# HIERARCHICAL BASES FOR ELLIPTIC PROBLEMS

## W. DÖRFLER

ABSTRACT. Linear systems of equations with positive and symmetric matrices often occur in the numerical treatment of linear and nonlinear elliptic boundary value problems. If the CG algorithm is used to solve these equations, one is able to speed up the convergence by "preconditioning." The method of preconditioning with hierarchical basis has already been considered for the Laplace equation in two space dimensions and for linear conforming elements. In the present work this method is generalized to a large class of conforming and non-conforming elements.

## 0. INTRODUCTION

Consider elliptic problems in variational form of order $2m$. Let $X := H_0^{m,2}(\Omega)$, and $\Omega \in \mathbb{R}^n$ be a bounded domain. We want to find $u \in X$ such that

$$(0.1) \qquad a(u, v) := \int_\Omega A[\nabla^m u, \nabla^m v] = F(v) \quad \forall v \in X.$$

Here, $A$ denotes a strictly positive (possibly $x$-dependent) matrix; $F$ is a functional on $X$, and $a(\cdot, \cdot)$ is assumed to satisfy

$$c_0 \|u\|_{m,2,0}^2 \le a(u, u) \le C_0 \|u\|_{m,2,0}^2$$

for all $u \in X$. The solution of this problem exists according to the Lax-Milgram theorem. One way to solve this problem numerically is to approximate $X$ by a sequence of finite-dimensional spaces $X_N$ for $N \in \mathbb{N}$. If $\{u_i\}_{i=1,N}$ is a basis of $X_N$, writing $u$ in the form

$$u = \sum_{i=1}^N \alpha_i u_i$$

leads to the equation

$$(0.2) \qquad\qquad S z = F,$$

where $z = [\alpha_i]_{i=1,N}$, $S_{ij} = a(u_i, u_j)$, and $F_j = F(u_j)$. The $N \times N$-matrix $S$ is positive and symmetric, and (0.2) can therefore be solved by the conjugate gradient (CG) algorithm. Given any initial vector $z^0$, we get a sequence of vectors $z^k$ for $k \in \mathbb{N}$. Theoretically, the exact solution is reached with $z^N$.

In practice, however, the procedure will stop on account of a stopping criterion for the residual,

$$|Sz^k - F|^2 \le \varepsilon,$$

where $\varepsilon$ is a given small number. The number of iteration steps is controlled by the error estimate

$$|z^k - \overline{z}|_S \le 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k |z^0 - \overline{z}|_S,$$

where $\overline{z}$ is the exact solution, $\kappa = \text{cond}(S)$ is the condition number of $S$, and $|z|_S^2 := z^* S z$ [1, p. 25].

Now let $X_N$ be a finite element space. For a regular discretization of $\Omega$ we get

(0.3)                                  $\kappa \sim h^{-2m},$

where $h$ is the minimal diameter of the discretization [7, p. 187]. In order to speed up the algorithm, one can try to find basis transformations $B$ for which the matrix $(B^{-1})^* S B^{-1}$ has a smaller condition number ("preconditioning"). On the other hand, we do not want the computation of $B^{-1} z$ and $(B^{-1})^* z$ to be too complex. We require that the additional work is $O(N)$ (that is what we usually have for $S$).

Many different preconditioning techniques can be found in the literature (e.g., SSOR-preconditioning, incomplete factorization [1, Chapter 1.4]). This work is devoted to the concept of hierarchical bases.

Before proceeding, we want to point out that speeding up solvers for linear problems is also of importance for nonlinear problems. We refer especially to the nonlinear CG-method [6, Chapter 7].

In this work we follow the theory of hierarchical bases developed by H. Yserentant [9]. He showed that using this method reduces the condition number of $S$ in the case $n = 2$, $m = 1$, and for conforming linear elements, from $O(N)$ to $O(|\log(N)|^2)$. Moreover, the transformation to this new basis requires $O(N)$ operations and the inverse matrix is also easily computed because this transformation can be represented as a product of simply structured matrices.

The idea of this method is to decompose the finite element space $V$ into a direct sum

$$V = \bigoplus_{k=1}^{J} W_k,$$

where the spaces $W_k$ belong to different levels of refinement of a given coarse discretization. They have the property that the conditioning of $S$, when restricted to $W_k$, is uniformly bounded and that the operator norm of the projection $V \to W_k$ is estimated by $C(1 + J - k)$. For this we make use of an estimate which states that for all $v \in H_0^{1,2}(B_R)$, $B_R$ being the ball with radius $R$, and $0 < \sigma \le \frac{1}{2} R$,

(0.4)           $\dfrac{1}{|B_\sigma|} \displaystyle\int_{B_\sigma} |v| \le C \left( 1 + \log \left( \frac{R}{\sigma} \right) \right)^{1/2} \left( \int_{B_R} |\nabla v|^2 \right)^{1/2}.$

(We call this estimate the "inverse estimate," because the left-hand side can approximate $\sup_{B_R} |v|$ for $\sigma \to 0$.)

Unfortunately, this procedure does not provide a similar improvement of the conditioning in higher space dimensions. Recently, a new class of preconditioners was found which gives dimension-independent results (see [3, 10]).

However, the aim of this work is to generalize the ideas of [9] to higher-order problems and to nonconforming finite element spaces. The result is that the condition numbers that can be achieved depend only on the space dimension and a certain defect, produced by the projections from one level to the next. The second phenomenon does not occur for linear conforming elements.

This work is organized as follows. In the first and second sections the notations and assumptions on the discretizations and the finite element spaces are introduced. The third section shows how to construct a hierarchical basis and states the requirements for the basis transformation which ensure the result. It is shown, by an example, how this can be applied to general conforming elements interpolating at the vertices of the discretization. This new basis defines a set of nested spaces $\widetilde{V}_k \subset \widetilde{V}_{k+1}$, for $k \geq 0$, such that $\widetilde{V}_J = V$ and $\widetilde{V}_k$ is isomorphic (but not identical as in the case of linear conforming elements) to the finite element space $V_k$ on the $k$th level discretization. In §4, we then study the norm of the isomorphism $\widetilde{V}_k \to V_k$. In §5 we consider the stiffness matrix when restricted to $\widetilde{V}_{k+1} \backslash \widetilde{V}_k$. The lowest and largest eigenvalues are estimated with respect to a suitable vector norm. We also achieve an estimate for the projection $V \to \widetilde{V}_k$. Here we use the inverse estimate (0.4), which has to be assumed to be valid also for nonconforming elements. Under the usual assumptions on $V$ and the discretization, such an estimate is proved for two dimensions in the supplement section at the end of this issue. In §6 the estimates for the smallest and the largest eigenvalue of $S$ corresponding to the hierarchical basis are derived. Recall that $\text{cond}(S) = \lambda_{\max}/\lambda_{\min}$. All these results are summarized in §7. In §8 we consider as an example the linear but nonconforming case and include some numerical results.

