UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Wavelet-Like Methods in the Design of Efficient Multilevel Preconditioners for Elliptic PDEs

Panayot S. Vassilevski
Junping Wang

August 1996

CAM Report 96-24

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

Wavelet-Like Methods in the Design of Efficient Multilevel Preconditioners for Elliptic PDEs

Panayot S. Vassilevski and Junping Wang

§1 Introduction

In this paper we are concerned with the construction of efficient numerical methods for matrix problems arising from finite element methods for elliptic partial differential equations. In practical computations, the standard nodal basis for the finite element space is often chosen as the computational basis and the resulting matrices are ill-conditioned. Our objective is to seek a substitution for the standard nodal basis so that the stiffness matrix arising from the new basis is well-conditioned.

A computationally feasible basis should possess the following properties: (a) the basis functions must be computable and have local support; (b) the resulting stiffness matrix is sparse and well-conditioned. This paper will introduce a wavelet-like method proposed by the authors in [41] which can be employed to construct a new basis with the above mentioned features for the finite element application to elliptic problems.

The method has a very close relation with the multiresolution (or multiscale) decompositions exploited especially in the wavelet literature (e.g., [17], [14]). In our approach, the requirements on the L^2-orthogonality and the existence of a single generating function ϕ (more precisely, ϕ generates an orthogonal basis by dilation and translation as in $\{\phi(\cdot - 2^{-k}\mathbf{i}),\ \mathbf{i} \in \boldsymbol{Z}\ \}$) commonly imposed in the wavelet literature (see Mallat [26] for details) are relaxed in order to have a computationally feasible basis. The basic idea lies behind approximating the wavelets without deteriorating their stability, yielding a stable Riesz basis for the finite element space under consideration.

put here title
put here (eds.), pp. 1–47.

put here ISBN

Attempts in the search of a stable Riesz basis with some restrictions, either on the mesh or on the analysis, have been made in Griebel and Oswald [20], Kotyczka and Oswald [24], and Stevenson [33, 34]. For a recent comparative study on the construction of economical Riesz bases for Sobolev spaces we refer to Lorentz and Oswald [25].

Our method is general and provides a satisfactory answer for most of the elliptic equations. The method is based on modifying the existing (unstable) hierarchical basis by using operators which are approximations of the L^2–projections onto coarse finite element spaces. A similar approach was used by Jaffard [22] in the seeking of a wavelet basis for finite element spaces. In [22], the construction starts with the standard nodal basis functions which are transformed to an orthogonal multiresolution basis based on an explicit orthogonalization procedure exploiting wavelets in each subspace W_j of V_j. Here W_j is the L^2–orthogonal complement of the finite element space V_{j-1} in the next fine finite element space V_j. The latter procedure generally leads to basis functions that are not locally supported but have good decay rates and hence allow for locally-supported approximations. The method proposed in Vassilevski and Wang [41] can be regarded as an approximation of the wavelet basis in [22]; the approximate-wavelet basis functions are locally-supported and, therefore, computationally feasible.

There are two alternatives in the implementation of the approximate-wavelet basis in practical computation. In the first approach, one first computes an explicit form for each new basis function and then assembles the global stiffness matrix corresponding to the new basis. The major drawback for this approach is the difficulty on the assembly of the global stiffness matrix; the robustness corresponding to the standard nodal basis is no longer retained for the new basis. In the second approach, one makes use of the standard nodal basis as the computational one and constructs a preconditioner for the stiffness matrix by using the new basis functions. This is possible because most of the iterative algorithms such as the CG method require only the action of the stiffness matrix on vectors. Thus, the technique of basis changing can be employed to compute the action of the matrix corresponding to the new basis. The second approach can be viewed as a "black–box fix" of the conditioning of the standard nodal basis stiffness matrices. Details can be found in §8.

The organization of the present paper is as follows. Section 2 contains some preliminaries and the problem formulation in an abstract Hilbert space setting. The importance of having a stable Riesz basis and its relation to the conditioning of discrete matrix problems, especially those from the finite element discretization, is the topic of §3. Specific examples of second order elliptic PDEs are presented in §4. Section 5 discusses the basic idea in constructing multilevel direct space decompositions. The hierarchical basis of Yserentant [44] is reviewed in §6. Section 7 contains a detailed discussion of the wavelet-like method in constructing a stable basis. The implementation issues, including two basic preconditioning schemes (namely the additive and multiplicative algorithms) and the construction of the approximate L^2–projections, are discussed in §8. Finally, some numerical results are presented in §9 for convection–diffusion problems to confirm the theory developed in previous sections.

§2 Preliminaries and basic problems

This preliminary section introduces the basic problem we are concerned about in this paper. Let H be a Hilbert space equipped with the inner product $(\cdot,\cdot)$ such that

$$H \subset V \subset H', \tag{2.1}$$

where H is dense in another Hilbert space V with inner product $(\cdot,\cdot)_0$, and H' is the dual of H with respect to the pairing of V defined by the inner product $(\cdot,\cdot)_0$. Assume that the imbedding from H to V is continuous. More precisely, there exists a constant $\varrho > 0$ satisfying

$$\|u\|_0 \le \varrho \|u\| \qquad \forall u \in H. \tag{2.2}$$

Here $\|\cdot\| = \sqrt{(\cdot,\cdot)}$ indicates the norm in the Hilbert space H and $\|\cdot\|_0 = \sqrt{(\cdot,\cdot)_0}$ that of V.

Let $a(\cdot,\cdot)$ be a bounded bilinear form defined on $H \times H$, satisfying the following *inf–sup* condition of Ladyzhenskaya, Babuška, and Brezzi:

$$\sup_{w\in H} \frac{a(v,w)}{\|w\|} \ge \beta \|v\| \qquad \forall v \in H, \tag{2.3}$$

where $\beta > 0$ is a fixed constant. We are interested in approximate solutions of the following problem:

Given $f \in H'$, find $u \in H$ satisfying

$$a(u,\psi) = f(\psi) \qquad \forall \psi \in H. \tag{2.4}$$

It is well–known that if $a(\cdot,\cdot)$ is symmetric and H-coercive (i.e., there exists a positive α satisfying $a(v,v) \ge \alpha\|v\|^2$ for all $v \in H$), then the problem (2.4) is equivalent to the minimization problem for the quadratic functional $\mathcal{J}(v) = \frac{1}{2}a(v,v) - f(v)$.

To approximate (2.4), let $\{H_k\}_{k=1}^{\infty}$ be a sequence of finite dimensional subspaces of H satisfying the following approximation property: *For any $v \in H$ one has*

$$\lim_{k\to\infty} \inf_{\phi\in H_k} \|v-\phi\| = 0. \tag{2.5}$$

The well–known Galerkin method for (2.4) seeks $u_k \in H_k$ such that

$$a(u_k, \psi_k) = f(\psi_k) \qquad \forall \psi_k \in H_k. \tag{2.6}$$

In order to have a well–posed discrete problem (2.6), we assume the following discrete *inf–sup* condition: There exists a constant $\beta_k > 0$ such that

$$\sup_{w\in H_k} \frac{a(v,w)}{\|w\|} \ge \beta_k \|v\| \qquad \forall v \in H_k. \tag{2.7}$$

It is known that if (2.7) holds true, then the Galerkin problem (2.6) has a unique solution in the subspace H_k. Interested readers are referred to [18] and [11] for more details.

In practical computation, we usually formulate a matrix problem for (2.6) by choosing a suitable basis for the subspace H_k. More precisely, let $\{\phi_i^{(k)} : \quad i = 1, 2, \cdots, n_k\}$ be a computational basis of H_k. Expand the approximate solution $u^{(k)}$ in terms of this basis, yielding

$$u^{(k)} = \sum_{i=1}^{n_k} c_i \phi_i^{(k)}. \tag{2.8}$$

Let $\mathbf{u}^{(k)} = (c_1, c_2, \ldots, c_{n_k})^T$ be the coordinates of $u^{(k)}$. It is not hard to see that the vector $\mathbf{u}^{(k)}$ is given as the solution of the following linear system:

$$A^{(k)}\mathbf{u}^{(k)} = \mathbf{f}^{(k)}, \tag{2.9}$$

where the right-hand side vector $\mathbf{f}^{(k)}$ is defined as follows:

$$\mathbf{f}^{(k)} = (b_1, b_2, \ldots, b_{n_k})^T \qquad \text{with } b_i = f(\phi_i^{(k)}). \tag{2.10}$$

The matrix is given by $A^{(k)} = \{a(\phi_j^{(k)}, \phi_i^{(k)})\}_{i,j=1}^{n_k}$. One may view $A^{(k)}$ as a linear operator on H_k defined, for any $v = \sum_{i=1}^{n_k} v_i \phi_i^{(k)} \in H_k$, by

$$A^{(k)}v = \sum_{i=1}^{n_k} c_i \phi_i^{(k)}, \quad \text{where } c_i = \sum_{j=1}^{n_k} a(\phi_j^{(k)}, \phi_i^{(k)})\, v_j. \tag{2.11}$$

Of main interest in this paper, we study techniques for solving the linear system (2.9) by iterative methods. It is known that the conditioning of the matrices $A^{(k)}$ is of great importance in practical computation. Let us see how the condition number of $A^{(k)}$ is related to the choice of the basis $\Phi^{(k)} \equiv \{\phi_i^{(k)}\}_{i=1}^{n_k}$ and the discrete *inf–sup* condition (2.7).

For any $v \in H_k$, denote in the bold face the coordinates of v with respect to the basis $\Phi^{(k)}$. Namely, the vectors $\mathbf{v} = (v_1, v_2, \cdots, v_{n_k})^T$ and v are related as follows:

$$v = \sum_{i=1}^{n_k} v_i \phi_i^{(k)}. \tag{2.12}$$

Introduce a new inner-product in H_k by using the basis $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ as follows:

$$\langle v, w\rangle_{k,0} = \sum_{i=1}^{n_k} v_i w_i \equiv \mathbf{v}\cdot\mathbf{w}. \tag{2.13}$$

Denote by $\|v\|_{k,0} = \sqrt{\mathbf{v}\cdot\mathbf{v}}$ the norm induced by the new inner-product. Recall that the condition number of the matrix $A^{(k)}$ is defined by

$$\kappa(A^{(k)}) = \|A^{(k)}\|\,\|A^{(k)^{-1}}\|, \tag{2.14}$$

where $\|A^{(k)}\|$ indicates the norm of the matrix $A^{(k)}$ with respect to the new norm $\|\cdot\|_{k,0}$ in the vector space H_k. The following result can be checked easily.

Theorem 1. *Assume that there are positive constants $\varrho_{k,1}$ and $\varrho_{k,2}$ satisfying*

$$\varrho_{k,1}\|v\|_{k,0} \le \|A^{(k)}v\|_{k,0} \le \varrho_{k,2}\|v\|_{k,0}. \tag{2.15}$$

Then, $\kappa(A^{(k)}) \le \frac{\varrho_{k,2}}{\varrho_{k,1}}$. The best estimate for $\varrho_{k,1}$ and $\varrho_{k,2}$ is given by

$$\begin{aligned}\varrho_{k,2} &= \sup_{v\in H_k,\ \|v\|_{k,0}=1} \|A^{(k)}v\|_{k,0},\\ \varrho_{k,1} &= \inf_{v\in H_k,\ \|v\|_{k,0}=1} \|A^{(k)}v\|_{k,0}.\end{aligned} \tag{2.16}$$

Moreover, one has $\kappa(A^{(k)}) = \frac{\varrho_{k,2}}{\varrho_{k,1}}$ for the best estimate of $\varrho_{k,i}$.

In fact, the inequality (2.15) implies the following

$$\|A^{(k)}\| \le \varrho_{k,2} \qquad \|A^{(k)^{-1}}\| \le 1/\varrho_{k,1}, \tag{2.17}$$

which verifies the validity of the theorem. Therefore, it suffices to establish an estimate like (2.15) in order to gain some knowledge on the condition number of the matrix $A^{(k)}$.

§3 Matrix conditioning and stable Riesz bases

Our objective in this section is to figure out the connection between the condition number of $A^{(k)}$ and the selection of the basis $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ for H_k. To this end, we derive the inequality (2.15) by using the *inf–sup* condition.

First, notice that for any $v \in H_k$

$$\|A^{(k)}v\|_{k,0} = \sup_{w\in H_k,\ \|w\|_{k,0}=1} \langle A^{(k)}v, w\rangle_{k,0}. \tag{3.1}$$

The definition of $A^{(k)}$ (see (2.11)) implies that

$$\langle A^{(k)}v, w\rangle_{k,0} = a\left(\sum_{j=1}^{n_k} v_j\phi_j^{(k)},\ \sum_{i=1}^{n_k} w_i\phi_i^{(k)}\right) = a(v,w). \tag{3.2}$$

It follows that

$$\|A^{(k)}v\|_{k,0} = \sup_{w\in H_k,\ \|w\|_{k,0}=1} a(v,w). \tag{3.3}$$

By letting

$$\varrho_{k,2} = \sup_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|_{k,0}\|w\|_{k,0}}, \tag{3.4}$$

one obtains the following

$$\|A^{(k)}v\|_{k,0} \le \varrho_{k,2}\|v\|_{k,0}. \tag{3.5}$$

Next, by letting

$$\varrho_{k,1} = \inf_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|_{k,0}\|w\|_{k,0}}, \tag{3.6}$$

one has the following estimate from below:

$$\|A^{(k)}v\|_{k,0} \ge \varrho_{k,1}\|v\|_{k,0}. \tag{3.7}$$

To summarize, we have proved the following result:

Theorem 2. *If $\varrho_{k,1}$ and $\varrho_{k,2}$ are given by (3.6) and (3.4), respectively, then the estimate (2.15) holds true. Consequently, the condition number of the matrix $A^{(k)}$ is bounded by $\frac{\varrho_{k,2}}{\varrho_{k,1}}$.*

The question now is on the determination of the constants $\varrho_{k,1}$ and $\varrho_{k,2}$. We would like to select a basis $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ for H_k so that the ratio (or the condition number of $A^{(k)}$) $\frac{\varrho_{k,2}}{\varrho_{k,1}}$ is as small as possible.

As our first consideration, we assume that $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ is an orthonormal basis with respect to the original inner product $(\cdot,\cdot)$ of H. It is clear that the corresponding discrete norm $\|\cdot\|_{k,0}$ is the same as the original norm $\|\cdot\|$ in H_k. Therefore, the constant $\varrho_{k,2}$ is bounded from above by the norm of the bilinear form $a(\cdot,\cdot)$. Similarly, the constant $\varrho_{k,1}$ is bounded from below by the parameter β_k in the ***inf–sup*** condition (2.7). Since in practical problems, the norm of $a(\cdot,\cdot)$ and the parameter β_k stays uniformly bounded in terms of k, then the condition number of $A^{(k)}$ is uniformly bounded.

It is, of course, impractical to assume the existence of an orthonormal basis $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ which is also computationally feasible. The next best thing to an orthonormal basis is the Riesz basis in any Hilbert space. We recall that in the Hilbert space H, a Riesz basis is a basis $\{\psi_k\}_{i=1}^{\infty}$ of H satisfying

$$\sigma_1\|v\|^2 \leq \sum_{i=1}^{\infty} c_i^2 \leq \sigma_2\|v\|^2 \qquad \forall v \in H, \tag{3.8}$$

where $v = \sum_{i=1}^{\infty} c_i\psi_i$ and $\|\cdot\|$ indicates the original norm in H; σ_1 and σ_2 are two absolute constants. Here we have assumed that the Hilbert space is separable.

