves the inner equations by m steps of the symmetric Gauß-Seidel method one has

$$\frac{1}{\mu_1} \leq (\alpha_0 + c(m)\alpha_1)^2$$

with

$$c(m)^2 = \frac{(1 - \frac{1}{1+q})^{2m}}{1 - (1 - \frac{1}{1+q})^{2m}}$$

and

$$q = |D^{-1/2} e^T_d{}^{-1} e D^{-1/2}| \quad .$$

Proof:

By Theorem 1, applied to the inner equations, the matrix $\tilde{D}$ has the representation

$$\tilde{D} = D+Z$$

with a positive semidefinite singular symmetric matrix Z satisfying

$$|D^{-1/2} Z D^{-1/2}| = \frac{(1 - \frac{1}{1+q})^m}{1 - (1 - \frac{1}{1+q})^m} \quad . \qquad (1)$$

Now we can proceed as in the proof of Lemma 3.7. The matrix B has the representation

$$B = (L+D+Z)^T(D+2Z)^{-1}(L+D+Z)$$

and therefore we have

$$\left(\frac{1}{\mu_1}\right)^{1/2} = |(D+2Z)^{-1/2}(L+D+Z)A^{-1/2}| \tag{2}$$

$$\leq |(D+2Z)^{-1/2}D^{1/2}||D^{-1/2}(L+D)A^{-1/2}| +$$

$$+ |(D+2Z)^{-1/2}ZD^{-1/2}||D^{1/2}A^{-1/2}| \quad .$$

As in the proof of Lemma 3.7 one obtains

$$|(D+2Z)^{-1/2}D^{1/2}| = 1 \quad . \tag{3}$$

If Λ denotes the set of eigenvalues of the positive semidefinite singular symmetric matrix

$$Q = D^{-1/2}ZD^{-1/2} \quad ,$$

as in the proof of Lemma 3.7 the set of the eigenvalues of the matrix

$$[(D+2Z)^{-1/2}ZD^{-1/2}]^T[(D+2Z)^{-1/2}ZD^{-1/2}]$$

is

$$\frac{\lambda^2}{1+2\lambda} \quad , \quad \lambda \in \Lambda \quad .$$

As the function

$$\lambda \to \frac{\lambda^2}{1+2\lambda} \quad , \quad \lambda \geq 0,$$

is monotonely increasing one obtains

$$|(D+2Z)^{-1/2}ZD^{-1/2}|^2 = \frac{|Q|^2}{1+2|Q|} \quad .$$

Substituting (1) for $|Q|=|D^{-1/2}ZD^{-1/2}|$ one gets

$$|(D+2Z)^{-1/2}ZD^{-1/2}| = c(m) \quad . \tag{4}$$

Inserting (3) and (4) into (2) and using (3.45), (3.46)

$$\frac{1}{\mu_1} \leq (\alpha_\emptyset + c(m)\alpha_1)^2$$

follows. ■

Applying Theorem 1 to the inner equations and using an argument as in the proof of Theorem 3.2 one gets

$$\mu_2 = 1 \quad .$$

Therefore

$$\kappa \leq (\alpha_\emptyset + c(m)\alpha_1)^2 \quad .$$

Because of

$$q = |D^{-1/2}e^T d^{-1} e D^{-1/2}| \leq (\beta\gamma_1)^2$$

this is a generalization of the condition number estimate (3.56).