*Notation.* The following notation will be used:

| | |
|---|---|
| $H^{m,p}(\Omega)$, $H_0^{m,p}(\Omega)$ | Sobolev spaces (with zero boundary data); |
| $\|\cdot\|_{m,p}$, $\|\cdot\|_{m,p,0}$ | corresponding norms: $\|\cdot\|_{m,p}^2 = \|\cdot\|_{m,p,0}^2 + \|\cdot\|_{m-1,p}^2$; |
| $\|\cdot\|_{m,p;G}$, $\|\cdot\|_{m,p,0;G}$ | as above, but restricted to $G \subset \Omega$; |
| $d_\Omega$ | $\text{diam}(\Omega)$; |
| $\|\cdot\|_{m,p,*;G}$ | "scale invariant norm": $\|\cdot\|_{m,p,*;G}^2$ $$= \sum_{k=0}^{m} d_G^{-2(m-k)} \|\cdot\|_{k,p,0;G}^2 \text{ for } G \text{ with } d_G > 0;$$ |
| $\mathbb{P}_m(\Omega)$ | the set of all polynomials of degree $m$ on $\Omega$; |
| $\|G\|$ | the $n$-dimensional Lebesgue measure of $G \subset \mathbb{R}^n$; |
| $|v|$ | the Euclidean vector norm for $v \in \mathbb{R}^n$. |

In this article we use the symbols $c$ and $C$ as unspecified constants, that is, in a sequence of estimates these constants may vary from one estimate to the

next (collecting all factors we are not interested in) although the symbol is not changed.

## 1. THE DISCRETIZATION

Let $\Omega$ be a bounded domain in $\mathbb{R}^n$ assumed to be representable by a set of linear inequalities:

$$\Omega = \{x \in \mathbb{R}^n : p_k \cdot x \le \beta_k, \; k = 1, \ldots, M\},$$

where $p_k$ and $\beta_k$ are given vectors and scalars, respectively, and $M \ge n$ is an integer. This means that $\partial\Omega$ has a piecewise flat boundary. Let $\Omega$ be divided into simplices

$$\overline{\Omega} = \bigcup_i T_i$$

(for $i$ in a finite set of indices), such that $|T_i \cap T_j| = 0$ for $i \ne j$ (consider the simplices as closed subsets of $\overline{\Omega}$). This partition will be referred to as *macro-discretization* $\mathscr{T}_0$.

Denote by $E$ the unit simplex. Let $\mathscr{P}_E$ be a set of *points* contained in $E$; for definiteness,

$$\mathscr{P}_E = \{\hat{p}_1, \ldots, \hat{p}_{N_E}\}.$$

Let $m$ be the degree of the problem we want to approximate (see §0). To each $\hat{p} \in \mathscr{P}_E$ we assign a nonempty set $\mathscr{X}_{\hat{p}}$:

$$\mathscr{X}_{\hat{p}} \subset \{(\nu, r) : \nu \in \mathbb{N}, \; 0 \le \nu \le m - 1, \; r \in \mathbb{N}^\nu, \; r = [r_1, \ldots, r_\nu]\}$$

(for $\nu = 0$ we let $r = 0$). An element of the set

$$\mathscr{N}_E := \{\hat{q} = (\hat{p}, \nu, r) : \hat{p} \in \mathscr{P}_E, \; (\nu, r) \in \mathscr{X}_{\hat{p}}\}$$

will be called a *node*. As will be pointed out in §2, the number $\nu$ gives the order of differentiation, and the vector $r$ gives a set of directions for an evaluation functional localized at $\hat{p}$.

For every simplex $T$, let $e_T : E \to T$ denote an affine linear isomorphism with $\det(\nabla e_T) > 0$. Define $e_T$ on the set $\mathscr{N}_E$ by $e_T(\hat{p}, \nu, r) := (e_T(\hat{p}), \nu, r)$. For convenience we assume the following symmetry in the definition of $\mathscr{N}_E$: for any $e_E$ assume

$$e_E(\mathscr{N}_E) = \mathscr{N}_E.$$

A *refinement* of $E$ is a partition $E = \bigcup_i E_i$ into a finite number of subsimplices $E_i$ ($|E_i \cap E_j| = 0$ if $i \ne j$). For simplicity we only consider refinements for which $d_{E_i} \ge \frac{1}{2} d_E$. For each $E_i$ fix a mapping $e_{E_i}$ (because of the assumed symmetry it does not matter which). A set of nodes on the refined simplex is defined by

$$\mathscr{N}_E' := \{\hat{q}' = e_{E_i}(\hat{q}), \; \hat{q} \in \mathscr{N}_E, \; E_i \subset E\}.$$

It follows that, if $E_i$ and $E_j$ have $\hat{p}'$ as a common point, then there are $\hat{q}_{i_1}, \hat{q}_{i_2} \in \mathscr{N}_E$ such that $(\hat{p}', \nu, r) = e_{E_i}(\hat{q}_{i_1}) = e_{E_j}(\hat{q}_{i_2})$.

Now let us construct a set of nodes on $\mathscr{T}_0$. For each $T \in \mathscr{T}_0$ we fix $e_T : E \to T$. Let

$$\mathscr{N}_0 := \{q = e_T(\hat{q}) : T \in \mathscr{T}_0, \; \hat{q} \in \mathscr{N}_E\}.$$

We proceed by refining the discretization $\mathscr{T}_0$. For this purpose we map a partition of the unit simplex via $e_T$ onto $T \in \mathscr{T}_0$. In this way we can construct local or global refinements. One step of refinement is called *complete* if $T$ is

partitioned into subsimplices $T_i$ such that any simplex $T_i$ with $d_{T_i} = d_T$ will not be refined in any further step. Applying one complete step of refinement to $\mathcal{T}_0$ yields $\mathcal{T}_1$ and, as above, also a set of nodes $\mathcal{N}_1$. After having applied $k$ steps, we get $\mathcal{T}_k$ and $\mathcal{N}_k$, respectively. It will be assumed that this process leads to what is called a "regular discretization." This means that the quotient of the diameter of $T$ and the radius of the inscribed ball remains uniformly bounded for all simplices in each discretization. For example, such algorithms are described in [2] or [8].

In many estimates it is important to separate the properties depending on the scale of a simplex $T$ from those which are scale-independent. This will be indicated in the following way. For any $T$ let $\widehat{T} := d_T^{-1}T$ denote the scaled simplex. If a constant depends on $\widehat{T}$, this means that it depends only on the scale-independent properties of $T$.

If we write $a \sim b$ for two expressions $a$ and $b$, which both depend on $T$, this means that there is an estimate $ca \le b \le Ca$ with constants $c, C$ depending only on $\widehat{T}$.

## 2. THE FINITE ELEMENTS

Let $V_k$ be a finite element space corresponding to $\mathcal{T}_k$. Following [4, p. 78], we can describe it by a triple $(\Omega, S_k, \Sigma_k)$, where $S_k$ is a set of functions on $\Omega$ and $\Sigma_k$ a set of functionals on $S_k$. Let $E$ be the unit simplex and $S_E$ a function space on $E$ such that

$$\mathbb{P}_{m-1}(E) \subset S_E \subset C^m(E).$$

On $T \in \mathcal{T}_k$ we define $S_k|_T$ via the affine linear isomorphisms $e_T$ by

$$S_k|_T := \{u = v \circ e_T^{-1} : v \in S_E\}.$$

Note that $u \in \mathbb{P}_{m-1}(T)$ whenever $v \in \mathbb{P}_{m-1}(E)$.