Remark 1: The strong norm $\|\cdot\|$ in the Hilbert space H was used in the definition of the Riesz stability (3.8). Notice that H is a subspace of V. Thus, it would be feasible to discuss the Riesz property with respect to the norm $\|\cdot\|_0$ of V. In practical applications, V often represents $L^2(\Omega)$ and $\|\cdot\|_0$ denotes the corresponding L^2–norm. In the wavelet literature, the Riesz property is commonly studied with respect to the L^2–norm. But the ultimate goal in the preconditioning analysis is to establish estimates such as (3.8) since (3.8) implies that the discretization matrix $A^{(k)}$ is well–conditioned.

We now go back to the finite dimensional subspace H_k of H. Similar to (3.8), we assume that there exist constants $\sigma_1^{(k)}$ and $\sigma_2^{(k)}$ such that

$$\sigma_1^{(k)}\|v\|^2 \leq \sum_{i=1}^{n_k} c_i^2 \leq \sigma_2^{(k)}\|v\|^2 \qquad \forall v \in H_k. \tag{3.9}$$

Observe that the condition (3.9) can be rewritten as

$$\sigma_1^{(k)}\|v\|^2 \leq \|v\|_{k,0}^2 \leq \sigma_2^{(k)}\|v\|^2 \qquad \forall v \in H_k. \tag{3.10}$$

Since $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ is assumed to be a basis of H_k, the equivalence relation (3.10) obviously holds for some constants $\sigma_1^{(k)}$ and $\sigma_2^{(k)}$ with

$$\sigma_2^{(k)} = \sup_{v\in H_k,\ \|v\|=1} \|v\|_{k,0}^2 \tag{3.11}$$

and

$$\sigma_1^{(k)} = \inf_{v\in H_k,\, \|v\|=1} \|v\|_{k,0}^2. \tag{3.12}$$

The important thing is that one should have some control on the ratio $\sigma_2^{(k)}/\sigma_1^{(k)}$.

Definition 1. *The family $\left\{\Phi^{(k)} \equiv \{\phi_i^{(k)}\}_{i=1}^{n_k}\right\}$ is said to be a uniformly stable family of Riesz bases for $\{H_k\}$ if the quotient $\sigma_2^{(k)}/\sigma_1^{(k)}$ of the Riesz bounds in (3.9) is bounded uniformly with respect to $k \to \infty$; i.e., if there exists a constant M independent of k such that*

$$\sigma_2^{(k)}/\sigma_1^{(k)} \le M, \tag{3.13}$$

for any $k = 1, 2, \cdots$.

The rest of this section explains the importance of having a stable Riesz basis.

First we estimate the parameter $\varrho_{k,2}$ defined in (3.4). Let $\|a\|$ denote the norm of the bilinear form $a(\cdot,\cdot)$. Then,

$$\begin{aligned}
\varrho_{k,2} &= \sup_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|_{k,0}\|w\|_{k,0}} \\
&\le \sup_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|\|w\|} \sup_{v\in H_k}\sup_{w\in H_k} \frac{\|v\|\|w\|}{\|v\|_{k,0}\|w\|_{k,0}} \\
&= \|a\| \sup_{v,w\in H_k} \frac{\|v\|\|w\|}{\|v\|_{k,0}\|w\|_{k,0}} \le \|a\|/\sigma_1^{(k)}.
\end{aligned} \tag{3.14}$$

Next, we estimate the parameter $\varrho_{k,1}$ as follows:

$$\begin{aligned}
\varrho_{k,1} &= \inf_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|_{k,0}\|w\|_{k,0}} \\
&\ge \inf_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|_{k,0}\|w\|} \inf_{w\in H_k} \frac{\|w\|}{\|w\|_{k,0}} \\
&\ge \inf_{v\in H_k}\sup_{w\in H_k} \frac{a(v,w)}{\|v\|\|w\|} \inf_{v\in H_k} \frac{\|v\|}{\|v\|_{k,0}} \inf_{w\in H_k} \frac{\|w\|}{\|w\|_{k,0}} \\
&\ge \beta_k/\sigma_2^{(k)}.
\end{aligned} \tag{3.15}$$

It follows that $\varrho_{k,2}/\varrho_{k,1} \le \frac{\|a\|}{\beta_k}\frac{\sigma_2^{(k)}}{\sigma_1^{(k)}}$. The result can be summarized as follows:

Theorem 3. *Let $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ be a computational basis of H_k satisfying (3.10). Then the condition number of the matrix $A^{(k)}$ arising from the basis $\{\phi_i^{(k)}\}_{i=1}^{n_k}$ is bounded by,*

$$\kappa(A^{(k)}) \leq \frac{\|a\|}{\beta_k} \frac{\sigma_2^{(k)}}{\sigma_1^{(k)}}. \tag{3.16}$$

Consequently, the condition number of $A^{(k)}$ is uniformly bounded for stable Riesz bases $\{\phi_i^{(k)}\}_{i=1}^{n_k}$, provided that the discrete inf–sup condition (2.7) holds true with uniformly bounded constants β_k from below by some $\beta^* > 0$.

The conditioning estimate in Theorem 3 contains two important factors: $\frac{\|a\|}{\beta_k}$ and $\frac{\sigma_2^{(k)}}{\sigma_1^{(k)}}$. The first one depends on the norm of the given bilinear form and the stability constant β_k from the discrete *inf–sup* condition. In practical computations, the space H_k must be so constructed that ensures the boundedness of β_k from below by some fixed $\beta^* > 0$. This is the case for the model problems and their discretization spaces to be considered in §4. The second factor is basis–dependent. More precisely, it is a characterization of the difference between the discrete coefficient norm $\|\cdot\|_{k,0}$ and the continuous norm $\|\cdot\|$. Stable Riesz bases are important because the corresponding discretization matrices are *well–conditioned*. Thus, simple iterative methods such as the conjugate gradient (CG) can be successfully applied to solve the matrix problem from the Galerkin discretization with a geometric rate of convergence. As is well–known, the convergence factor is bounded by $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$, where $\kappa = \kappa(A^{(k)}) \leq \frac{\|a\|}{\beta_k} \frac{\sigma_2^{(k)}}{\sigma_1^{(k)}}$. For nonsymmetric problems, one could apply the CG method to the normal equation or the GMRES method to $A^{(k)}$. For symmetric and indefinite problems, the MINRES (minimum residual) algorithm would be a good choice. The convergence rate is no morse than that of the CG–method applied to $A^{(k)^2}$.

There is another practical criterion in the choice of the basis $\{\phi_i^{(k)}\}_{i=1}^{n_k}$. It is of great practical importance to represent the matrix entries of $A^{(k)}$ as sparse as possible. This is trivially achieved (assuming $a(.,.)$ is symmetric and positive definite) if the basis is $a(.,.)$–orthogonal. Thus, the matrix $A^{(k)}$ admits diagonal forms and only n_k entries need to be stored. Such a situation is too special and rarely happens in practice. In general, we assume that the basis is computationally feasible in the sense that the basis function are computable and the corresponding matrix $A^{(k)}$ is sparse (the number of nontrivial entries in the matrix is of order $O(n_k)$). This is the case in practice for the standard nodal bases of finite element spaces H_k if the bilinear form arises from partial differential equations. However, this choice will not make a stable Riesz basis for most of the PDEs. The following section contains a detailed discussion on this aspect.

§4 Model problems

Here we consider some model problems of (2.4) in partial differential equations. Boundary value problems for the second order elliptic equations are of major consideration in this discussion. Finite element methods will be applied to approximate the solution defined on an open bounded domain Ω in $\mathbf{R}^d$ with $d = 2$ or 3.

4.1 Second-order elliptic equations

Consider the homogeneous Dirichlet boundary value problem for the following second-order elliptic equation:

$$\mathcal{L}(u) \equiv -\nabla \cdot (a(x)\nabla u) + \mathbf{b}(x) \cdot \nabla u + c(x)u = f(x), \quad x \in \Omega, \tag{4.1.1}$$

where $a = a(x)$ is a symmetric and positive definite matrix with bounded and measurable entries, $\mathbf{b} = \mathbf{b}(x)$ and $c = c(x)$ are given bounded functions, $f = f(x)$ is a function in $H^{-1}(\Omega)$.

Note that we do not intend to consider problems which are convection–dominated.

Let $H^1(\Omega)$ be the standard Sobolev space equipped with the norm:

$$\|u\|_1 = \left(\|u\|_0^2 + \|\nabla u\|_0^2\right)^{1/2} \qquad \forall u \in H^1(\Omega). \tag{4.1.2}$$

Here $\|\cdot\|_0$ stands for the L^2-norm. Let $H_0^1(\Omega)$ be the closed subspace of $H^1(\Omega)$ consisting of functions with vanishing boundary values. The following relation is well–known:

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega). \tag{4.1.3}$$

A weak form for the Dirichlet problem of (4.1.1) seeks $u \in H_0^1(\Omega)$ satisfying

$$b(u, v) = f(v) \qquad \forall v \in H_0^1(\Omega), \tag{4.1.4}$$

where

$$b(u, v) = \int_\Omega (a(x)\nabla u \cdot \nabla v + \mathbf{b}(x) \cdot \nabla u \; v + c(x)uv)\, dx \tag{4.1.5}$$

and $f(v)$ is the action of the linear functional f on v.

Assume that the problem (4.1.4) has a unique solution. Then the *inf–sup* condition (2.3) is satisfied for the bilinear form $b(\cdot,\cdot)$ defined on $H_0^1(\Omega)\times H_0^1(\Omega)$. Let us approximate (4.1.4) by using the Galerkin method with continuous piecewise polynomials. If S_h denotes the finite element space associated with a prescribed triangulation of Ω with mesh size h, then the Galerkin approximation is given as the solution of the following problem: *Find* $u_h \in S_h$ *satisfying*

$$b(u_h, \phi) = f(\phi) \qquad \forall \phi \in S_h. \tag{4.1.6}$$

It has been shown that the discrete problem (4.1.6) has a unique solution when the mesh size h is sufficiently small. Thus, the discrete *inf–sup* condition (2.7) is satisfied for this problem. Details can be found from [35, 36].

Choose the standard nodal basis as the computational basis for the finite element space S_h. Let $\{\phi_i\}_{i=1}^n$ be the set of nodal basis functions and A_h be the corresponding discrete matrix (also called the global stiffness matrix). The condition number of A_h can be estimated by using Theorem 3. To this end, let us establish the inequality (3.10) for the standard nodal basis. For any $v \in S_h$, let

$$v = \sum_{i=1}^{n} v_i \phi_i \qquad \text{with } v_i = v(x_i), \tag{4.1.7}$$

where x_i are the interior nodal points of the finite element partition. The relation (3.10) is equivalent to the following:

$$\hat{\sigma}_1 \|v\|_1^2 \leq \sum_{i=1}^{n} v_i^2 \leq \hat{\sigma}_2 \|v\|_1^2, \qquad \forall v \in S_h, \tag{4.1.8}$$

for some positive constants $\hat{\sigma}_1$ and $\hat{\sigma}_2$. It is not hard to see that

$$\sum_{i=1}^{n} v_i^2 \simeq h^{-d} \|v\|_0^2. \tag{4.1.9}$$

It follows that $\hat{\sigma}_2 = \mathcal{O}(h^{-d})$. Also, by the standard inverse inequality one sees that $\hat{\sigma}_1$ is bounded from below by a constant proportional to h^{2-d}. Thus, from Theorem 3, the condition number of A_h is bounded from above by Ch^{-2} for some constant C; the lower bound for its condition number is also bounded from below by some Ch^{-2}.

4.2 Stokes equations

Consider the problem which seeks $\mathbf{u} \in \left[H_0^1(\Omega)\right]^d$ and $p \in L^2(\Omega)$ such that

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f}, && \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0, && \text{in } \Omega, \\ \mathbf{u} &= 0, && \text{on } \partial\Omega, \end{aligned} \tag{4.2.1}$$

where $\mathbf{f} \in \left[L^2(\Omega)\right]^d$ is a given vector-valued function, and $\partial\Omega$ denotes the boundary of Ω.

A weak form for the problem (4.2.1) involves the following bilinear form:

$$\mathcal{A}(\mathbf{u}, p; \mathbf{v}, w) \equiv (\nabla\mathbf{u}, \nabla\mathbf{v})_0 - (\nabla \cdot \mathbf{v}, p)_0 - (\nabla \cdot \mathbf{u}, w)_0 \tag{4.2.2}$$

defined on $\mathcal{W} \times \mathcal{W}$ with $\mathcal{W} = \left[H_0^1(\Omega)\right]^d \times L_0^2(\Omega)$. Here $L_0^2(\Omega)$ is the closed subspace of $L^2(\Omega)$ consisting of functions with vanishing mean value. The weak problem seeks $(\mathbf{u}, p) \in \mathcal{W}$ satisfying

$$\mathcal{A}(\mathbf{u}, p; \mathbf{v}, w) = (\mathbf{f}, \mathbf{v}) \qquad \forall (\mathbf{v}, w) \in \mathcal{W}. \tag{4.2.3}$$

The *inf–sup* condition is satisfied for the bilinear form defined in (4.2.2). Details can be found from [18], [11].

As to the finite element method for (4.2.3), we employ the Hood–Taylor element [21] which satisfies the discrete *inf–sup* condition (with a mild restriction on the triangulation). The Hood–Taylor element is a combination of continuous piecewise linear functions for the pressure variable p and continuous piecewise quadratic functions for the velocity variable $\mathbf{u}$. Denote by $\mathcal{W}_h = X_h \times S_h$ the corresponding finite element space, where X_h contains continuous piecewise quadratic functions and S_h contains continuous piecewise linear functions for the pressure variable. The finite element approximation $(\mathbf{u}_h, p_h) \in \mathcal{W}_h$ satisfies

$$\mathcal{A}(\mathbf{u}_h, p_h; \mathbf{v}, w) = (\mathbf{f}, \mathbf{v}) \qquad \forall (\mathbf{v}, w) \in \mathcal{W}_h. \tag{4.2.4}$$

If the standard nodal basis is selected in formulating a matrix problem for (4.2.4), then the condition number of the global stiffness matrix can be estimated by using Theorem 3. More precisely, let $\{\psi_j\}_{j=1}^m$ be the standard nodal basis of X_h and $\{\phi_i\}_{i=1}^n$ that of S_h as discussed in the previous section. For any $(\mathbf{v}, w) \in X_h \times S_h$, let

$$w = \sum_{i=1}^n w_i \phi_i, \tag{4.2.5}$$

and

$$\mathbf{v} = \sum_{j=1}^m v_j \psi_j. \tag{4.2.6}$$

Using the inverse inequality and the Poincaré inequality one can derive the following relations:

$$\sum_{i=1}^n w_i^2 \simeq h^{-d} \|w\|_0^2, \tag{4.2.7}$$

and

$$c_1 h^{2-d}\|\mathbf{v}\|_1^2 \le \sum_{j=1}^{m} v_j^2 \le c_2 h^{-d}\|\mathbf{v}\|_1^2, \tag{4.2.8}$$

where c_1 and c_2 are two absolute constants. Thus, we have from Theorem 3 that the condition number of the global stiffness matrix is bounded by Ch^{-2}.