Now we want to construct the functionals. First, consider the discretization $\mathcal{T}_0$. Define $\mathcal{F}_0$ to be the set of all points in $\Omega$ lying in the $(n-1)$-dimensional interior of the boundary of a simplex. Define vector functions $a_1, \ldots, a_n$ with the following properties:

    1. $a_n$ is a unit normal vector on the corresponding face;
    2. $a_1, \ldots, a_{n-1}$ are linear independent tangential vectors;
    3. the vector fields are constant on each connected component.

For $\mathcal{T}_k$ (where $\mathcal{F}_k$ is defined analogously) such vector fields are defined by the following recursion:

        if $p \in \mathcal{F}_{k-1}$, then $a_1, \ldots, a_n$ are already defined;
        if $p \in \mathcal{F}_k \backslash \mathcal{F}_{k-1}$, then define $a_1, \ldots, a_n$ using criteria 1–3.

The canonical basis of $\mathbb{R}^n$, interpreted as a constant vector field, is denoted by $a_{n+1}, \ldots, a_{2n}$. A global regularity condition on $S_k$ may be given by $S_k \subset C^{m-1}(\overline{\Omega})$, in which case we speak of *conforming elements*. If, on the other hand, $S_k \not\subset C^{m-1}(\overline{\Omega})$, the elements are called *nonconforming*. In the latter case there are of course some regularity conditions which assure the usefulness of these elements for approximating problem (0.1).

The set $\Sigma_k$ can now be defined as follows. Let $\mathcal{N}_k$ be the set introduced in §1. For each $q = (p, \nu, r) \in \mathcal{N}_k$ we define a functional

$$F_q^k : C^\nu(\Omega) \to \mathbb{R},$$

where for $v \in S_k$

$$F_q^k(v) = \nabla^\nu v(p)[a_{r_1}(p), \ldots, a_{r_\nu}(p)]$$

with $1 \le r_j \le 2n$. Of course, we have to assume that the functionals are well defined on functions in $S_k$. This follows immediately for conforming elements. For nonconforming elements, $F_q^k$ is a functional on $S_k$ if $p$ and the vector fields $a_{r_j}$ are chosen properly.

A basis of $V_k$ shall be given by a set of functions $\Phi_q^k \in S_k$, $q \in \mathcal{N}_k$, for which $F_q^k(\Phi_{\bar{q}}^k) = \delta_{q\bar{q}}$. Recall that a function $u$ in $S_k$ is assumed to be uniquely defined by its values $F_q^k(u)$.

Now assume that we work with $J + 1$ levels, that is, $k$ varies from 0 to $J$ for a given $J \in \mathbb{N}$. For convenience we will omit the level index in the special case $k = J$, thus writing $V = V_J$ and so on.

For $u \in V$ let

$$\|u\|_V^2 := \sum_{T \in \mathcal{T}} \|u\|_{m,2,0;T}^2.$$

This expression is assumed to define a norm on $V$, which means that this finite element space is an appropriate approximation of $H_0^{m,2}(\Omega)$. In the sequel we assume that it is sufficient to require that

$$V_k := \text{span}\{\Phi_q^k : q \in (\mathcal{N}_k \backslash \partial \mathcal{N}_k)\},$$

where $\partial \mathcal{N}_k := \{q \in \mathcal{N}_k : p \in \partial \Omega\}$. Note that $\|\cdot\|_V$ is also a norm on $V_k$. Moreover, for any $\Omega' \subset \Omega$ we define

$$\|u\|_{V;\Omega'}^2 := \sum_{T \in \mathcal{T}} \|u\|_{m,2,0;T \cap \Omega'}^2.$$

For later use we state the scaling properties of the norm of the basis functions. For this purpose, consider $\Phi_q^k$ on $T \in \mathcal{T}_k$. Defining

$$\widehat{\Phi}_q^k(\hat{x}) := d_T^{-\nu} \Phi_q^k(d_T \hat{x}), \qquad \hat{x} \in \widehat{T},$$

we obtain

$$\|\Phi_q^k\|_{V;T}^2 \sim d_T^{n+2(\nu-m)} \|\widehat{\Phi}_q^k\|_{m,2,0;\widehat{T}}^2.$$

Note that $\|\widehat{\Phi}_q^k\|_{m,2,0;\widehat{T}}^2 \sim 1$.

## 3. THE HIERARCHICAL STRUCTURE

**a. The hierarchical structure on the nodes.** In the following we want to construct a hierarchical order on $\mathcal{N} = \mathcal{N}_J$. Therefore, consider again $E$, the unit simplex. Applying one step of refinement, we obtain a discretization $\mathcal{T}_E$ and a set of nodes $\mathcal{N}_E'$ on $E$ (see §1). Consider the set of injective mappings $\pi: \mathcal{N}_E \to \mathcal{N}_E'$ of the form $(p, \nu, r) \mapsto (\tilde{\pi}(p), \nu, r)$ (for appropriate mappings $\tilde{\pi}$). With this, we get a set of similar mappings on

$$\mathcal{N}_{k;T} := \{q \in \mathcal{N}_k : p \in T\}$$

for each $T \in \mathcal{T}$. Now we fit these mappings together in a suitable way to obtain a well-defined injective mapping $\pi_k: \mathcal{N}_k \to \mathcal{N}_{k+1}$. As a consequence, we have that $\pi_k$ is the identity on each vertex, and if a node $q$ is located on a face of a simplex $T \in \mathcal{T}_k$, then the same holds for $\pi_k(q)$.

Having defined $\pi_k$ for $k = 0, \ldots, J - 1$, we now define

$$\widetilde{\mathcal{N}_k} := \{q \in \mathcal{N} : q = \pi_{J-1} \circ \cdots \circ \pi_k(q'), \quad q' \in \mathcal{N}_k\}$$

for $k = 0, \ldots, J - 1$. Let $\widetilde{\mathcal{N}_J} := \mathcal{N}$. Noting that $\widetilde{\mathcal{N}_0} \subset \cdots \subset \widetilde{\mathcal{N}_J}$, we define the nonempty sets of *nodes of the kth level* by $\widetilde{\mathcal{R}_k} := \widetilde{\mathcal{N}_k} \backslash \widetilde{\mathcal{N}_{k-1}}$ for $k = 1, \ldots, J$. In particular, $\widetilde{\mathcal{R}_0} := \widetilde{\mathcal{N}_0}$. This leads to a decomposition of $\mathcal{N}$ into disjoint sets

$$\mathcal{N} = \bigcup_{k=0}^{J} \widetilde{\mathcal{R}_k},$$

or more generally, $\widetilde{\mathcal{N}_k} = \bigcup_{l=0}^{k} \widetilde{\mathcal{R}_l}$. In general, $\widetilde{\mathcal{N}_k} \neq \mathcal{N}_k$, because $\widetilde{\mathcal{N}_k}$ may differ from $\mathcal{N}_k$ in the first argument (compare with the construction of $\pi$ above). But using the mappings $\pi_k$, we may identify both. (More precisely, we may identify the set of points with the corresponding set of indices and put the numbering in a hierarchical order; this is what is done in applications.) Therefore, we write in the sequel $\widetilde{\mathcal{N}_k} = \mathcal{N}_k$ and $\widetilde{\mathcal{R}_k} = \mathcal{R}_k$.

**b. The hierarchical structure on the basis functions.** Let $0 \leq k \leq J - 1$. On $V_{k+1}$ we define the following functions:

$$(3.1)$$
$$\overline{\Phi}_q^{k+1} := \sum_{\overline{q} \in \mathcal{N}_{k+1}} \zeta_{q\overline{q}}^k \Phi_{\overline{q}}^{k+1} \qquad \text{for } q \in \mathcal{R}_{k+1},$$

$$\overline{\Phi}_q^k := \Phi_q^{k+1} + \sum_{\overline{q} \in \mathcal{R}_{k+1}} \zeta_{q\overline{q}}^k \Phi_{\overline{q}}^{k+1} \quad \text{for } q \in \mathcal{N}_k,$$

where $\zeta_{q\overline{q}}^k \neq 0$ implies that $\{q, \overline{q}\} \subset T$ for some $T \in \mathcal{T}_k$. Concerning the coefficients, we assume the following:

(A1) (*definiteness*) For arbitrary $u \in V_{k+1}$ there are uniquely defined coefficients $\overline{\alpha}_q$ and $\beta_q$ such that

$$u = \sum_{q \in \mathcal{N}_{k+1}} \alpha_q \Phi_q^{k+1} = \sum_{q \in \mathcal{N}_k} \overline{\alpha}_q \overline{\Phi}_q^k + \sum_{q \in \mathcal{R}_{k+1}} \beta_q \overline{\Phi}_q^{k+1}.$$

(A2) (*consistency*) For $T \in \mathcal{T}_k$, let $w \in \mathbb{P}_{m-1}(T)$. We know that $w \in V_{k+1}|_T$. Expressing $w$ in the new basis, we obtain that $w \in \text{span}(\overline{\Phi}_q^k : q \in \mathcal{N}_{k;T})$ on $T$ and, moreover,

$$w = \sum_{q \in \mathcal{N}_{k;T}} \gamma_q \overline{\Phi}_q^k = \sum_{q \in \mathcal{N}_{k;T}} \gamma_q \Phi_q^k \quad \text{on } T.$$

(A3) (*stability*) For $u$ as in (A1) and $T \in \mathcal{T}_k$ it is assumed that if $q \in \mathcal{N}_{k;T}$, then

$$|\overline{\alpha}_q| \leq \max_{\overline{q} \in \mathcal{N}_{k+1;T}} |\alpha_{\overline{q}}|.$$

(A4) (*scale independency*) The coefficients in definition (3.1) do not depend on the diameter of $T$ after having applied the transformation

$$T \mapsto d_T^{-1} T, \quad \Phi_q^k(x) \mapsto d_T^{-\nu} \Phi_q^k(d_T \hat{x}) \qquad (x = d_T \hat{x}).$$

In this sense, they are assumed to vary only in a bounded subset of $\mathbb{R}$ (independently of $k$).

We now define a new basis on $V$. This is done recursively. Let formally $\overline{\Psi}_q^J = \Phi_q$ for all $q \in \mathcal{N}$. Then for $k = J-1, \ldots, 0$, set

$$\Psi_q^{k+1} := \sum_{\overline{q} \in \mathcal{N}_{k+1}} \zeta_{q\overline{q}}^k \overline{\Psi}_{\overline{q}}^{k+1} \qquad \text{for } q \in \mathcal{R}_{k+1},$$

(3.2)

$$\overline{\Psi}_q^k := \overline{\Psi}_q^{k+1} + \sum_{\overline{q} \in \mathcal{R}_{k+1}} \zeta_{q\overline{q}}^k \overline{\Psi}_{\overline{q}}^{k+1} \quad \text{for } q \in \mathcal{N}_k,$$

with the coefficients given in (3.1) and fulfilling (A1)–(A4).

We now express $u$ recursively in the new basis. Starting with

$$u = \sum_{q \in \mathcal{N}_J} \alpha_q \Phi_q = \sum_{q \in \mathcal{N}_{J-1}} \overline{\alpha}_q \overline{\Psi}_q^{J-1} + \sum_{q \in \mathcal{R}_J} \beta_q \Psi_q^J,$$

we obtain for $0 < k \leq J$

$$u = \sum_{q \in \mathcal{N}_{k-1}} \overline{\alpha}_q \overline{\Psi}_q^{k-1} + \sum_{l=k}^J \sum_{q \in \mathcal{R}_l} \beta_q \Psi_q^l.$$

Defining $\overline{\Psi}_q^0 = \Psi_q^0$, we get the representation

$$u = \sum_{l=0}^J \sum_{q \in \mathcal{R}_l} \beta_q \Psi_q^l.$$

This new basis will be called the *hierarchical basis*. This recursive process also shows that the matrices $B$ and $B^{-1}$ (where $B^{-1}$ transforms the original into the hierarchical basis) can be represented as a product of "simpler" matrices $B_k$, $k = 0, \ldots, J$. Here, "simpler" means sparse and easy to invert (this is due to the required localization of the coefficients in (3.1)). This is of great importance for the application of this basis transformation as a preconditioner.

We are led to the following definitions for $0 \leq k \leq J$:

$$\widetilde{V}_k := \operatorname{span}\{\Psi_q^k : 0 \leq l \leq k, \ q \in \mathcal{R}_l\}, \qquad \widetilde{W}_k := \operatorname{span}\{\Psi_q^k : q \in \mathcal{R}_k\},$$

$$\widetilde{L}_k u := \sum_{q \in \mathcal{R}_k} \beta_q \Psi_q^k, \qquad \widetilde{I}_k u := \sum_{l=0}^k \widetilde{L}_l u.$$

One verifies readily that $V = \bigoplus_{l=0}^J \widetilde{W}_l$, $\widetilde{I}_J u = u$, and $\widetilde{L}_k u = \widetilde{I}_k u - \widetilde{I}_{k-1} u$ for all $u \in V$ and $1 \leq k \leq J$.

**c. Example.** Let $\mathcal{P}_E$ be the set of vertices of $E$ and let the spaces $S_k$ be conforming. Then if

$$\zeta_{q\overline{q}}^k = \begin{cases} \delta_{q\overline{q}} & \text{for } q \in \mathcal{R}_{k+1}, \\ F_{\overline{q}}^k(\Phi_q^k) & \text{for } q \in \mathcal{N}_k, \ \overline{q} \in \mathcal{R}_{k+1}, \end{cases}$$

the assumptions (A1)–(A4) are satisfied.