We emphasize that the poor conditioning for the Stokes problem is caused by the relation (4.2.8) where the H^1-norm of $\mathbf{v}$ was approximated by a discrete norm stable in L^2 only. The equivalence (4.2.7) indicates that the standard nodal basis is a good choice for the pressure variable in the Stokes problem. Therefore, attention should be focused on stabilizing the velocity component in the Stokes equation.

A direct wavelet approach to the Stokes problem has been developed in Dahmen, Kunoth, and Urban [16]. One could also use block-diagonal preconditioners for the saddle-point discretization matrices $A^{(k)}$ with one block corresponding to Laplace-like preconditioners for the velocity component and a second block corresponding to mass-matrix preconditioners for the pressure unknown in the MINRES method. Details in this approach can be found from Rusten and Winther [31] and Silvester and Wathen [32].

4.3 Mixed methods

Here we consider the mixed method for the second order elliptic equation (4.1.1). For simplicity, assume that $\mathbf{b} \equiv 0$, $c \equiv 0$, and the following Dirichlet boundary condition

$$u = -g \qquad \text{on } \partial\Omega \tag{4.3.1}$$

is imposed on the solution. Let

$$H(\text{div};\ \Omega) = \left\{\mathbf{v} :\ \mathbf{v} \in \left[L^2(\Omega)\right]^d,\ \nabla\cdot\mathbf{v} \in L^2(\Omega)\right\},$$

which is equipped with the following norm:

$$\|\mathbf{v}\|_{H(\text{div};\ \Omega)} = \left(\int_\Omega (|\mathbf{v}|^2 + |\nabla\cdot\mathbf{v}|^2)dx\right)^{1/2}.$$

Let $\alpha(x) = a^{-1}(x)$ be the inverse of the coefficient matrix $a = a(x)$ and

$$\mathcal{A}(\mathbf{q}, u; \mathbf{v}, w) \equiv (\alpha(x)\mathbf{q}, \mathbf{v})_0 - (\nabla\cdot\mathbf{v}, p)_0 - (\nabla\cdot\mathbf{q}, w)_0$$

be a bilinear form defined on $\mathcal{W}\times\mathcal{W}$ where $\mathcal{W} \equiv H(\text{div};\ \Omega)\times L^2(\Omega)$. Then, a mixed weak form for (4.1.1) with the boundary condition (4.3.1) seeks $(\mathbf{q}, u) \in \mathcal{W}$ satisfying

$$\mathcal{A}(\mathbf{q}, u; \mathbf{v}, w) = \langle g, \mathbf{v}\cdot\mathbf{n}\rangle_{\partial\Omega} - (f, w)_0 \qquad \forall(\mathbf{v}, w) \in \mathcal{W}, \tag{4.3.2}$$

where $\langle \cdot, \cdot \rangle_{\partial\Omega}$ denotes the inner product in $L^2(\partial\Omega)$.

The *inf–sup* condition (2.3) can be verified for the bilinear form $\mathcal{A}(\cdot;\cdot)$. Furthermore, finite element spaces satisfying the discrete *inf–sup* condition (2.7) are available for this bilinear form. Details can be found from the book by Brezzi and Fortin [11]. If the standard nodal basis for the finite element space of Raviart and Thomas is employed in practical computation, the condition number of the global stiffness matrix is known to be proportional to h^{-2}.

The question is how to stabilize the nodal basis. Observe that in this application, one needs to construct a basis for the mixed finite element space (called $\mathcal{W}_h$) so that the discrete norm is equivalent to the following norm:

$$\|(\mathbf{v}; w)\|^2_{\mathcal{W}} = \|\mathbf{v}\|^2_{H(\mathrm{div};\,\Omega)} + \|w\|^2_0.$$

One sees from above that the standard nodal basis is a good choice for the pressure unknown. Thus, the difficulty is on stabilizing the flux component with the norm in $H(\mathrm{div};\ \Omega)$. We do not have a positive answer yet for this stabilization. But the approach of Helmholtz decomposition for vectors might provide a partial answer for problems of two-space variables. The method decomposes each vector-valued function $\mathbf{v}$ as follows:

$$\mathbf{v} = \mathbf{curl}\,\phi + \nabla\psi, \tag{4.3.3}$$

where ϕ and ψ are functions in $H^1(\Omega)$. A discrete version of (4.3.3) can be studied in order to apply it to the mixed method. Results for the standard conforming and non-conforming finite elements should be investigated first in this direction. We point out that this approach may have some difficulty for problems of three–space variables.

For multilevel methods relying on the above Helmholtz decomposition, see Vassilevski and Wang [40] and Arnold, Falk, and Winther [1].

In the following sections we will devote ourselves to modifying the nodal bases in the application to second order elliptic problems. Our goal is to construct some stable Riesz bases that are computationally feasible for elliptic equations. More precisely, the basis should be so constructed that the resulting discretization matrices are both well–conditioned and sparse.

§5 Multilevel direct decompositions

To construct a computationally stable basis for the second-order elliptic and the Stokes equations, we take advantage of the fact that the weak problem (2.4) is discretized on a sequence of finite element subspaces. In particular, a sequence of nested subspaces may be possible in practical computations. Our objective in this section is to exploit ways of constructing stable basis by using the information from each approximating subspace.

5.1 The basic idea

The basic idea comes from the fact that an L^2 orthonormal basis of wavelets is also H^1-stable in applications to partial differential equations. Therefore, wavelet bases are good candidates in formulating matrix problems for (2.6). Since the conventional wavelet bases have complicated structures which limit their application in the numerical methods, we shall focus our attention on approximations of wavelet bases. Below we present a detailed discussion.

Assume that we have a sequence of nested subspaces $\{V_j\}_{j=0}^{\infty}$ satisfying

$$V_0 \subset V_1 \subset \ldots \subset V_k \subset \cdots. \tag{5.1.1}$$

Each vector space V_j shall be referred to as a coarse subspace of V_k when $j < k$. In applications, they are finite element spaces consisting of continuous piecewise polynomials over a sequence of finite element partitions for the domain Ω. Upon viewing V_j as a subspace of $L^2(\Omega)$, one has

$$\overline{\bigcup_{j=0}^{\infty} V_j} = L^2(\Omega), \tag{5.1.2}$$

where the closure was taken in the strong topology induced by the L^2-norm. We assume that V_0 is a very coarse subspace of $L^2(\Omega)$ whose dimension is a small number.

For every $j \geq 1$, define W_j to be the L^2-orthogonal complement of V_{j-1} in V_j. We have

$$V_j = V_{j-1} \oplus W_j \tag{5.1.3}$$

and

$$W_i \perp W_j \quad \text{if} \quad i \neq j, \tag{5.1.4}$$

where we have assumed that $W_0 = V_0$. It follows that

$$V_k = \bigoplus_{j=0}^{k} W_j, \tag{5.1.5}$$

where all these subspaces are orthogonal. By virtue of (5.1.2) and (5.1.5), this implies

$$L^2(\Omega) = \bigoplus_{j=1}^{\infty} W_j,$$

which is a decomposition of $L^2(\Omega)$ into mutually orthogonal subspaces. A wavelet basis for $L^2(\Omega)$ can be constructed if one is able to find an orthonormal basis for each subspace W_j. Such a basis would be ideal in preconditioning the discrete problem (2.6) if it is computationally feasible.

In practice, it is very hard to find an L^2-orthonormal wavelet basis which is also computable. Therefore, orthogonality requirement in the decomposition (5.1.5) shall be relaxed to allow only a direct decomposition of the following form:

$$V_k = V_0 \oplus V_1^1 \oplus V_2^1 \oplus \ldots \oplus V_k^1, \tag{5.1.6}$$

where each V_j^1 is a complement of V_{j-1} in V_j such that the corresponding two–level decomposition is direct. But in order to attain the stability of the wavelet basis, it is crucial to have some approximate orthogonality among the subspaces V_j^1.

5.2 A general approach

A general method for deriving the hierarchical complement of each V_{j-1} in V_j is based on the existence of some computationally feasible projections π_j from $\mathcal{C}$, a dense subspace of the Hilbert space H, onto V_j. In particular, we assume that $\mathcal{C} \supset \cup_{l \geq l} V_l$ and $\pi_j \psi = \psi$ for any $\psi \in V_j$. Thus, one has $\pi_j \pi_i = \pi_i$ for $j \geq i$ if $V_i \subset V_j$. With $V_j^1 = (I - \pi_{j-1}) V_j$, one has the following two–level direct decomposition:

$$V_j = (I - \pi_{j-1}) V_j \oplus V_{j-1}. \tag{5.2.1}$$

Definition 2. (Multilevel Hierarchical Basis) *For $j = 0, 1, \ldots, k$, let $\{\phi_i^{(j)},\ i = 1, \ldots, n_j\}$ be a computationally feasible basis of V_j. Assume that $\{\phi_i^{(j-1)},\ i = 1, \ldots, n_{j-1}\} \bigcup \{\phi_i^{(j)},\ i = n_{j-1}+1, \ldots, n_j\}$ forms a basis of V_j. A multilevel hierarchical basis for V_k is defined as follows:*

$$\Phi_k \equiv \bigcup_{j=0}^{k} \{(I - \pi_{j-1}) \phi_i^{(j)},\ i = n_{j-1}+1, \ldots, n_j\}. \tag{5.2.2}$$

Here we have assumed that $\pi_{-1} = 0$ and $n_{-1} = 0$.

We now discuss the stability of the multilevel hierarchical basis. The following result sets a guideline for the selection of the operators π_j.

Theorem 4. *A necessary condition for Φ_k to be a stable Riesz basis of V_k is that the projection operators π_r be uniformly bounded on V_k with respect to r and k for any $r \leq k$.*

Proof: Let $v \in V_k$ be expanded as follows:

$$v = \sum_{i=1}^{n_r} c_i^{(r)} \hat{\phi}_i^{(r)} + \sum_{j=r+1}^{k} \sum_{i=n_{j-1}+1}^{n_j} c_i^{(k)} \bar{\phi}_i^{(k)}$$

where $\{\hat{\phi}_i^{(r)}\}_{i=1}^{n_r}$ is the multilevel hierarchical basis of V_r and $\bar{\phi}_i^{(s)} = (I - \pi_{s-1})\phi_i^{(s)}$. Introduce the notation

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}^{(r)} \\ \mathbf{c}_1^{(r+1)} \\ \vdots \\ \mathbf{c}_1^{(k)} \end{bmatrix},$$

where $\mathbf{c}_1^{(j)} = (c_{n_{j-1}+1}^{(j)}, \dots, c_{n_j}^{(j)})^T$ for $j = r+1, \dots, k$, and

$$\mathbf{c}^{(r)} = (c_1^{(r)}, \dots, c_{n_r}^{(r)})^T.$$

Since, by assumption, we have a stable Riesz basis, then there exist σ_1 and σ_2 independent of k such that

$$\sigma_1 \mathbf{c}^T \mathbf{c} \le \|v\|^2 \le \sigma_2 \mathbf{c}^T \mathbf{c}. \tag{5.2.3}$$

Observe that $\pi_r v = \sum_{i=1}^{n_r} c_i^{(r)} \hat{\phi}_i^{(r)}$. Thus, from (5.2.3) with v replaced by $\pi_r v$ we obtain

$$\|\pi_r v\|^2 \le \sigma_2 \mathbf{c}^{(r)^T} \mathbf{c}^{(r)} \le \sigma_2 \left(\mathbf{c}^{(r)^T} \mathbf{c}^{(r)} + \sum_{j=r+1}^{k} \mathbf{c}_1^{(j)^T} \mathbf{c}_1^{(j)} \right) \le \frac{\sigma_2}{\sigma_1} \|v\|^2.$$

Here we have used the lower bound of (5.2.3). This shows the boundedness of the projection operators π_r for any r.

§6 The Hierarchical basis

In this section we review the classical hierarchical basis decomposition. First, partition the domain Ω into large elements. Let $\mathcal{T}_0$ denote this initial coarse triangulation and V_0 be the corresponding finite element space of continuous piecewise linear functions. The fine finite element space $V = V_J$ corresponds to the triangulation $\mathcal{T}_J$ which is obtained by $J \ge 1$ successive refinements of the coarse triangulation $\mathcal{T}_0$. For problems of two space variables, one can use the triangular element and the refinement of one triangle at level $k-1$ will generate four congruent triangles of level k by connecting the midpoints of its edges. Similar techniques are available for problems of three space variables with tetrahedra being used as elements. We refer to Ong [29] for more details in this discussion. It is also possible to use bisection refinement for both two– and three–space problems. Details for this technique can be found from Mitchell [28] and Maubach [27].

Let $\mathcal{T}_i$ denote the finite element partition at level i, and V_i be the corresponding finite element space of continuous piecewise linear functions. Thus, we obtain a nested sequence of conforming finite element spaces

$$V_0 \subset V_1 \subset \ldots \subset V_J,$$

which can be used to discretize the second-order elliptic equation.

To describe the classical hierarchical basis, let $\mathcal{N}_i$ be the node set of nodal degrees of freedom at level i which consists of vertices of triangles (or tetrahedra in 3–d) in $\mathcal{T}_i$. One has the following natural direct decomposition:

$$\mathcal{N}_i = \mathcal{N}_i^1 \cup \mathcal{N}_{i-1},$$

with $\mathcal{N}_i^1$ being the set of newly-introduced nodal points. For example, in 2–d, the nodal set $\mathcal{N}_k^1$ contains the vertices of the triangles at level k that are midpoints of the edges of the triangles from level $k-1$. We also introduce the mesh size $h_i = 2^{-i} h_0$ for the ith level triangulation $\mathcal{T}_i$. Here, h_0 stands for the maximum diameter of the elements in $\mathcal{T}_0$. Recall that the standard nodal basis functions $\{\phi_i^{(k)},\ x_i \in \mathcal{N}_k\}$ are defined to satisfy the condition $\phi_i^{(k)}(x_j) = \delta_{i,j}$, where $\delta_{i,j}$ is the Kronecker symbol, with x_j running over all the nodes in $\mathcal{N}_k$. One can then define a two–level hierarchical basis of V_k by adding to the nodal basis of V_{k-1} the nodal basis functions of V_k corresponding to the complementary nodal set $\mathcal{N}_k^1 = \mathcal{N}_k \setminus \mathcal{N}_{k-1}$.

Consider now the nodal interpolation operator $I_k : C(\bar{\Omega}) \to V_k$ defined by $I_k v = \sum_{x_i \in \mathcal{N}_k} v(x_i) \phi_i^{(k)}$. It is clear that I_k is a projection and $I_k \psi = \psi$ if $\psi \in V_k$. With the choice of $\pi_j \equiv I_j$, one obtains the classical hierarchical basis by using the general decomposition method discussed in §5.

We comment briefly on the stability of the classical hierarchical basis in applications to the second order elliptic problems. In this application, the natural norm for the finite element space is the norm in the Sobolev space $H^1(\Omega)$. Since the interpolation operator I_k is not bounded in the H^1-norm (e.g., [13]), then Theorem 4 implies that the classical hierarchical basis is not absolutely stable. Thus, the resulting stiffness matrices computed with respect to the hierarchical basis will not be well–conditioned. However, as is well–known, the condition number for problems of two space variables is practically acceptable. In fact, the condition number for the second-order elliptic problems can be verified to be proportional to k^2 at level k. This condition number grows much slower than that from the standard nodal basis, which behaves like $\mathcal{O}(h_k^{-2})$. For problems of three space variables, the condition number for the classical hierarchical basis is of order $\mathcal{O}(h_k^{-1})$ which is better than the one from the standard nodal basis but worse than the application for problems in 2-d.