## 4. The norm of the hierarchical basis functions

In this section we study the dependence of $\|\Psi_q^k\|_V^2$ on the level index $k$. Therefore, we introduce the operator $Q_k : \widetilde{W}_k \to W_k$ by

$$\sum_{q \in \mathcal{R}_k} \beta_q \Psi_q^k \mapsto \sum_{q \in \mathcal{R}_k} \beta_q \Phi_q^k.$$

Now we prove the following

**Lemma.** *Let* $u \in \widetilde{W}_k$ *be given. Then there exist constants* $\sigma$ *and* $\tau$ *(not depending on* $k$ *and* $J$*) such that*

$$\|u\|_V^2 \leq \sigma\tau^{J-k}\|Q_k u\|_V^2.$$

*Proof.* Let $0 \leq k \leq J-1$ and $T \in \mathscr{T}_k$. We know that for $w \in \mathbb{P}_{m-1}(T)$

$$w = \sum_{q \in \mathscr{N}_{k;T}} \gamma_q \overline{\Phi}_q^k = \sum_{q \in \mathscr{N}_{k;T}} \gamma_q \Phi_q^k.$$

If, for arbitrary $u \in V_k$, we define a mapping by

$$u = \sum_{q \in \mathscr{N}_k} \omega_q \Phi_q^k \mapsto \tilde{u} := \sum_{q \in \mathscr{N}_k} \omega_q \overline{\Phi}_q^k,$$

we have just seen that $u = \tilde{u}$ on $T$ for $u \in \mathbb{P}_{m-1}(T)$. Using the scaling required in assumption (A4) of §3, we can derive the estimate

$$\|\tilde{u}\|_{V;T}^2 \leq \tau(\widehat{T}, \hat{\zeta})\|u\|_{V;T}^2,$$

because we showed that $\|u\|_{V;T}^2 = 0$ implies $\|\tilde{u}\|_{V;T}^2 = 0$. Note that $V_k|_T$ is of uniformly bounded dimension. In the estimate above, $\hat{\zeta}$ stands for the set of coefficients $\zeta_{q\bar{q}}^k$ involved in the definitions of $\overline{\Phi}_q^k$ and $\overline{\Phi}_q^{k+1}$ but scaled to $\widehat{T}$. Recall that it is assumed in (A4) that these do not depend on the size of $T$. Take $\tau$ to be the maximum of all $\tau(\widehat{T}, \hat{\zeta})$ which can occur for $T \in \mathscr{T}_k$, $k = 0, \dots, J$. As a consequence, we get the estimate

$$\|\tilde{u}\|_{V;T}^2 \leq \tau\|u\|_{V;T}^2.$$

On $\widetilde{V}_k$, viewed as $\text{span}(\overline{\Psi}_q^k : q \in \mathscr{N}_k)$, we define the mapping $\overline{Q}_k : \widetilde{V}_k \to V_k$ by

$$\sum_{q \in \mathscr{N}_k} \omega_q \overline{\Psi}_q^k \mapsto \sum_{q \in \mathscr{N}_k} \omega_q \Phi_q^k.$$

Now let $k = J-1$ and $T \in \mathscr{T}_{J-1}$. The first part of the proof shows that for $u \in \widetilde{V}_{J-1}$

$$\|u\|_{V;T}^2 \leq \tau\|\overline{Q}_{J-1} u\|_{V;T}^2.$$

Summing over all $T \in \mathscr{T}_{J-1}$ gives

$$\|u\|_V^2 \leq \tau\|\overline{Q}_{J-1} u\|_V^2.$$

Let us now assume that the estimate

$$\|u\|_{V;T}^2 \leq \lambda_k\|\overline{Q}_k u\|_{V;T}^2$$

holds for $1 \leq k < J-1$, $T \in \mathscr{T}_k$, and $u \in \widetilde{V}_k$. Consider $u \in \widetilde{V}_{k-1}$ and $T \in \mathscr{T}_{k-1}$. Using (3.2), we can rewrite $u$ as a function in $\widetilde{V}_k$. But then the foregoing assumption can be applied, and the estimate

$$\|u\|_{V;T}^2 \leq \lambda_k\|\overline{Q}_k u\|_{V;T}^2$$

follows. We can now apply the first step of the induction again to obtain

$$\|\overline{Q}_k u\|_{V;T}^2 \leq \tau\|\overline{Q}_{k-1} u\|_{V;T}^2.$$

Therefore, $\lambda_{k-1} = \tau\lambda_k$ and $\lambda_{J-1} = \tau$. Summation over all $T \in \mathscr{T}_{k-1}$ yields

$$\|u\|_V^2 \leq \tau^{J-k}\|\overline{Q}_k u\|_V^2.$$

Now let $u \in \widetilde{W}_k$, that is, $u = \sum_{q \in \mathscr{R}_k} \beta_q \Psi_q^k$. Utilizing (3.2), we rewrite $u$ as a function on $\widetilde{V}_k$. By the preceding result we have

$$\|u\|_V^2 \leq \tau^{J-k}\|\overline{Q}_k u\|_V^2 = \tau^{J-k}\left\|\sum_{q \in \mathscr{R}_k} \beta_q \overline{\Phi}_q^k\right\|_V^2.$$

So it remains to estimate the norm of the mapping

$$v = \sum_{q \in \mathscr{R}_k} \omega_q \Phi_q^k \mapsto \overline{v} = \sum_{q \in \mathscr{R}_k} \omega_q \overline{\Phi}_q^k.$$

If the right-hand side is a polynomial, it is zero because of (A2). Using the same arguments as at the beginning of this section yields

$$\|\overline{v}\|_{V;T}^2 \leq \sigma\|v\|_{V;T}^2,$$

where $\sigma$ can be chosen independently of all $T$, $k$, and coefficients $\zeta_{q\overline{q}}^k$. This gives

$$\|u\|_V^2 \leq \sigma\tau^{J-k}\|Q_k u\|_V^2. \quad \square$$

## 5. Estimates for the projection operators

Let us define the operator $L_k := Q_k\widetilde{L}_k \colon V \to W_k$. In the following we give two estimates for $L_k u$.

**a. Spectral estimate for $\|L_k u\|_V^2$.** Let

$$v := L_k u = \sum_{q \in \mathscr{R}_k} \beta_q \Phi_q^k.$$

Fix $T \in \mathscr{T}_{k-1}$ and consider $v|_T$. If $v \in \mathbb{P}_{m-1}(T)$, we conclude from (A2) that $v = 0$. This shows that $\|v\|_{V;T}^2$ defines a norm on the coefficients of $v|_T$. Scaling to $\widehat{T}$ yields

$$\hat{v} = \sum_{q \in \mathscr{R}_{k;T}} \beta_q d_T^\nu \widehat{\Phi}_q^k.$$

Recall that $\|\widehat{\Phi}_q^k\|_{m,0;\widehat{T}}^2 \sim 1$. Therefore, there are constants depending only on $\widehat{T}$ and the number of coefficients of $v|_T$ such that

$$\|\hat{v}\|_{m,2,0;\widehat{T}}^2 \sim \sum_{q \in \mathscr{R}_{k;T}} d_T^{2\nu}|\beta_q|^2.$$

Rescaling establishes

$$\|v\|_{V;T}^2 \sim \sum_{q \in \mathscr{R}_{k;T}} d_T^{2(\nu-m)+n}|\beta_q|^2.$$

If we introduce the vector norm

$$|z|_{*;T}^2 := \sum_{q \in \mathscr{R}_{k;T}} d_T^{2(\nu-m)+n}|\beta_q|^2$$

for the vector representation $z$ of $u$, this result can be written as $\|L_k u\|_{V;T}^2 \sim |z|_{*;T}^2$.

**b. The operator norm of $L_k$.** Let $u \in V$ and $T \in \mathscr{T}_{k-1}$. We have proved in the first part that $\|L_k u\|_{V;T}^2 \le C|z|_{*;T}^2$ for $z$ the representation of $u$ in the hierarchical basis. Recall how the basis transformation to the hierarchical basis is performed and that it is stable in the $L^\infty$-norm by (A3). Using this, we can derive for each $\overline{\alpha}_q$, $q \in \mathscr{N}_k$, that

$$|\overline{\alpha}_q| \le \max_{\overline{q} \in \mathscr{N}_{J;T}} |\alpha_{\overline{q}}|,$$

and from the last transformation step we obtain

$$|\beta_q| \le C d_T^{-\nu} \sum_{\overline{\nu}=0}^{m-1} d_T^{\overline{\nu}} \max_{\overline{q} \in \mathscr{N}_{J;T}, \overline{q}=(\overline{\nu},...)} |\alpha_{\overline{q}}|,$$

where we used (A4). The constant in this estimate depends on the coefficients $\hat{\zeta}_{q\overline{q}}^k$. From this, we obtain immediately

$$|z|_{*;T}^2 \le C \sum_{\nu=0}^{m-1} d_T^{2(\nu-m)+n} \|\nabla^\nu u\|_{0,\infty;D_\nu}^2$$

for some appropriate triangles $D_\nu \in \mathscr{T}_J$, $D_\nu \subset T$.

Let us first consider the case of conforming elements. In this case we refer to an estimate that is well known (e.g., [9]; see also Lemma 2 in the supplement of this issue). It states that

$$(5.1) \qquad \left( \frac{1}{|D|} \int_D |v| \right)^2 \le C d_T^{2-n} \kappa \left( \frac{d_T}{d_D} \right) \|v\|_{1,2,*;T}^2$$

holds for any $v \in H_0^{1,2}(T)$, where for $\lambda \ge 1$

$$\kappa(\lambda) = \begin{cases} 1 + \log(\lambda), & n = 2, \\ \lambda^{n-2}, & n > 2. \end{cases}$$

Applying this estimate to $\nabla^\nu v$ instead of $v$, for $0 \le \nu \le m-1$, yields

$$(5.2) \qquad \sum_{\nu=0}^{m-1} d_T^{2(\nu-(m-1))} \left( \frac{1}{|D|} \int_D |\nabla^\nu v| \right)^2 \le C d_T^{2-n} \kappa \left( \frac{d_T}{d_D} \right) \|v\|_{m,2,*;T}^2.$$

In case of a nonconforming space we assume that such an estimate generally holds. But for a certain class of elements, Lemma 3 in the supplement provides a proof that such an estimate can be obtained.

The usual scaling technique shows that

$$\frac{1}{|D|} \int_D |\nabla^\nu u| \sim |u|_{\nu,\infty;D}.$$

Thus, we get

$$|z|_{*;T}^2 \le C \max_{D \in \mathscr{T}_J, D \subset T} \left\{ \kappa \left( \frac{d_T}{d_D} \right) \right\} \|u\|_{V,*;T}^2.$$

For arbitrary $w \in \mathbb{P}_{m-1}(T)$, $T \in \mathscr{T}_{k-1}$, we have $L_k u = L_k(u - w)$, and therefore

$$\|L_k u\|_{V;T}^2 \le C \max_{D \in \mathscr{T}_J, D \subset T} \left\{ \kappa \left( \frac{d_T}{d_D} \right) \right\} \|u\|_{V;T}^2,$$

using the Poincaré inequality for functions having mean value zero (see the remark following Lemma 1 in the supplement).

## 6. Estimates for the eigenvalues

**a. The minimal eigenvalue.** Consider $T \in \mathscr{T}_{k-1}$ for $k = 1, \dots, J$. Utilizing the results of §5 yields

$$|z|^2_{*;T} \leq C \max_{D \in \mathscr{T}_J, D \subset T} \left\{ \kappa \left( \frac{d_T}{d_D} \right) \right\} \|u\|^2_{V;T},$$

and summing over all $T \in \mathscr{T}_{k-1}$ gives

$$\sum_{T \in \mathscr{T}_k} |z|^2_{*;T} \leq C \max_{T \in \mathscr{T}_k} \max_{D \in \mathscr{T}_J, D \subset T} \left\{ \kappa \left( \frac{d_T}{d_D} \right) \right\} \|u\|^2_V.$$

We introduce for $k \geq 1$

$$d_q := \max\{d_T : T \in \mathscr{T}_{k-1}, \ q \in T\},$$

sum over $k$, and obtain

$$\|L_0 u\|^2_V + \sum_{k=1}^{J} \sum_{q \in \mathscr{R}_k} d_q^{2(\nu-m)+n} |\beta_q|^2 \leq C \sum_{k=0}^{J} \|L_k u\|^2_V$$

$$\leq C \sum_{k=0}^{J} \max_{T \in \mathscr{T}_k} \max_{D \in T; D \subset \mathscr{T}} \kappa \left( \frac{d_T}{d_D} \right) \|u\|^2_V =: C \Lambda_J^{-1} \|u\|^2_V.$$

**b. The maximal eigenvalue.** The following estimate holds without any further restrictions:

$$\|u\|^2_V \leq \left\{ \sum_{k=0}^{J} \|\widetilde{L}_k u\|_V \right\}^2 \leq C \sum_{k=0}^{J} \tau^{J-k} \sum_{k=0}^{J} \|L_k u\|^2_V$$

$$\leq C K_{\tau,J} \left\{ \|L_0 u\|^2_V + \sum_{k=1}^{J} \sum_{q \in \mathscr{R}_k} d_q^{2(\nu-m)+n} |\beta_q|^2 \right\}$$

(using §5a), where $K_{\tau,J} := \sum_{k=0}^{J} \tau^k$. If we can prove that for $j > k$ there exists $\sigma < 1$ such that

$$(\widetilde{L}_k u, \widetilde{L}_j u)_V \leq C \sigma^{j-k} \|u\|_V \|v\|_V$$

for all $u, v \in V$, the previous estimate can be slightly improved. Using the arguments in [9], we obtain

$$\|u\|^2_V \leq C \max(1, \tau^J) \sum_{k=0}^{J} \|L_k u\|^2_V.$$

## 7. Results

The previous sections are summarized in the following theorem.