§7 A stable Riesz basis by wavelet method

In this section we will construct appropriate projections π_k which are H^1–stable and provide computationally feasible Riesz basis for V_J. The bilinear forms of main interest are those arising from the Hilbert space method for second-order elliptic problems discussed in §4. The method to be presented here was proposed by Vassilevski and Wang in [41].

7.1 On the basis construction

Define the L^2–projection operators $Q_k \,:\, L^2(\Omega) \to V_k$ as follows:

$$(Q_k v, \psi)_0 = (v, \psi)_0 \qquad \forall \psi \in V_k.$$

Also, assume that there are computationally feasible approximations $Q_k^a \,:\, L^2(\Omega) \to V_k$ of Q_k such that for some small tolerance $\tau > 0$ the following estimate holds:

$$\|(Q_k - Q_k^a)v\|_0 \le \tau \|Q_k v\| \qquad \forall v \in L^2(\Omega). \tag{7.1.1}$$

The projection operators of major interest are defined as follows:

$$\pi_k = \prod_{j=k}^{J-1} (I_j + Q_j^a(I_{j+1} - I_j)), \tag{7.1.2}$$

with $\pi_J = I$. It is clear that $\pi_k \psi = \psi$ if $\psi \in V_k$ since $I_j\psi + Q_j^a(I_{j+1} - I_j)\psi = I_j \psi = \psi$ for $j \ge k$ based on $(I_{j+1} - I_j)\psi = 0$ and $I_j \psi = \psi$. This also implies that $\pi_k^2 = \pi_k$.

Note that $\pi_{k-1}(I_k - I_{k-1})\phi = Q_{k-1}^a(I_k - I_{k-1})\phi$ and $\pi_k - \pi_{k-1} = (I - Q_{k-1}^a)(I_k - I_{k-1})\pi_k$. Then, the components in the definition (5.2.2) for the wavelet–like multilevel hierarchical basis read as follows:

$$\{\phi_i^{(0)}, i = 1, \ldots, n_0\} \bigcup_{j=1}^{k} \{(I - Q_{j-1}^a)\phi_i^{(j)},\ i = n_{j-1} + 1, \ldots, n_j\}. \tag{7.1.3}$$

The above components $\{(I - Q_{j-1}^a)\phi_i^{(j)},\ i = n_{j-1} + 1, \ldots, n_j\}$ can be seen as a modification of the classical hierarchical basis components based on the interpolation operator I_k since $(I - Q_{j-1}^a)\phi_i^{(j)} = (I - Q_{j-1}^a)(I_j -$

$I_{j-1})\phi_i^{(j)}$; the modification of the classical hierarchical basis components $\{(I_j - I_{j-1})\phi_i^{(j)},\ i = n_{j-1}+1,\dots,n_j\}$ comes from the additional term $Q_{j-1}^a(I_j - I_{j-1})\phi_i^{(j)}$. In other words, the modification was made by subtracting from each nodal hierarchical basis function $\phi_i^{(j)}$ its approximate L^2–projection $Q_{j-1}^a\phi_i^{(j)}$ onto the coarse level $j-1$. Such modifications of the hierarchical basis function $\phi_i^{(k)}$ for some particular choices of Q_{j-1}^a will be shown in Figures 1–3 in §8. It can be seen that the modified hierarchical basis functions are close relatives of the known Battle–Lemarié wavelets [17].

Observe that in the limit case of $Q_k^a = Q_k$ so that $\tau = 0$ in (7.1.1), we get

$$\pi_k v = Q_k I_{k+1} Q_{k+1} I_{k+2} \dots Q_{J-1} I_J v = Q_k Q_{k+1} \dots Q_{J-1} v = Q_k v.$$

Therefore, π_k reduces to the exact L^2–projection Q_k. As is well–known, the L^2–projection operators are bounded in both $H_0^1(\Omega)$ and $L^2(\Omega)$. This gives us a hope that the hierarchical multilevel basis corresponding to the above choice of the operators π_k may yield a stable Riesz basis if τ is sufficiently small.

7.2 Preliminary estimates

For an analysis of the multilevel basis (7.1.3), we need some auxiliary estimates already presented in Vassilevski and Wang [41]. The following result on estimating the error $e_j = (\pi_j - Q_j)v$ will play an important role in our analysis.

Lemma 1. *There exits an absolute constant C such that*

$$\sum_{j=1}^{k} h_j^{-2}\|e_j\|_0^2 \le C\tau^2 \sum_{j=1}^{k} h_j^{-2}\|(Q_j - Q_{j-1})v\|_0^2, \quad \forall v \in V_k. \qquad (7.2.1)$$

The estimate (7.2.1) relies on the following recursive relation:

$$e_{s-1} = (Q_{s-1} + R_{s-1})e_s + R_{s-1}(Q_s - Q_{s-1})v, \qquad (7.2.2)$$

where $R_{s-1} = (Q_{s-1} - Q_{s-1}^a)(I_{s-1} - I_s)$. It can be seen as follows:

$$\begin{aligned} e_{s-1} &= \pi_{s-1}v - Q_{s-1}v \\ &= (I_{s-1} + Q_{s-1}^a(I_s - I_{s-1}))\pi_s v - Q_{s-1}v \\ &= (Q_{s-1} - Q_{s-1}^a)I_{s-1}\pi_s v + Q_{s-1}^a \pi_s v - Q_{s-1}v. \end{aligned}$$

Thus, one has

$$\begin{aligned}
e_{s-1} &= (Q_{s-1} - Q^a_{s-1})I_{s-1}(\pi_s v - Q_s v) + Q^a_{s-1}(\pi_s v - Q_s v) \\
&\quad +(Q_{s-1} - Q^a_{s-1})I_{s-1}e_{s-1} + Q^a_{s-1}I_s e_s \\
&\quad +Q_{s-1}(I_{s-1}Q_s v - Q_s v) - Q^a_{s-1}(I_{s-1}Q_s v - Q_s v) \\
&= (Q_{s-1} - Q^a_{s-1})(I_{s-1} - I_s)e_s + (Q_{s-1} - Q^a_{s-1})e_s + Q^a_{s-1}e_s \\
&\quad +(Q_{s-1} - Q^a_{s-1})(I_{s-1} - I_s)Q_s v \\
&= \left[Q_{s-1} + (Q_{s-1} - Q^a_{s-1})(I_{s-1} - I_s)\right] e_s \\
&\quad +(Q_{s-1} - Q^a_{s-1})(I_{s-1} - I_s)Q_s v\cdot
\end{aligned}$$

The latter together with the fact that $(I_{s-1} - I_s)Q_{s-1} = 0$ implies the desired recursive relation (7.2.2).

Proof: We now verify Lemma 1. Let C_R be a mesh-independent upper bound of the L^2-norm for the operator $I_s - I_{s-1} : V_s \to V_s$. Then,

$$\|R_{s-1}v\|_0 \le C_R\tau\|v\|_0 \quad \text{for all } v \in V_s\cdot \tag{7.2.3}$$

We assume that,

$$C_R\tau \le q_0 = \text{Const } < 1\cdot \tag{7.2.4}$$

Then,

$$(1 + C_R\tau)\frac{1}{2} \le q = \frac{1+q_0}{2} = \text{Const } < 1\cdot \tag{7.2.5}$$

Next, observe that $e_k = 0$ (since $v \in V_k$). Then a recursive use of (7.2.2) leads to

$$\begin{aligned}
\|e_{s-1}\|_0 &\le (1 + C_R\tau)\|e_s\|_0 + C_R\tau\|(Q_s - Q_{s-1})v\|_0 \\
&\le C_R\tau \sum_{j=s}^{k} (1 + C_R\tau)^{j-s}\|(Q_j - Q_{j-1})v\|_0\cdot
\end{aligned}$$

Therefore, with $h_j = 2^{-j} h_0$ and h_0 being the coarsest mesh size,

$$\begin{aligned}
\|e_{s-1}\|_0 &\leq C_R \tau h_{s-1} \sum_{j=s}^{k} (1 + C_R\tau)^{j-s} h_s^{-1} \|(Q_j - Q_{j-1})v\|_0 \\
&= C_R \tau h_{s-1} \sum_{j=s}^{k} (1 + C_R\tau)^{j-s} h_s^{-1} h_j h_j^{-1} \|(Q_j - Q_{j-1})v\|_0 \\
&= C_R \tau h_{s-1} \sum_{j=s}^{k} (1 + C_R\tau)^{j-s} \left(\tfrac{1}{2}\right)^{j-s} h_j^{-1} \|(Q_j - Q_{j-1})v\|_0 \\
&\leq C_R \tau h_{s-1} \sum_{j=s}^{k} q^{j-s} h_j^{-1} \|(Q_j - Q_{j-1})v\|_0 \\
&\leq C_R \tau h_{s-1} \tfrac{1}{\sqrt{1-q}} \left[\sum_{j=s}^{k} q^{j-s} h_j^{-2} \|(Q_j - Q_{j-1})v\|_0^2 \right]^{\frac{1}{2}}.
\end{aligned}$$

The latter inequality shows

$$\begin{aligned}
\sum_{s=1}^{k} h_{s-1}^{-2} \|e_{s-1}\|_0^2 &\leq C_R^2 \tau^2 \tfrac{1}{1-q} \sum_{s=1}^{k} \sum_{j=s}^{k} q^{j-s} h_j^{-2} \|(Q_j - Q_{j-1})v\|_0^2 \\
&\leq C_R^2 \tau^2 \tfrac{1}{(1-q)^2} \sum_{j=1}^{k} h_j^{-2} \|(Q_j - Q_{j-1})v\|_0^2 \cdot
\end{aligned}$$

which proves the lemma.

The above proof also shows the following corollary.

Corollary 1. *For any $\sigma \in (0, 1]$, the following estimate holds,*

$$\sum_{s=1}^{k} h_{s-1}^{-2\sigma} \|e_{s-1}\|_0^2 \leq C_R^2 \tau^2 \frac{1}{(1-q)^2} \sum_{j=1}^{k} h_j^{-2\sigma} \|(Q_j - Q_{j-1})v\|_0^2,$$

provided that τ satisfies the estimate

$$C_R \tau \leq q 2^{\sigma} - 1, \tag{7.2.6}$$

for a mesh–independent constant $q \in (0, 1)$ (actually $q > 2^{-\sigma}$).

Remark 2: Corollary 1 indicates that in order to have the L^2–stability of the deviations one has to assume a level dependence on the tolerance τ. More precisely, there exists a $\tau_0 > 0$ such that if $\tau < \tau_0 J^{-1}$, then

$$\sum_{s=1}^{k} \|e_{s-1}\|_0^2 \leq C \|v\|_0^2 \qquad \text{for all } v \in V_k \cdot \tag{7.2.7}$$

Lemma 2. *Let $V_k^1 = (I - M_{k-1})V_k^{(1)}$, with $V_k^{(1)} \equiv (I_k - I_{k-1})V_k$, be the modified hierarchical subspace of level k for any given L^2–bounded operator M_{k-1}. Then, there are positive constants c_1 and c_2 independent of k such that*

$$c_1\|\phi^1\|_r^2 \le \|\psi^1\|_r^2 \le c_2\|\phi^1\|_r^2, \qquad r = 0, 1, \tag{7.2.8}$$

for any $\psi^1 = (I - M_{k-1})\phi^1 \in V_k^1$, $\phi^1 \in V_k^{(1)}$. Here $\|\cdot\|_1$ stands for the norm in the Sobolev space $H_0^1(\Omega)$ and $\|\cdot\|_0$ denotes the $L^2(\Omega)$–norm.

Proof: The following strengthened Cauchy inequality is valuable: There exists a constant $\gamma \in (0,1)$, independent of the mesh size or the level index k such that

$$(\nabla\phi^1, \nabla\tilde{\phi}) \le \gamma(\nabla\phi^1, \nabla\phi^1)^{\frac{1}{2}}\ (\nabla\tilde{\phi}, \nabla\tilde{\phi})^{\frac{1}{2}}, \tag{7.2.9}$$

for all $\phi^1 \in V_k^{(1)}$ and $\tilde{\phi} \in V_{k-1}$. In fact, we shall make use of the following version of (7.2.9). For any $\phi^1 \in V_k^{(1)}$ and $\tilde{\phi} \in V_{k-1}$, one has

$$(\nabla(\phi^1 + \tilde{\phi}), \nabla(\phi^1 + \tilde{\phi})) \ge (1 - \gamma^2)(\nabla\phi^1, \nabla\phi^1). \tag{7.2.10}$$

A derivation of (7.2.9) and (7.2.10) can be found from Bank and Dupont [5] or Axelsson and Gustafsson [3].

We first establish (7.2.8) for the case $r = 1$. With $\tilde{\phi} = -M_{k-1}\phi^1$ we see from (7.2.10) that

$$(1 - \gamma^2)\|\phi^1\|_1^2 \le \|\psi^1\|_1^2.$$

Thus, the inequality on the left-hand side of (7.2.8) is valid with $c_1 = 1 - \gamma^2$. To derive the part on the right-hand side, we use the standard inverse inequality to obtain

$$\|\psi^1\|_1^2 \le Ch_k^{-2}\|\psi^1\|_0^2 \le Ch_k^{-2}\|\phi^1\|_0^2,$$

where we have used the L^2–boundedness of the linear operator M_{k-1}. Observe now that since $\phi^1 \in V_k^{(1)}$, there exists a constant C such that

$$\|\phi^1\|_0^2 \le Ch_k^2\|\phi^1\|_1^2. \tag{7.2.11}$$

It follows that $\|\psi^1\|_1^2 \le C\|\phi^1\|_1^2$ for some constant C. This completes the proof of (7.2.8) for $r = 1$. Similar arguments can be applied to verify the case $r = 0$.

Lemma 3. *For any $\psi^1 = (I - M_{k-1})\phi^1 \in V_k^1$ and $\varphi^1 = (I - M_{k-1})\chi^1 \in V_k^1$, with $\phi^1, \chi^1 \in V_k^{(1)} = (I_k - I_{k-1})V_k$, define*

$$(A_{11}^{(k)}\psi^1, \varphi^1) = a(\psi^1, \varphi^1) = \int_\Omega a(x)\nabla\psi^1 \cdot \nabla\varphi^1.$$

Here, the bilinear form $a(\cdot,\cdot)$ is equivalent to the H_0^1–inner product. Then there are positive constants τ_i such that

$$\tau_1 h_k^{-2}\|\phi^1\|_0^2 \le (A_{11}^{(k)}\psi^1, \psi^1) \le \tau_2 h_k^{-2}\|\phi^1\|_0^2.$$

Proof: Since $a(.,.)$ is equivalent to the H_0^1–inner product, there are two positive constants $\tilde{\tau}_1$ and $\tilde{\tau}_2$ such that

$$\tilde{\tau}_1\|\psi^1\|_1^2 \le (A_{11}^{(k)}\psi^1, \psi^1) \le \tilde{\tau}_2\|\psi^1\|_1^2.$$

Using the norm equivalence (7.2.8), estimate (7.2.11), and the inverse inequality, we obtain with possibly different constants τ_1 and τ_2,

$$\tau_1 h_k^{-2}\|\phi^1\|_0^2 \le (A_{11}^{(k)}\psi^1, \psi^1) \le \tau_2 h_k^{-2}\|\phi^1\|_0^2. \tag{7.2.12}$$

The above inequalities conclude the lemma.

Lemma 4. *Given v and let $v^{(k)^1} = (\pi_k - \pi_{k-1})v$. There exists a sufficiently small constant $\tau_0 > 0$ such that if the approximate projections Q_k^a satisfy (7.1.1) with $\tau \in (0, \tau_0)$ (see (7.2.4) and (7.2.5)), then*

$$\|v\|_1^2 \simeq \sum_{k=0}^{J} h_k^{-2}\|v^{(k)^1}\|_0^2. \tag{7.2.13}$$

Proof: Let $v \in V$. Starting with $v^{(J)} = v$, for $s = J$ down to 1, one defines $v^{(s-1)} = (I_{s-1} + Q_{s-1}^a(I_s - I_{s-1})v^{(s)} = \pi_{s-1}v$. Then the decomposition v in terms of entries in $V_s^1 = (I - Q_{s-1}^a)V_s^{(1)} = (I - Q_{s-1}^a)(I_s - I_{s-1})V$ reads as,

$$v = v^{(0)} + \sum_{j=1}^{J} v^{(j)^1}, \quad v^{(s)^1} = v^{(s)} - v^{(s-1)}. \tag{7.2.14}$$

From the representation $v^{(s)^1} = v^{(s)} - v^{(s-1)} = (Q_s - Q_{s-1}) + e_s - e_{s-1}$ one arrives at the estimate,

$$\begin{aligned}\sum_{k=0}^{J} h_k^{-2}\|v^{(k)^1}\|_0^2 &\le C\sum_{k=0}^{J} h_k^{-2}\|(Q_k - Q_{k-1})v\|_0^2 + C\sum_{k=0}^{J} h_k^{-2}\|e_k\|_0^2 \\ &\le C(1+\tau^2)\sum_{k=0}^{J} h_k^{-2}\|(Q_k - Q_{k-1})v\|_0^2 \\ &\le C\|v\|_1^2.\end{aligned}$$

Here we have used the norm equivalence result of Oswald [30] and Lemma 1.

Thus, it suffices to establish the upper bound:

$$\|v\|_1^2 \le C \sum_{k=0}^{J} h_k^{-2} \|v^{(k)^1}\|_0^2. \tag{7.2.15}$$

It is possible to give a direct proof for (7.2.15) by using the strengthened Cauchy–Schwarz inequality (see Vassilevski and Wang [41]). Here we would like to adopt an alternative approach by using the following characterization of the H_0^1–norm for finite element functions:

$$\|v\|_1^2 \simeq \inf_{v=\sum_{k=0}^{J} v_k,\ v_k \in V_k} \sum_{k=0}^{J} h_k^{-2} \|v_k\|_0^2. \tag{7.2.16}$$

A proof of the above equivalence can be found from [30]. Thus, for the particular decomposition $v = \sum_{k=0}^{J} v^{(k)^1}$, $v^{(k)^1} \in V_k$, one immediately has

$$\|v\|_1^2 \le C \sum_{k=0}^{J} h_k^{-2} \|v^{(k)^1}\|_0^2,$$

which completes the proof of the lemma.

7.3 Stability analysis

Here we study the Riesz property of the wavelet–like multilevel hierarchical basis defined in (7.1.3). For any $v \in V$ let

$$v = \sum_{x_i \in \mathcal{N}_0} c_{0,i} \phi_i^{(0)} + \sum_{k=1}^{J} \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i} (I - Q_{k-1}^a) \phi_i^{(k)} \tag{7.3.1}$$

be its representation with respect to the given wavelet basis. The corresponding coefficient norm of v is given by

$$|||v||| = \left(h_0^{d-2} \sum_{x_i \in \mathcal{N}_0} c_{0,i}^2 + \sum_{k=1}^{J} h_k^{d-2} \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}^2 \right)^{1/2}, \tag{7.3.2}$$

where $d = 2$ or 3 according to the number of space variables. Our main result in this section is the following norm equivalence:

Theorem 5. *There exists a small (but fixed) $\tau_0 > 0$ such that if the approximate projections Q_k^a satisfy (7.1.1) with $\tau \in (0, \tau_0)$, then there are positive constants c_1 and c_2 satisfying*

$$c_1 |||v|||^2 \leq \|v\|_1^2 \leq c_2 |||v|||^2 \qquad \forall v \in V. \tag{7.3.3}$$

In other words, the modified hierarchical basis is a stable Riesz basis for the second order elliptic and Stokes problems. The equivalence relation (7.3.3) shall be abbreviated as $|||v|||^2 \simeq \|v\|_1^2$.

Proof: We first rewrite (7.3.1) as follows:

$$v = \sum_{k=0}^{J} v^{(k)^1}, \tag{7.3.4}$$

where, with $Q_{-1}^a = 0$,

$$v^{(k)^1} = \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}(I - Q_{k-1}^a)\phi_i^{(k)} \in V_k^1. \tag{7.3.5}$$

Furthermore, by letting $\phi^{(k)} = \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}\phi_i^{(k)} \in V_k^{(1)}$ we see that $v^{(k)^1} = (I - Q_{k-1}^a)\phi^{(k)}$. Thus, by using (7.2.8) in Lemma 2 (with $r = 0$ and $M_{k-1} = Q_{k-1}^a$) we obtain

$$\|\phi^{(k)}\|_0^2 \simeq \|v^{(k)^1}\|_0^2. \tag{7.3.6}$$

Since $\phi^{(k)} \in V_k^{(1)}$, then

$$\|\phi^{(k)}\|_0^2 \simeq h_k^d \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}^2.$$

Combining the above with (7.3.6) yields

$$|||v|||^2 \simeq \sum_{k=0}^{J} h_k^{-2} \|v^{(k)^1}\|_0^2.$$

This, together with lemma 4, completes the proof of the theorem.

Remark 3: For any fixed $\sigma \in (0, 1]$, define

$$|||v|||_\sigma = \left(h_0^{d-2\sigma} \sum_{x_i \in \mathcal{N}_0} c_{0,i}^2 + \sum_{k=1}^{J} h_k^{d-2\sigma} \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}^2 \right)^{1/2},$$

with $d = 2$ or 3. Then, for sufficiently small τ, one has

$$|||v|||_\sigma \simeq \|v\|_\sigma \tag{7.3.7}$$

for all the finite element functions $v \in V_J$. Here $\|\cdot\|_\sigma$ denotes the $H_0^\sigma(\Omega)$–norm defined by interpolating $H_0^1(\Omega)$ with $L^2(\Omega)$. The constants in the norm equivalence depend on σ as indicated by Corollary 1. For $\sigma = 0$, the equivalence (7.3.7) holds true provided that the tolerance satisfies $\tau < \tau_0 J^{-1}$ for some small τ_0.

§8 Implementations for a model problem

In this section we discuss techniques which implement the wavelet–like multilevel hierarchical basis for approximate solutions of PDEs. For simplicity, we consider the selfadjoint second-order elliptic equation discussed in §4. The Stokes equation can be covered in a similar manner.

The bilinear form under consideration is defined as follows:

$$a(\varphi, \psi) = \int_\Omega a(x)\nabla\varphi \cdot \nabla\psi, \quad \forall\varphi,\ \psi \in H_0^1(\Omega). \tag{8.1}$$

Let $d(\varphi, \psi) \equiv \int_\Omega \nabla\varphi \cdot \nabla\psi$ be the Dirichlet form defined on $H_0^1(\Omega) \times H_0^1(\Omega)$. Since the two bilinear forms $a(\cdot, \cdot)$ and $d(\cdot, \cdot)$ are equivalent, then we have from (7.3.3) that

$$c_1 |||v|||^2 \le a(v, v) \le c_2 |||v|||^2 \tag{8.2}$$

for some positive constants c_1 and c_2. Moreover, the following result holds:

Lemma 5. *Let $\epsilon > 0$ be a parameter and*

$$a_\epsilon(\varphi, \psi) = \epsilon \int_\Omega a(x)\nabla\varphi \cdot \nabla\psi + \int_\Omega \varphi\, \psi, \quad \forall\varphi,\ \psi \in H_0^1(\Omega). \tag{8.3}$$

If the approximation in (7.1.1) is sufficiently accurate such that $\tau \le \tau_0 J^{-1}$ for some constant τ_0, then there exist τ_1 and τ_2 independent of ϵ and the mesh size h_k such that

$$\tau_1 \sum_{k=0}^{J} h_k^d \alpha_k \sum_{x_i \in \mathcal{N}_k^1} c_{k,i}^2 \le a_\epsilon(v, v) \le \tau_2 \sum_{k=0}^{J} h_k^d \alpha_k \sum_{x_i \in \mathcal{N}_k^1} c_{k,i}^2 \tag{8.4}$$

for any finite element function $v \in V_J$. Here $\alpha_k = \epsilon h_k^{-2} + 1$ and $c_{k,i}$ are the coefficients of v in the expansion (7.3.1).

The spectral bounds (8.4) can be used for the bilinear form arising from discretizing time–dependent Stokes problems. The appearance of ϵ is due to the time stepping parameter Δt. Similar bilinear forms can be obtained for the pressure unknown by eliminating the vector unknown $\mathbf{u}$ for the steady–state Stokes equation. For more details, we refer to Bramble and Pasciak [10].

We remark that the bilinear form $a_\epsilon(\cdot,\cdot)$ is equivalent to the Dirichlet form $d(\cdot,\cdot)$ if the parameter ϵ is bounded away from zero by a fixed positive constant C_0 (i.e., $\epsilon \geq C_0$). In this case, the equivalence (8.4) holds true under the condition of Theorem 5.

In practical computations, one has two alternatives in solving the discrete problem (2.6) with the wavelet–like basis presented in previous sections. The first one makes use of the explicit form of the basis functions $(I - Q^a_{j-1})\phi^{(j)}_i$ to assemble the corresponding stiffness matrix. The second one uses the stiffness matrix assembled from using the standard nodal basis and then performs change of basis in the process of iterations. The first approach has a difficulty in that the assembly of the global stiffness matrix is no longer local (element–wise). In general, one is recommended to adopt the second approach because the corresponding stiffness matrix is much easier to assemble. In this case, the wavelet–like hierarchical basis actually provides a preconditioning technique for solving the discrete problem (2.6). The rest of this section will describe some preconditioning procedures for the model problem. A more detailed discussion can be found from [41] and [42].

8.1 Preconditioners

Here we outline two preconditioners for the elliptic operator $A^{(k)} : V_k \to V_k$ arising from the bilinear form $a(\cdot,\cdot)$. The preconditioners will be constructed by using the following wavelet–like multilevel hierarchical decomposition of V_k:

$$V_k = V_0 \oplus V^1_1 \oplus V^1_2 \oplus \ldots \oplus V^1_k,$$

where $V^1_j = (I - Q^a_{j-1})(I_j - I_{j-1})V$. The preconditioners (to be defined below) shall be called AWM–HB (Approximate-Wavelet Modified Hierarchical Basis) preconditioners.

The following operators are needed in the construction of the AWM–HB preconditioners:

• In each coordinate space V^1_k, there exists a discretization operator $A^{(k)}_{11} : V^1_k \to V^1_k$ as the restriction of $A^{(k)}$ onto the subspace V^1_k defined by

$$(A^{(k)}_{11}\psi^1, \varphi^1) = a(\psi^1, \varphi^1) \qquad \forall \varphi^1,\ \psi^1 \in V^1_k. \tag{8.1.1}$$

- Similarly, we define $A_{12}^{(k)}: V_{k-1} \to V_k^1$ and $A_{21}^{(k)}: V_k^1 \to V_{k-1}$ by

$$(A_{12}^{(k)}\widetilde{\psi}, \varphi^1) = (\widetilde{\psi}, A_{21}^{(k)}\varphi^1) = a(\varphi^1, \widetilde{\psi}) \qquad \forall \widetilde{\psi} \in V_{k-1}, \varphi^1 \in V_k^1. \quad (8.1.2)$$

Thus, the operator $A^{(k)}$ naturally admits the following partition:

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A^{(k-1)} \end{bmatrix} \begin{matrix} \} & V_k^1 \\ \} & V_{k-1} \end{matrix} . \qquad (8.1.3)$$

Let $B_{11}^{(k)}$ be given approximations (symmetric positive definite operators) to $A_{11}^{(k)}$ such that for some positive constant b_1 the following holds:

$$(A_{11}^{(k)}\varphi^1, \varphi^1) \le (B_{11}^{(k)}\varphi^1, \varphi^1) \le (1 + b_1)(A_{11}^{(k)}\varphi^1, \varphi^1), \quad \forall \varphi^1 \in V_k^1. \quad (8.1.4)$$

Let $A = A^{(J)}$ be the discretization operator for which a preconditioner is necessary in practical computation. The following explains how the preconditioners can be constructed based on the block structure of $A^{(k)}$ in (8.1.3).

Definition 3. (Multiplicative AWM–HB preconditioners) *Let $B = B^{(J)}$ be the multiplicative AWM–HB preconditioner of A. It is defined as follows:*

- *Set $B^{(0)} = A^{(0)}$.*
- *For $k = 1, \cdots, J$, set*

$$B^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ A_{21}^{(k)} & B^{(k-1)} \end{bmatrix} \begin{bmatrix} I & B_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & I \end{bmatrix}.$$

Definition 4. (Additive AWM–HB preconditioners)
Let $D = D^{(J)}$ be the additive AWM–HB preconditioner of A. It is defined as follows:

- *Set $D^{(0)} = A^{(0)}$.*
- *For $k = 1, \cdots, J$, set*

$$D^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ 0 & D^{(k-1)} \end{bmatrix}.$$

8.2 Main results for the AWM–HB preconditioners

A spectral equivalence between A and its preconditioners B and D has been established in [41]. The result can be stated as follows. If the tolerance in (7.1.1) is sufficiently small such that $\tau \leq \tau_0$ for some small τ_0, then there are two absolute constants c_1 and c_2 satisfying

$$c_1(Sv, v) \leq (Av, v) \leq c_2(Sv, v) \qquad \forall v \in V_J, \tag{8.2.1}$$

where $S = B^{(J)}$ or $D^{(J)}$. The estimate (8.2.1) is based on the following results:

(A) There exists a constant $\sigma_N > 0$ such that

$$|Q_0 v|_1^2 + \sum_{s=1}^{J} 2^{2s} \|(Q_s - Q_{s-1})v\|_0^2 \leq \sigma_N \|v\|_1^2 \qquad \forall v \in V. \tag{8.2.2}$$

(B) There exist constants $\sigma_I > 0$ and $\delta \in (0,1)$ (in fact, if $h_i = \frac{1}{2}h_{i-1}$, then $\delta = \frac{1}{\sqrt{2}}$) such that the following strengthened Cauchy–Schwarz inequality holds for any $i \leq j$:

$$a(\varphi_i, \varphi_j)^2 \leq \sigma_I \delta^{2(j-i)} a(\varphi_i, \varphi_i) \lambda_j \|\varphi_j\|_0^2 \qquad \forall \varphi_i \in V_i,\ \varphi_j \in V_j. \tag{8.2.3}$$

Here $\lambda_j = O(h_j^{-2})$ is the largest eigenvalue of the operator $A^{(j)}$.