**Theorem.** *Let $\Omega$ be a polygonally bounded domain and $(\mathscr{T}_k)_{k=0,J}$ a sequence of discretizations of $\Omega$ fulfilling the assumptions stated in §1. Equip the discretizations with finite element spaces as described in §2.*

*Assume that there exists a hierarchical structure fulfilling assumptions (A1)–(A4) of §3b and that an estimate of the form (5.2) can be established in case of nonconforming elements. Then there exists a basis transformation $B$ that can be represented as a product of matrices with small bandwidth and a diagonal matrix $M$, such that for $S$ being the matrix in (0.2),*

$$(7.1) \qquad \mathrm{cond}(M^*B^*SBM) \le CK_{\tau,J}\Lambda_J.$$

*The constant $C$ depends on scale-invariant properties of the triangulation and the conditioning of the stiffness matrix on the coarsest level. Moreover,*

$$\Lambda_J = \begin{cases} O((J+1)^2), & n=2, \\ O(2^{J(n-2)}), & n>2, \end{cases}$$

*and $K_{\tau,J} = \sum_{k=0}^{J}\tau^k$, for $\tau$ as described in §4. If we have, in particular, a uniformly refined discretization consisting of simplices of approximately the same size, and if $h = 2^{-J}$ is the average diameter of the simplices of the finest level, then (with $\mathrm{ld}$ denoting the logarithm to base 2)*

$$\Lambda_J = \begin{cases} O(|\mathrm{ld}(h)|^2), & n=2, \\ O(h^{-(n-2)}), & n>2, \end{cases} \qquad K_{\tau,J} = \begin{cases} O(|\mathrm{ld}(h)|), & \tau=1, \\ O(h^{-\mathrm{ld}(\tau)}), & \tau>1. \end{cases}$$

*Under the additional assumption stated in §6b, one has the improved estimate*

$$K_{\tau,J} = \max(1,\tau^J),$$

*so that for $\tau=1$ we have $K_{\tau,J}=1$.*

*Proof.* Using the notation of §6, we have already shown that

$$c\Lambda_J^{-1}\left(\|L_0 u\|_V^2 + \sum_{k=1}^{J}\sum_{q\in\mathscr{R}_k} d^{2(\nu-m)+n}|\beta_q|^2\right) \le \|u\|_V^2$$

and

$$\|u\|_V^2 \le CK_{\tau,J}\left(\|L_0 u\|_V^2 + \sum_{k=1}^{J}\sum_{q\in\mathscr{R}_k} d_q^{2(\nu-m)+n}|\beta_q|^2\right).$$

Take $z$ to be the vector of the coefficients for $u$ in the hierarchical basis. Recall that $\|u\|_V^2 \sim z^*B^*SBz$, where $S$ is the stiffness matrix for the original basis.

With constants depending on the smallest and largest eigenvalue of the problem on the coarsest level and an appropriate diagonal matrix $M$, we derive

$$c\Lambda_J^{-1}|z|^2 \le z^*M^*B^*SBMz \le CK_{\tau,J}|z|^2.$$

Notice that by the definition of the refinement procedure we have for $T \in \mathscr{T}_k$ and $D \in \mathscr{T}_J$ with $D \subset T$

$$\frac{d_T}{d_D} \le C2^{J-k}. \qquad \square$$

*Remarks.* (i) As proposed in [9], one can invert $S_0$, the stiffness matrix for the coarsest problem, to improve the constants in the estimate above. In this case, replace the submatrix in $M$ corresponding to the coefficients for the coarsest level by $S_0^{-1}$.

(ii) We observe (as was known before [9]), that the improvement in the behavior of the condition number in two space dimensions is quite better than in three dimensions. Note that the conditioning of the stiffness matrix in (0.2) does not depend on the space dimension. On the other hand, the condition number due to the hierarchical basis does not depend on the order of the elliptic problem. Therefore, one may obtain satisfactory improvement for 4th-order problems in three dimensions (if $\tau$ is close to 1).

(iii) If a conforming element is a polynomial when restricted to a single simplex and if we use the hierarchical structure proposed in §3c, we find that $\tau = 1$, because we can exactly reproduce the basis functions on the next level. As an example, we refer to the case of the linear element [9] (see also [5]) and as a contrast to the RHCT-element [4], where $\tau \approx 2$.

(iv) As one can see in §4, $\tau$ can be obtained by means of local considerations. In practice it may be enough to consider only a few situations to obtain a reasonable value for $\tau$. For this purpose we consider $T$, the refined simplex $T'$, and the corresponding stiffness matrices $S$ and $S'$ (restricted to $T$). Given the coefficients of the basis transformation, we can compute $\widetilde{S}$, the stiffness matrix with respect to the basis functions of the coarse level. $\tau$ is now the maximal eigenvalue of the matrix $S^{-1}\widetilde{S}$. But $S$ and $\widetilde{S}$ are not bijective, since $Sz = 0$ if $z$ corresponds to a polynomial function. However, both matrices are symmetric, and therefore we can restrict ourselves to the subspace $\text{Image}(S)$. Notice that we have required that $\text{Ker}(S) = \text{Ker}(\widetilde{S})$.

(v) The assumption that $\Omega$ is a polygonally bounded domain can be removed. Assume that $\partial\Omega$ is sufficiently smooth. We start again with a macro-discretization $\mathscr{T}_0$ and perform a complete refining step to obtain $\mathscr{T}_1$. Now we improve the approximation of the boundary of $\Omega$ by modifying $\mathscr{T}_1$ to a discretization $\mathscr{T}_1^*$ (see, for example, [4, Chapter 4.3]). Then we refine $\mathscr{T}_1^*$ to obtain $\mathscr{T}_2$ and modify $\mathscr{T}_2$ to get $\mathscr{T}_2^*$, and so on. We introduce a hierarchical structure as before. In order to compare the spaces $\widetilde{V}_k$ to the spaces $V_k$, we have to take into account the modification $\mathscr{T}_k \to \mathscr{T}_k^*$, that is, we have to compare $\widetilde{V}_k$ with $V_k^*$ and $V_k^*$ with $V_k$. Notice that even in the case of conforming linear elements this will produce a defect $(\tau > 1)$. However, using the regularity assumption on $\partial\Omega$, we notice that in each step we get an additional factor of the form $1 + ch_k$ for $0 \le k \le J$. But the product of these factors is uniformly bounded, and we end up with the same results as presented here.

## 8. THE NONCONFORMING LINEAR ELEMENT

Here we want to discuss the case of nonconforming linear elements in $\mathbb{R}^2$. Let $(\mathscr{T}_k)_k$ be a sequence of global refined triangulations (that is, every triangle is partitioned into four similar triangles). A basis function looks as follows:
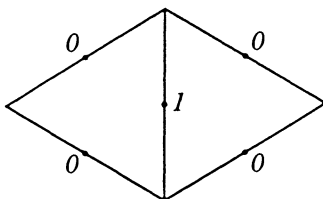
It is continuous on the inner edge (where it is 1) and on the other edges it is only continuous at the midpoints. The set of all midpoints of the edges belonging to $\Omega$ will be denoted by $\mathcal{M}$. In general, a function in $V$ is only continuous at points in $\mathcal{M}$. In this example, the set of nodes is given by
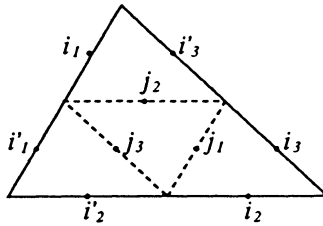
$$\mathcal{N} = \{(p, 0, 0) : p \in \mathcal{M}\}.$$

The problem mainly consists in finding a hierarchical structure on $\mathcal{M}$.