The inequalities in (A) and (B) have been respectively verified by Oswald [30] and Yserentant [44], [45].

One important feature in the partition (8.1.3) is that the block $A_{11}^{(k)}$ is well–conditioned; this can be seen from Lemma 3. In particular, the block $A_{11}^{(k)}$ is spectrally equivalent to its diagonal part. Thus, the Jacobi preconditioner would be a good choice for $B_{11}^{(k)}$ (see (8.1.4)) in approximating $A_{11}^{(k)}$.

Remark 4: If one does not assume the strengthened Cauchy–Schwarz inequality (B), then the estimate (8.2.1) for $S = B$ still holds with constants $c_1 = O(1)$ and $c_2 = O(\log_2 2J)$. In the case $S = D$, the condition (B) is not required in the equivalence (8.2.1). See Griebel and Oswald [19] for related details.

8.3 On the approximate L^2-projection

Denote by Q_{k-1}^a the approximate L^2-projections onto the subspace V_{k-1}. We begin with describing algorithms for computing the action of Q_{k-1}^a.

For any $v \in V_k^{(1)}$, let $\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ 0 \end{bmatrix} \begin{matrix} \}\mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \}\mathcal{N}_{k-1} \end{matrix}$ be its coefficient vector with respect to the standard nodal basis of V_k; the second block-component of $\mathbf{v}$ is zero since v vanishes on $\mathcal{N}_{k-1}$. The operator Q^a_{k-1} can be designed by approximately solving the following equation:

$$(Q_{k-1}v, w) = (v, w), \qquad \forall w \in V_{k-1}. \tag{8.3.1}$$

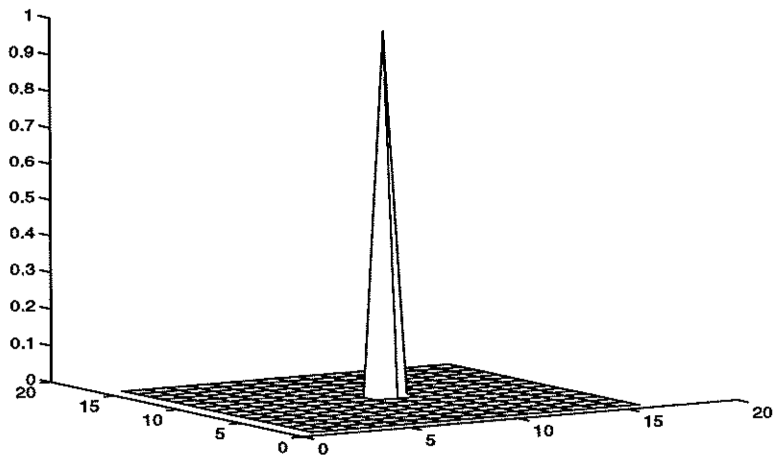

Figure 1. Plot of a HB function (no modification)

Let $I^k_{k-1} = \begin{bmatrix} J_{12} \\ I \end{bmatrix} \begin{matrix} \}\mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \}\mathcal{N}_{k-1} \end{matrix}$ (with the abbreviation $J_{12} = J^{(k)}_{12}$) and $I^{k-1}_k = I^{k^T}_{k-1}$ be the natural coarse-to-fine, and respectively, fine-to-coarse transformation matrices. For example, if the nodal basis coefficient vector of a function $v_2 \in V_{k-1}$ in terms of the nodal basis of V_{k-1} is $\mathbf{v}_2$, then its coefficient vector with respect to the nodal basis of V_k (note that $v_2 \in V_{k-1} \subset V_k$) will be $I^k_{k-1}\mathbf{v}_2 = \begin{bmatrix} J_{12}\mathbf{v}_2 \\ \mathbf{v}_2 \end{bmatrix} \begin{matrix} \}\mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \}\mathcal{N}_{k-1} \end{matrix}$.

Denote now by $G_k = \{(\phi^{(k)}_j, \phi^{(k)}_i)\}_{x_j, x_i \in \mathcal{N}_k}$ the mass (or Gram) matrix at the k-th level. Then (8.3.1) admits the following matrix–vector form:

$$\mathbf{w}_2^T G_{k-1}\mathbf{v}_2 = (I^k_{k-1}\mathbf{w}_2)^T G_k \mathbf{v}, \qquad \forall \mathbf{w}_2.$$

Here $\mathbf{v}_2$ and $\mathbf{w}_2$ are the nodal coefficient vectors of $Q_{k-1}v$ and $w \in V_{k-1}$ at the $(k-1)$th level respectively. Therefore, one needs to solve the following mass matrix problem:

$$G_{k-1}\mathbf{v}_2 = I^{k-1}_k G_k \mathbf{v}. \tag{8.3.2}$$

In other words, the exact L^2-projection $Q_{k-1}v$ is actually given by

$$G_{k-1}^{-1} I_k^{k-1} G_k \mathbf{v}.$$

Hence

$$\begin{aligned} \|Q_{k-1}v\|_0^2 &= \left(G_{k-1}^{-1} I_k^{k-1} G_k \mathbf{v}\right)^T G_{k-1} \left(G_{k-1}^{-1} I_k^{k-1} G_k \mathbf{v}\right) \\ &= \|G_{k-1}^{-\frac{1}{2}} I_k^{k-1} G_k \mathbf{v}\|^2. \end{aligned} \tag{8.3.3}$$

We used here the notation $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$.

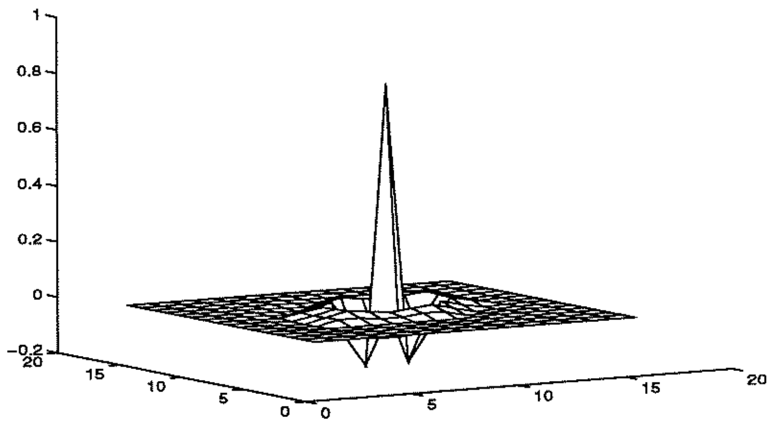

Figure 2. Plot of a wavelet-modified HB function; $m = 2$

To have a computationally feasible basis, we replace G_{k-1}^{-1} by some approximations $\widetilde{G}_{k-1}^{-1}$ whose action can be computed by simple iterative methods applied to (8.3.2). Such iterative methods lead to following polynomial approximations of G_{k-1}^{-1},

$$\widetilde{G}_{k-1}^{-1} = \left[I - \pi_m\left(G_{k-1}\right)\right] G_{k-1}^{-1},$$

where π_m is a polynomial of degree $m \geq 1$. The polynomial π_m also satisfies $\pi_m(0) = 1$ and $0 \leq \pi_m(t) < 1$ for $t \in [\alpha, \beta]$, where the latter interval contains the spectrum of the mass matrix G_{k-1}. Since G_{k-1} is well-conditioned, one can choose the interval $[\alpha, \beta]$ independent of k. Thus, the polynomial degree m can be chosen to be mesh-independent so that a given prescribed accuracy $\tau > 0$ in (7.1.1) is guaranteed. More precisely, given a tolerance $\tau > 0$, one can choose $m = m(\tau)$ satisfying

$$\begin{aligned} \|Q^a_{k-1}v - Q_{k-1}v\|_0 &= \|G^{\frac{1}{2}}_{k-1}\left(G^{-1}_{k-1} - \widetilde{G}^{-1}_{k-1}\right) I^{k-1}_k G_k \mathbf{v}\| \\ &= \|G^{\frac{1}{2}}_{k-1}\pi_m\left(G_{k-1}\right) G^{-1}_{k-1} I^{k-1}_k G_k \mathbf{v}\| \\ &\le \max_{t\in[\alpha,\beta]} \pi_m(t) \|G^{-\frac{1}{2}}_{k-1} I^{k-1}_k G_k \mathbf{v}\| \\ &= \max_{t\in[\alpha,\beta]} \pi_m(t) \|Q_{k-1}v\|_0. \end{aligned}$$

Here we have used identity (8.3.2) and the properties of π_m. The last estimate implies the validity of (7.1.1) with

$$\tau \ge \max_{t\in[\alpha,\beta]} \pi_m(t).$$

A simple choice of $\pi_m(t)$ is the truncated series

$$(1 - \pi_m(t))t^{-1} = p_{m-1}(t) \equiv \beta^{-1} \sum_{k=0}^{m-1} (1 - \frac{1}{\beta} t)^k, \tag{8.3.4}$$

which yields $\widetilde{G}^{-1}_{k-1} = p_{m-1}(G_{k-1})$. We remark that (8.3.4) was obtained from the following expansion:

$$1 = t\beta^{-1} \sum_{k=0}^{\infty} (1 - t\beta^{-1})^k, \qquad t \in [\alpha, \beta].$$

With the above choice on the polynomial $\pi_m(t)$, we have

$$\pi_m(t) = 1 - tp_{m-1}(t) = t\beta^{-1} \sum_{k\ge m} (1 - \beta^{-1}t)^k = (1 - \beta^{-1}t)^m.$$

It follows that

$$\max_{t\in[\alpha,\beta]} \pi_m(t) = \left(1 - \frac{\alpha}{\beta}\right)^m.$$

In general, by a careful selection on π_m we have $\max_{t\in[\alpha,\beta]} \pi_m(t) \le Cq^m$ for some constants $C > 0$ and $q \in (0, 1)$, both independent of k. Since the restriction on τ was that τ be sufficiently small, then one must have

$$m = O(\log \tau^{-1}). \tag{8.3.5}$$

The requirement (8.3.5) obviously imposes a very mild restriction on m. In practice, one expects to use reasonably small m (e.g., $m = 1, 2$). This observation is confirmed by our numerical experiments performed in Vassilevski and Wang [42]. We show in Figure 1 a typical plot of a nodal basis function of $V_k^{(1)}$ and its approximate-wavelet modification for $m = 2$ in Figure 2. The conjugate gradient method was employed to provide polynomial approximations for the solution of the mass–matrix problem (8.3.2).

8.4 Matrix formulations of the AWM–HB preconditioners

We now turn to the description of the multiplicative and additive AWM–HB methods in a matrix–vector form. Let us first derive matrix representations for the operators $A_{11}^{(k)}$, $A_{12}^{(k)}$, and $A_{21}^{(k)}$ introduced in (8.1.1) and (8.1.2). In what follows of this section, capital letters without overhats will denote matrices corresponding to the standard nodal basis of the underlined finite element space. For example, $A^{(k)}$ denotes the standard nodal basis stiffness matrix with entries $\{a(\phi_i^{(k)}, \phi_j^{(k)})\}_{x_i, x_j \in \mathcal{N}_k}$.

For any $v \in V_k$ and its nodal coefficient vector $\mathbf{v}$, we decompose v as follows:

$$v = (I - Q_{k-1}^a)(I_k - I_{k-1})v + w_2,$$

where $w_2 \in V_{k-1}$ is uniquely determined as $w_2 = I_{k-1}v + Q_{k-1}^a(I_k - I_{k-1})v$. Our goal is to find a vector representation for components of v. Since the above decomposition is direct, it is clear that there are vectors $\widehat{\mathbf{v}}_1$ and $\widehat{\mathbf{v}}_2$ satisfying

$$\mathbf{v} = (I - I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1} G_k) \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ 0 \end{bmatrix} \begin{matrix} \}\mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \}\mathcal{N}_{k-1} \end{matrix} + I_{k-1}^k \widehat{\mathbf{v}}_2. \tag{8.4.1}$$

The vectors $\widehat{\mathbf{v}}_1$ and $\widehat{\mathbf{v}}_2$ represent the components of our wavelet-modified two–level HB coefficient vector $\widehat{\mathbf{v}} = \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \widehat{\mathbf{v}}_2 \end{bmatrix}$ of v.

Now, consider the following problem

$$A\mathbf{v} = \mathbf{d}, \tag{8.4.2}$$

which is in the standard nodal basis matrix–vector form. We transform it into the approximate wavelet modified two–level HB by testing (8.4.2) with the two components $(I - I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1} G_k) \begin{bmatrix} \widehat{\mathbf{w}}_1 \\ 0 \end{bmatrix}$ and $I_{k-1}^k \widehat{\mathbf{w}}_2$ for arbitrary $\widehat{\mathbf{w}}_1$ and $\widehat{\mathbf{w}}_2$. By doing so, we get the following two–by–two block system for the approximate wavelet modified two–level HB components of $\widehat{\mathbf{v}}$ (denoted by $\widehat{\mathbf{v}}_1$ and $\widehat{\mathbf{v}}_2$),

$$\begin{bmatrix} \widehat{A}_{11}^{(k)} & \widehat{A}_{12}^{(k)} \\ \widehat{A}_{21}^{(k)} & \widehat{A}_{22}^{(k)} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \widehat{\mathbf{v}}_2 \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{d}}_1 \\ \widehat{\mathbf{d}}_2 \end{bmatrix}, \tag{8.4.3}$$

where

$$\widehat{A}_{11}^{(k)} = [I \quad 0] \left(I - G_k I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1}\right) A^{(k)} \left(I - I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1} G_k\right) \begin{bmatrix} I \\ 0 \end{bmatrix};$$
$$\widehat{A}_{12}^{(k)} = [I \quad 0] \left(I - G_k I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1}\right) A^{(k)} I_{k-1}^k;$$
$$\widehat{A}_{21}^{(k)} = I_k^{k-1} A^{(k)} \left(I - I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1} G_k\right) \begin{bmatrix} I \\ 0 \end{bmatrix};$$
$$\widehat{A}_{22}^{(k)} = I_k^{k-1} A^{(k)} I_{k-1}^k = A^{(k-1)}.$$

Note that having computed $\widehat{\mathbf{v}}_1$ and $\widehat{\mathbf{v}}_2$, the solution $\mathbf{v}$ of (8.4.2) can be recovered by using the formula (8.4.1), i.e.,

$$\mathbf{v} = Y_1 \widehat{\mathbf{v}}_1 + Y_2 \widehat{\mathbf{v}}_2,$$

where,

$$Y_1 = \left(I - I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1} G_k\right) \begin{bmatrix} I \\ 0 \end{bmatrix},$$
$$Y_2 = I_{k-1}^k.$$

We have,

$$\mathbf{v} = Y\widehat{\mathbf{v}}, \quad \widehat{\mathbf{v}} = \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \widehat{\mathbf{v}}_2 \end{bmatrix}, \; Y = [Y_1, \; Y_2], \; Y_1 = Y_1^{(k)}, \; Y_2 = Y_2^{(k)}.$$

The transformed right–hand side vectors of (8.4.3) read similarly as follows:

$$\widehat{\mathbf{d}}_1 = [I \quad 0] \left(I - G_k I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1}\right) \mathbf{d} = Y_1^T \mathbf{d},$$
$$\widehat{\mathbf{d}}_2 = I_k^{k-1} \mathbf{d} = Y_2^T \mathbf{d}.$$

Therefore, the multiplicative AWM–HB preconditioner $B^{(k)}$ from Definition 3, starting with $B^{(0)} = A^{(0)}$, takes the following block–matrix form:

$$\widehat{B}^{(k)} = \begin{bmatrix} \widehat{B}_{11}^{(k)} & 0 \\ \widehat{A}_{21}^{(k)} & B^{(k-1)} \end{bmatrix} \begin{bmatrix} I & \widehat{B}_{11}^{(k)^{-1}} \widehat{A}_{12}^{(k)} \\ 0 & I \end{bmatrix}. \qquad (8.4.4)$$

The preconditioner $B^{(k)}$ is related to $\widehat{B}^{(k)}$ in the same way as $A^{(k)}$ to $\widehat{A}^{(k)}$. More precisely, one has

$$\widehat{B}^{(k)} = [Y_1, Y_2]^T B^{(k)} [Y_1, Y_2], \quad B^{(k)^{-1}} = [Y_1, Y_2] \widehat{B}^{(k)^{-1}} [Y_1, Y_2]^T.$$

We will show below that the inverse actions of $B^{(k)}$ can be computed only via the actions of $A^{(k)}$, Y_1, Y_2, and Y_1^T, Y_2^T in addition to the inverse actions of $\widehat{B}_{11}^{(k)}$.

We point out that (8.4.4) has precisely the same form as the algebraic multilevel method studied in Vassilevski [37] (see also Axelsson and Vassilevski [4] and Vassilevski [38]).

Observe that, in (8.4.4), $\widehat{B}_{11}^{(k)}$ is an appropriately scaled approximation of $\widehat{A}_{11}^{(k)}$. We have shown that $\widehat{A}_{11}^{(k)}$ is well–conditioned (see Lemma 3). Thus, it is possible to utilize some simple polynomial approximation $\widehat{B}_{11}^{(k)}$ for $\widehat{A}_{11}^{(k)}$ in the implementation. However, in order to take into account any possible jumps in the coefficient of the differential operator, it would be preferable to compute the diagonal part of $\widehat{A}_{11}^{(k)}$. This is computationally feasible since the basis functions of $V_k^1 = (I - Q_{k-1}^a)V_k^{(1)}$ have reasonably narrow support if m is not too large, which should be the case in practice. Nevertheless, one can employ in actual implementation the CG method to compute reliable approximate actions of $\widehat{A}_{11}^{(k)^{-1}}$.

Algorithm 1 *(*COMPUTING INVERSE ACTIONS OF $B^{(k)}$*)*
The inverse actions of $B^{(k)}$ are computed by solving the system

$$B^{(k)}\mathbf{w} = \mathbf{d},$$

with the change of basis $\mathbf{w} = Y\widehat{\mathbf{w}}$. Namely, by setting

$$\begin{aligned} \mathbf{w} &= Y_1\widehat{\mathbf{w}}_1 + Y_2\widehat{\mathbf{w}}_2 = [Y_1,\ Y_2]\begin{bmatrix} \widehat{\mathbf{w}}_1 \\ \widehat{\mathbf{w}}_2 \end{bmatrix}, \\ \widehat{\mathbf{d}}_1 &= Y_1^T\mathbf{d}, \\ \widehat{\mathbf{d}}_2 &= Y_2^T\mathbf{d}, \end{aligned}$$

$\mathbf{w} = B^{(k)^{-1}}\mathbf{d}$ *is computed via the solution of* $\widehat{B}^{(k)}\widehat{\mathbf{w}} = \widehat{\mathbf{d}}$ *as follows:*

FORWARD RECURRENCE*:*

1. *compute* $\widehat{\mathbf{z}}_1 = \widehat{B}_{11}^{(k)^{-1}}\widehat{\mathbf{d}}_1$*;*
2. *change the basis; i.e., compute* $\mathbf{z} = Y_1\widehat{\mathbf{z}}_1$*;*
3. *compute* $\widehat{\mathbf{d}}_2 := \widehat{\mathbf{d}}_2 - \widehat{A}_{21}^{(k)}\widehat{\mathbf{z}}_1 = Y_2^T(\mathbf{d} - \mathbf{A}^{(\mathbf{k})}\mathbf{z})$;
4. *compute* $\widehat{\mathbf{w}}_2 = B^{(k-1)^{-1}}\widehat{\mathbf{d}}_2$*;*
5. *change the basis, i.e., compute* $\mathbf{v} = Y_2\widehat{\mathbf{w}}_2$*;*

BACKWARD RECURRENCE*:*

1. *update the fine–grid residual, i.e., compute*

$$\widehat{\mathbf{d}}_1 := \widehat{\mathbf{d}}_1 - \widehat{A}_{12}^{(k)}\widehat{\mathbf{w}}_2 = Y_1^T(\mathbf{d} - A^{(k)}Y_2\widehat{\mathbf{w}}_2) = Y_1^T(\mathbf{d} - A^{(k)}\mathbf{v});$$

2. compute $\widehat{\mathbf{w}}_1 = \widehat{B}_{11}^{(k)^{-1}}\widehat{\mathbf{d}}_1$*;*

3. get the solution by $\mathbf{w} = Y_1\widehat{\mathbf{w}}_1 + Y_2\widehat{\mathbf{w}}_2 = Y_1\widehat{\mathbf{w}}_1 + \mathbf{v}$.

End

Note that the above algorithm requires only the actions of the standard stiffness matrix $A^{(k)}$, the actions of the transformation matrices Y_1 and Y_2 and their transposition Y_1^T and Y_2^T, the inverse action of $\widehat{B}_{11}^{(k)}$, and some suitable approximations to the well–conditioned matrices $\widehat{A}_{11}^{(k)}$. The actions of Y^{-1} are not required in the algorithm.

We now formulate the solution procedure for one preconditioning step using the multiplicative AWM–HB preconditioner $B = B^{(J)}$.

Algorithm 2 *(*MULTIPLICATIVE AWM–HB PRECONDITIONING*)*
Given the problem

$$B\mathbf{v} = \mathbf{d}\cdot$$

Initiate:

$$\mathbf{d}^{(J)} = \mathbf{d}\cdot$$

(A) **Forward recurrence.** *For* $k = J$ *down to* 1 *perform:*

1. Compute:

$$\widehat{\mathbf{d}}_1^{(k)} = [I \quad 0]\left(I - G_k I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1}\right)\mathbf{d}^{(k)};$$

2. Solve:

$$\widehat{B}_{11}^{(k)}\widehat{\mathbf{w}}_1 = \widehat{\mathbf{d}}_1^{(k)};$$

3. Transform basis:

$$\mathbf{w} = \left(I - I_{k-1}^k \widetilde{G}_{k-1}^{-1} I_k^{k-1} G_k\right)\begin{bmatrix} \widehat{\mathbf{w}}_1 \\ 0 \end{bmatrix} \begin{matrix} \}\mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \}\mathcal{N}_{k-1} \end{matrix};$$

4. Coarse–grid defect restriction:

$$\begin{aligned} \mathbf{d}^{(k-1)} &= I_k^{k-1}\mathbf{d}^{(k)} - \widehat{A}_{21}^{(k)}\widehat{\mathbf{w}}_1 \\ &= I_k^{k-1}(\mathbf{d}^{(k)} - A^{(k)}\mathbf{w}); \end{aligned}$$

5. Set $k = k-1$*. If* $k > 0$ *go to (1), else:*

6. Solve on the **coarsest level***:*

$$A^{(0)}\mathbf{x}^{(0)} = \mathbf{d}^{(0)};$$

(B) **Backward recurrence.**

1. *Interpolate result: Set* $k := k + 1$ *and compute*

$$\mathbf{x}^{(k)} = I_{k-1}^{k}\mathbf{x}^{(k-1)};$$

2. *Update fine-grid residual:*

$$\begin{aligned}\widehat{\mathbf{d}}_1^{(k)} &:= \widehat{\mathbf{d}}_1^{(k)} - \widehat{A}_{12}^{(k)}\mathbf{x}^{(k-1)} \\ &= \widehat{\mathbf{d}}_1^{(k)} - [I \quad 0](I - G_k I_{k-1}^{k}\widetilde{G}_{k-1}^{-1}I_k^{k-1})A^{(k)}\mathbf{x}^{(k)} \\ &= [I \quad 0](I - G_k I_{k-1}^{k}\widetilde{G}_{k-1}^{-1}I_k^{k-1})(\mathbf{d}^{(k)} - A^{(k)}\mathbf{x}^{(k)});\end{aligned}$$

3. *Solve:*

$$\widehat{B}_{11}^{(k)}\widehat{\mathbf{w}}_1 = \widehat{\mathbf{d}}_1^{(k)};$$

4. *Change the basis:*

$$\mathbf{w} = (I - I_{k-1}^{k}\widetilde{G}_{k-1}^{-1}I_k^{k-1}G_k)\begin{bmatrix}\widehat{\mathbf{w}}_1 \\ 0\end{bmatrix};$$

5. *Finally set:*

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k)} + \mathbf{w}.$$

6. *Set* $k := k + 1$. *If* $k < J$ *go to step (1) of (B), else* set

$$\mathbf{v} = \mathbf{x}^{(J)}.$$

END

Similarly, one preconditioning solution step for the additive AWM–HB preconditioner $D = D^{(J)}$ takes the following form:

Algorithm 3 *(*Additive AWM–HB preconditioning*)*
Given the problem

$$D\mathbf{v} = \mathbf{d}\cdot$$

Initiate:

$$\mathbf{d}^{(J)} = \mathbf{d}\cdot$$

(A) **Forward recurrence.** *For* $k = J$ *down to* 1 *perform:*

1. *Compute:*

$$\widehat{\mathbf{d}}_1^{(k)} = [I \quad 0]\left(I - G_k I_{k-1}^{k}\widetilde{G}_{k-1}^{-1}I_k^{k-1}\right)\mathbf{d}^{(k)};$$

2. Solve:

$$\widehat{B}_{11}^{(k)}\widehat{\mathbf{w}}_1 = \widehat{\mathbf{d}}_1^{(k)};$$

3. Transform basis:

$$\mathbf{x}^{(k)} = \left(I - I_{k-1}^{k}\widetilde{G}_{k-1}^{-1}I_k^{k-1}G_k\right)\begin{bmatrix} \widehat{\mathbf{w}}_1 \\ 0 \end{bmatrix} \begin{matrix} \}\mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \}\mathcal{N}_{k-1} \end{matrix} ;$$

4. Coarse–grid defect restriction:

$$\mathbf{d}^{(k-1)} = I_k^{k-1}\mathbf{d}^{(k)};$$

5. Set $k = k - 1$. If $k > 0$ go to (1), else :

6. Solve on the **coarsest level***:*

$$A^{(0)}\mathbf{x}^{(0)} = \mathbf{d}^{(0)};$$

(B) **Backward recurrence.**

1. Interpolate result: Set $k := k + 1$ and compute

$$\mathbf{w} = I_{k-1}^{k}\mathbf{x}^{(k-1)};$$

2. Update at level k:

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k)} + \mathbf{w};$$

3. Set $k := k + 1$. If $k < J$ go to step (1) of (B), else set

$$\mathbf{v} = \mathbf{x}^{(J)}.$$

END

For both the additive and multiplicative preconditioners, it is readily seen that the above implementations require only actions of the stiffness matrices $A^{(k)}$, the mass matrices $G^{(k)}$, and the transformation matrices I_{k-1}^{k} and I_k^{k-1}. The approximate inverse actions of $\widehat{A}_{11}^{(k)}$ can be computed via some inner iterative algorithms. Similarly, the action of $\widetilde{G}_{k-1}^{-1}$ can be computed as approximate solutions of the corresponding mass–matrix problem using m steps of some simple iterative methods. Therefore, at each discretization level k, one performs a number of arithmetic operations proportional to the degrees of freedom at that level denoted by n_k. In the case of local mesh refinement, the corresponding operations involve only the stiffness and mass matrices computed for the subdomains where local refinement was made. Hence, even in the case of locally refined meshes, the cost of the AWM–HB methods is proportional to $N = n_J$.

The proportionality constant depends linearly on $m = O(\log \tau^{-1})$, but is independent of J (or h). Some numerical results for the AWM–HB preconditioners can be found from Vassilevski and Wang [42].

A performance comparison with the BPX method [8] and Stevenson's method [34] on more difficult elliptic problems in three–dimensions and in other applications such as interface domain decomposition preconditioning is yet to be seen.

§9 Numerical experiments

In this section we present some numerical results to illustrate the efficiency of the method discussed in §7. Consider the boundary value problem of seeking u satisfying

$$\begin{aligned} \mathcal{L}_\epsilon u \equiv -\epsilon\Delta u + \underline{b}\cdot\nabla u &= f \quad \text{in } \Omega, \\ u &= g \quad \text{on } \partial\Omega\cdot \end{aligned} \tag{9.1}$$

Here, $f \in L^2(\Omega)$, $g \in H^{\frac{1}{2}}(\partial\Omega)$, and $\underline{b} = [b_1 b_2]$ are given single–valued or vector–valued functions. We assume that all given functions are sufficiently smooth on their domains. For simplicity, we take Ω to be a square domain and $g = 0$ on $\partial\Omega$.

If u is a solution of (9.1), then it solves the following problem:

$$\mathcal{M}_\epsilon u \equiv -\delta\nabla\cdot(\underline{b}\,\mathcal{L}_\epsilon u) + \mathcal{L}_\epsilon u = -\delta\nabla\cdot(\underline{b}\,f) + f \quad \text{in } \Omega, \tag{9.2}$$

subject to the boundary condition $u = 0$ on $\partial\Omega$. Here $\delta > 0$ is a parameter. The purpose of considering the problem (9.2) is to get the so–called streamline derivative $\frac{\partial u}{\partial \underline{b}} = \underline{b}\cdot\nabla u$ in a variational formula for u. More precisely, by testing (9.2) against any $\psi \in H_0^1(\Omega)$ one obtains

$$\begin{aligned} b_\epsilon(u,\ \psi) &\equiv \epsilon(\nabla u, \nabla\psi) + (\psi, \underline{b}\cdot\nabla u) + \delta(\underline{b}\cdot\nabla u,\ \underline{b}\cdot\nabla\psi) \\ &\quad -\epsilon\delta(\Delta u,\ \underline{b}\cdot\nabla\psi) \\ &= (f_\delta, \psi), \end{aligned} \tag{9.3}$$

for all $\psi \in H_0^1(\Omega)$. Here $f_\delta = f - \delta\nabla\cdot(\underline{b}f)$.

9.1 Galerkin discretization

The Galerkin method for the approximation of u is based on the variational problem (9.3). Let $V = V_h$ be a C^0-conforming finite element space of piecewise polynomials corresponding to a quasiuniform triangulation $\mathcal{T}_h$ of

Ω. The Galerkin approximation is a function $u_h \in V_h$ such that

$$\begin{aligned} b_\epsilon(u_h, \psi) &\equiv \epsilon(\nabla u_h, \nabla \psi) + (\psi, \underline{b} \cdot \nabla u) \\ &\quad + \delta(\underline{b} \cdot \nabla u, \underline{b} \cdot \nabla \psi) - \epsilon\delta \sum_{T \in T_h} \int_T \Delta u_h(\underline{b} \cdot \nabla \psi)\, dx \\ &= (f_\delta, \psi), \end{aligned} \tag{9.1.1}$$

for all $\psi \in V_h$. For continuous piecewise linear functions, one has $\Delta u_h = 0$ on each element. It follows that the discrete problem seeks $u_h \in V_h$ such that

$$\epsilon(\nabla u_h, \nabla \psi) + (\psi, \underline{b} \cdot \nabla u_h) + \delta(\underline{b} \cdot \nabla u_h, \underline{b} \cdot \nabla \psi) = (f_\delta, \psi), \tag{9.1.2}$$

for all $\psi \in V_h$.

The convection term is assumed to satisfy

$$\nabla \cdot \underline{b} \leq 0 \quad \text{in } \Omega. \tag{9.1.3}$$

For a convergence analysis of the streamline diffusion finite element approximation u_h, we refer to [23] and [2].

9.2 Numerical tests

We choose the same test examples as in [2]. Namely,

$$\underline{b} = [(1 - x\cos\alpha)\cos\alpha(1 - y\sin\alpha)\sin\alpha],$$

for various angles α. Note that $\nabla \cdot \underline{b} = -1$. The right hand side f is chosen so that $u = x(1-x)y(1-y)$ is the exact solution. Thus, the right-hand side function f is ϵ-dependent. The stopping criterion was that the relative error of the residual be less than 10^{-8} in the discrete L^2–norm. The objective is to test the number of iterations in the solution procedure by using the wavelet-like hierarchical basis.

The matrix form of the discretized problem (9.1.2) reads as follows:

$$A\mathbf{u} = \mathbf{f}. \tag{9.2.1}$$

Here A is a nonsymmetric matrix. For small ϵ, it is very difficult to find a good preconditioner for A. It was seen in [2] that a block–ILU factorization method turns out to be very robust with respect to arbitrary positive ϵ, though very little is known theoretically on this good performance.

For any finite element function v, we use the bold face $\mathbf{v}$ to denote the vector with respect to the nodal basis and $\widehat{\mathbf{v}}$ the vector representation with respect to the modified hierarchical basis. In the implementation, the modified hierarchical basis is employed to provide a preconditioner for the global stiffness matrix A as follows. Let Y be the transformation from $\widehat{\mathbf{v}}$

to $\mathbf{v}$ such that $\mathbf{v} = Y\widehat{\mathbf{v}}$. The preconditioner is given by $P \equiv (YY^T)^{-1}$ with Y^T being the conjugate of Y. Below we discuss the computation of Y and Y^T.

Let

$$Y_1^{(k)} = \left(I - I_{k-1}^k \widetilde{G}^{(k-1)^{-1}} I_k^{k-1} G^{(k)}\right) \begin{bmatrix} I \\ 0 \end{bmatrix} \begin{matrix} \} \mathcal{N}_k^1 = \mathcal{N}_k \setminus \mathcal{N}_{k-1} \\ \} \mathcal{N}_{k-1} \end{matrix} \tag{9.2.2}$$

and

$$Y_2^{(k)} = I_{k-1}^k. \tag{9.2.3}$$

Here, $G^{(k)}$ stands for the mass matrix at kth discretization level, I_{k-1}^k is the natural coarse-to-fine interpolation matrix from $(k-1)$th grid to the kth one and $I_k^{k-1} = (I_{k-1}^k)^T$. Also, $\widetilde{G}^{(s)^{-1}}$ is an approximate inverse of the mass matrix $G^{(s)}$. For example, a good choice for $\widetilde{G}^{(s)^{-1}}\mathbf{w}$ would be an approximate solution of $G^{(s)}\mathbf{x} = \mathbf{w}$ by polynomial iterative methods. In practice, the Jacobi and conjugate gradient methods are good candidates. The numerical results in this section are based on the Jacobi iterative method with two iterations.

Algorithm 4 *For any given* $\widehat{\mathbf{v}} = (\widehat{\mathbf{v}}_1^{(J)}, \cdots, \widehat{\mathbf{v}}_1^{(1)}, \widehat{\mathbf{v}}^{(0)})^T$, *the action* $\mathbf{v} = Y\widehat{\mathbf{v}}$ *is computed as follows:*

- *Set* $\mathbf{v}^{(0)} = \widehat{\mathbf{v}}^{(0)}$.
- *For* $k = 1$ *to* J *do*

$$\mathbf{v}^{(k)} = Y_2^{(k)}\mathbf{v}^{(k-1)} + Y_1^{(k)}\widehat{\mathbf{v}}_1^{(k)}.$$

- $\mathbf{v} = \mathbf{v}^{(J)}$.

Algorithm 5 *For any given* $\mathbf{d} = (\mathbf{d}_1^{(J)}, \cdots, \mathbf{d}_1^{(1)}, \mathbf{d}^{(0)})^T$, *the action* $\widehat{\mathbf{d}} = Y^T\mathbf{d}$ *is computed as follows:*

- *Set* $\widehat{\mathbf{d}}^{(J)} = \mathbf{d}^{(0)}$.
- *For* $k = J$ *down to* 1 *do*

$$\begin{aligned} \widehat{\mathbf{d}}_1^{(k)} &= Y_1^{(k)^T}\widehat{\mathbf{d}}^{(k)} \\ \widehat{\mathbf{d}}^{(k-1)} &= Y_2^{(k)^T}\widehat{\mathbf{d}}^{(k)}. \end{aligned}$$

- $\widehat{\mathbf{d}} = (\widehat{\mathbf{d}}_1^{(J)}, \cdots \widehat{\mathbf{d}}_1^{(1)}, \widehat{\mathbf{d}}^{(0)})^T$.

Table 1. Iteration counts for $h^{-1} = 64$, $\alpha = 75^o$.

	$\epsilon = 1$ $\delta = 0.001$	$\epsilon = 0.1$ $\delta = 0.01$	$\epsilon = 10^{-2}$ $\delta = 0.1$	$\epsilon = 10^{-3}$ $\delta = 1.0$	$\epsilon = 10^{-4}$ $\delta = 10$
iter	30	28	42	59	74

Table 2. Iteration counts for $h^{-1} = 64$, $\alpha = 105^o$.

	$\epsilon = 1$ $\delta = 0.001$	$\epsilon = 0.1$ $\delta = 0.01$	$\epsilon = 10^{-2}$ $\delta = 0.1$	$\epsilon = 10^{-3}$ $\delta = 1.0$	$\epsilon = 10^{-4}$ $\delta = 10$
iter	31	27	70	$\star$	$\star$

Table 3. Iteration counts for $h^{-1} = 128$, $\alpha = 75^o$.

	$\epsilon = 1$ $\delta = 0.0025$	$\epsilon = 0.1$ $\delta = 0.025$	$\epsilon = 10^{-2}$ $\delta = 0.25$
iter	32	29	48

Table 4. Iteration counts for $h^{-1} = 128$, $\alpha = 105^o$.

	$\epsilon = 1$ $\delta = 0.0025$	$\epsilon = 0.1$ $\delta = 0.025$	$\epsilon = 10^{-2}$ $\delta = 0.25$
iter	32	29	71

Once the preconditioner $P = (YY^T)^{-1}$ is known, one can solve the discrete problem (9.2.1) by using the generalized conjugate gradient method employed in [2]. We comment that this simple preconditioner may not work well for convection-dominated diffusion problems. This fact can be seen from the numerical results illustrated in Tables 1–4.

In the Tables 1–4, "$\star$" is used to indicate a non-convergence in 100 iterations. It is clear that the use of the wavelet–like hierarchical basis gives an efficient preconditioner if the problem is not convection-dominated.

The present implementation of the modified hierarchical basis is comparable with the additive preconditioning method discussed in [42]. The information is also contained in ALGORITHM 3 of this paper with $\widehat{B}_{11}^{(k)} = Ch_k^{-2}I_k$, where I_k stands for the identity matrix.

Acknowledgments. The research of Vassilevski is supported in part by the U.S. NSF grant INT-95-06184 and also in part by Bulgarian Ministry for Education grant MM–415, 1994. The research of Wang is supported in part by the NSF grant INT-93-09286. The authors are also grateful to Dr. Oswald for his careful comments and remarks on the paper.

References

[1] Arnold, D. N., Falk, R. S., and Winther, R., Preconditioning in $H(\mathrm{div})$ and applications, *Math. Comp.*, to appear.

[2] Axelsson, O., Eijkhout, V., Polman, B., and Vassilevski, P. S., Incomplete block–matrix factorization iterative methods for convection–diffusion problems, *BIT* **29** (1989), 867–889.

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

[4] Axelsson, O. and Vassilevski, P. S., Algebraic multilevel preconditioning methods, II. *SIAM J. Numer. Anal.* **27** (1990), 1569–1590.

[5] Bank, R. E. and Dupont, T., Analysis of a two–level scheme for solving finite element equations, *Technical Report CNA–159, Center for Numerical Analysis, The University of Texas at Austin*, 1980.

[6] Bramble, J. H., Pasciak, J. E., Wang, J., and Xu, J., Convergence estimates for product iterative methods with applications to domain decomposition, *Math. Comp.* **57** (1991), 1–21.

[7] Bramble, J. H., Pasciak, J. E., Wang, J., and Xu, J., Convergence estimates for multigrid algorithms without regularity assumptions, *Math. Comp.* **57** (1991), 23–45.

[8] Bramble, J. H., Pasciak, J. E., and Xu, J., Parallel multilevel preconditioners, *Math. Comp.* **55** (1990), 1–22.

[9] Bramble, J. H., *Multigrid methods*, Pitman Research Notes in Mathematical Series, # 234, Longman Scientific & Technical, UK, 1995 (second edition).

[10] Bramble, J. H. and Pasciak, J. E., Iterative techniques for time dependent Stokes problems, *Preprint, ISC, Texas A&M University*, 1994.

[11] Brezzi, F. and Fortin, M., *Mixed and Hybrid Finite Element Methods*, Springer Verlag, New York, 1991.

[12] Carnicer, J. M., Dahmen, W., and Pena, J. M., Local decompositions of refinable spaces and wavelets, *Appl. Comp. Harm. Anal.* **3** (1996), 127–153.

[13] Ciarlet, P., *The finite element method for elliptic problems*, North Holland, 1977.

[14] Chui, C. K., *An Introduction to Wavelets*, Academic Press, Boston, MA, 1992.

[15] Dahmen, W. and Kunoth, A., Multilevel preconditioning, *Numer. Math.* **63** (1992), 315–344.

[16] Dahmen, W., Kunoth, A., and Urban, K., A wavelet–Galerkin method for the Stokes equations, *Bericht Nr. 111, Institut für Geometrie und Praktische Mathematik, RWTH, Aachen*, 1995.

[17] Daubechies, I., *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[18] Girault, V. and Raviart, P.–A., *Finite Element Methods for Navier–Stokes Equations*, Springer Verlag, Berlin-Heidelberg-New York-Tokyo, 1986.

[19] Griebel, M. and Oswald, P., On the abstract theory of additive and multiplicative Schwarz algorithms, *Numer. Math.* **70** (1995), 163–180.

[20] Griebel M. and Oswald, P., Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems, *Advances in Computational Mathematics* **4** (1994), 171-206.

[21] Hood, P. and Taylor, C., A numerical solution of the Navier–Stokes equations using the finite element technique, *Comp. in Fluids* **1** (1973), 73–100.

[22] Jaffard, S., Wavelet methods for fast resolution of elliptic problems, *SIAM J. Numer. Anal.* **29** (1992), 965–986.

[23] Johnson, C., *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, 1994.

[24] Kotyczka, U. and Oswald, P., Piecewise linear prewavelets of small support, in: *Approximation Theory VIII*, vol. 2 (wavelets and multilevel approximation), C. Chui, L.L. Schumaker, eds., World Scientific, Singapore, 1995, 235-242.

[25] Lorentz, R. and Oswald, P., Constructing economical Riesz bases for Sobolev spaces, *talk presented at the Domain Decomposition Conference held in Bergen*, Norway, June 3–8, 1996.

[26] Mallat, S., Multiresolution approximations and wavelet orthonormal bases of $L^2(\boldsymbol{R})$, *Trans. Amer. Math. Soc.* **315** (1989), 69–88.

[27] Maubach, J. M., Local bisection refinement for n–simplicial grids generated by reflection, *SIAM J. Sci. Comput.* **16** (1995), 210–227.

[28] Mitchell, W. F., Optimal multilevel iterative methods for adaptive grids, *SIAM J. Sci. Stat. Comput.* 13(1992), 146–167.

[29] Ong, M.-E. G., Hierarchical basis preconditioners in three dimensions, *SIAM J. Sci. Comput.*, to appear.

[30] Oswald, P., *Multilevel Finite Element Approximation: Theory and Applications*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[31] Rusten, T. and Winther, R., A preconditioned iterative method for saddle–point problems, *SIAM J. Matrix Anal. Appl.* **13** (1992), 887–904.

[32] Silvester, D. S. and Wathen, A. J., Fast iterative solution of stabilized Stokes systems, part II: using general block preconditioners, *SIAM J. Numer. Anal.* **31** (1994), 1–16.

[33] Stevenson, R., Robustness of the additive and multiplicative frequency decomposition multilevel method, *Computing* **54** (1995), 331–346.

[34] Stevenson, R., A robust hierarchical basis preconditioner on general meshes, *Technical Report # 9533, Department of Mathematics, University of Nijmegen*, Nijmegen, The Netherlands, 1995.

[35] Schatz, A. H., An observation concerning Ritz–Galerkin methods with indefinite bilinear forms, *Math. Comp.* **28** (1974), 959–962.

[36] Schatz, A. H. and Wang, J., Some new error estimates for Ritz-Galerkin methods with minimal regularity assumptions, *Math. Comp.* **65** (1996), 19–27.

[37] Vassilevski, P. S., Nearly optimal iterative methods for solving finite element elliptic equations based on the multilevel splitting of the matrix, *Technical Report # 1989–09, Institute for Scientific Computation, University of Wyoming*, Laramie, WY, USA, 1989.

[38] Vassilevski, P. S., Hybrid V–cycle algebraic multilevel preconditioners, *Math. Comp.* **58** (1992), 489–512.

[39] Vassilevski, P. S., On two ways of stabilizing the hierarchical basis multilevel methods, *SIAM Review*, to appear.

[40] Vassilevski, P. S. and Wang, J., Multilevel iterative methods for mixed finite element discretizations of elliptic problems, *Numer. Math.* **63** (1992), 235–248.

[41] Vassilevski, P. S. and Wang, J., Stabilizing the hierarchical basis by approximate wavelets, I: Theory, *Numer. Linear Alg. Appl.*, to appear, 1996.

[42] Vassilevski, P. S. and Wang, J., Stabilizing the hierarchical basis by approximate wavelets, II: Implementation and Numerical Experiments, *SIAM J. Sci. Comput.*, submitted, 1996.

[43] Xu, J., Iterative methods by space decomposition and subspace correction, *SIAM Review* **34** (1992), 581–613.

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

[45] Yserentant, H., Old and new convergence proofs for multigrid algorithms, *Acta Numerica*, 1993, 285–326.

Panayot S. Vassilevski
Center of Informatics and Computing Technology
Bulgarian Academy of Sciences
"Acad. G. Bontchev" street, Block 25 A
1113 Sofia, Bulgaria
panayot@iscbg.acad.bg

Junping Wang
Department of Mathematics
University of Wyoming
Laramie, Wyoming 82071
wang@schwarz.uwyo.edu