**a. The hierarchical structure.** Assume that an orientation is given for each edge. The definition of $\pi$ is shown in the following picture:
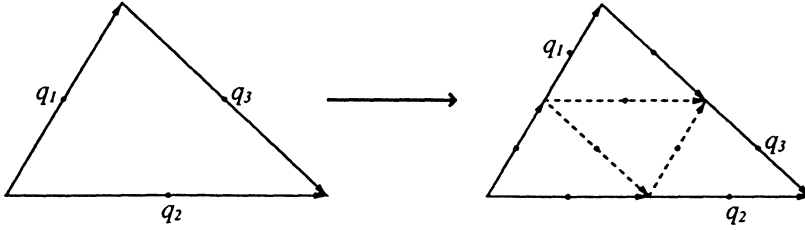


As explained in §3a, this will define a hierarchical structure on the nodes. Consider the following triangle:



Here,

$i_1, i_2, i_3$      are the nodes on the next higher level and

$i'_1, i'_2, i'_3, j_1, j_2, j_3$    are the nodes on the actual level.

The new basis functions (restricted to one triangle) are given by

$$\overline{\Phi}_{i_1} := \Phi_{i_1} + \Phi_{i'_1} + \xi(\Phi_{j_1} + \Phi_{j_3}) + \zeta\Phi_{j_2},$$

$$\overline{\Phi}_{i'_1} := \tfrac{1}{2}(\Phi_{i_1} - \Phi_{i'_1}),$$

$$\overline{\Phi}_{i_2} := \Phi_{i_2} + \Phi_{i'_2} + \xi(\Phi_{j_1} + \Phi_{j_2}) + \zeta\Phi_{j_3},$$

$$\overline{\Phi}_{i'_2} := \tfrac{1}{2}(\Phi_{i_2} - \Phi_{i'_2}),$$

$$\overline{\Phi}_{i_3} := \Phi_{i_3} + \Phi_{i'_3} + \xi(\Phi_{j_2} + \Phi_{j_3}) + \zeta\Phi_{j_1},$$

$$\overline{\Phi}_{i'_3} := \tfrac{1}{2}(\Phi_{i_3} - \Phi_{i'_3}).$$

If we require that $2\xi + \zeta = 1$, it is easily seen that this definition fulfills the assumptions (A1)–(A4). According to this definition, the transformation from the standard to the hierarchical basis looks like

$$\overline{\alpha}_{i_1} = \tfrac{1}{2}(\alpha_{i_1} + \alpha_{i'_1}), \ldots,$$

$$\beta_{i'_1} = \alpha_{i_1} - \alpha_{i'_1}, \ldots,$$

$$\beta_{j_1} = \alpha_{j_1} - \xi(\overline{\alpha}_{i_1} + \overline{\alpha}_{i_2}) - \zeta\overline{\alpha}_{i_3}, \ldots.$$

The dots refer to analogous relations for the other coefficients.

In proving the estimate (7.1) for the case of nonconforming elements we have to show that the estimate (5.1) holds. This is proved in Lemma 3 in the supplement.

**b. Numerical results.** To give an example, we restrict ourselves to a triangulation of $\Omega = [0, 1] \times [0, 1]$ consisting entirely of triangles having two sides of equal length enclosing an angle of $90°$. Calculating $\tau$ numerically gives the following results:

| $\xi$ | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|
| $\tau$ | 4.00 | 3.00 | 2.32 | 2.02 | 2.08 | 2.5 |
| $\mathrm{ld}(\tau)$ | 2.00 | 1.58 | 1.21 | 1.01 | 1.06 | 1.32 |

The smallest and the largest eigenvalue were computed by vector iteration and inverse vector iteration, respectively. Furthermore, the Dirichlet problem

$$\Delta u = 0 \qquad \text{in } \Omega,$$
$$u(x, y) = x^2 - y^2 \quad \text{in } \partial\Omega$$

was solved (*it* denotes the number of iterations). PGG I refers to our hierarchical basis preconditioning with $\xi = 0.4$, while PCG II stands for the corresponding preconditioning with $\xi = 0.5$.

| | CG | | | PCG I | | | PCG II | | |
|---|---|---|---|---|---|---|---|---|---|
| $J$ | $\lambda_{\min}$ | $\lambda_{\max}$ | $it$ | $\lambda_{\min}$ | $\lambda_{\max}$ | $it$ | $\lambda_{\min}$ | $\lambda_{\max}$ | $it$ |
| 2 | 0.43 | 11.5 | 10 | 1.2 | 18 | 14 | 1.3 | 24 | 14 |
| 3 | 0.10 | 11.7 | 30 | 0.85 | 28 | 23 | 0.93 | 42 | 25 |
| 4 | 0.025 | 11.8 | 61 | 0.66 | 44 | 36 | 0.72 | 74 | 39 |
| 5 | 0.006 | 11.8 | 121 | 0.53 | 69 | 48 | 0.59 | 130 | 58 |
| 6 | – | – | 235 | – | – | 67 | – | – | 87 |

If $c_k$ is the condition number of the $k$th refinement, then

$$E_k := \frac{\log(c_k/c_{k-1})}{\log(N_k/N_{k-1})}$$

is an approximation for the exponent $\nu$ in the assumed law $c_k = C N_k^\nu$ ($N_k = \dim(V_k)$). For PCG I we get $E_5 \approx 0.95$ and for PCG II we get $E_5 \approx 1.1$. For the 6th refinement, the computation time for solving the Dirichlet problem with PCG I preconditioning is about 0.6 times the time needed for the ordinary CG algorithm.

## BIBLIOGRAPHY

1. O. Axelsson and V. A. Barker, *Finite element solutions of boundary value problems*, Academic Press, London, 1984.

2. E. Bänsch, *Local mesh refinement in 2 and 3 dimensions*, SFB 256, report no. 6, Universität Bonn, 1989.

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

4. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, New York, 1978.

5. W. Dörfler, *The conditioning of the stiffness matrix for certain elements approximating the incompressibility condition in fluid dynamics*, Numer. Math. **58** (1990), 203–214.

6. R. Glowinski, *Numerical methods for nonlinear variational problems*, Springer, New York, 1984.

7. W. Hackbusch, *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner, Stuttgart, 1986.

8. M. C. Rivara, *Algorithms for refining triangular grids suitable for adaptive and multigrid techniques*, Internat. J. Numer. Methods Engrg. **20** (1984), 745–756.

9. H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math. **49** (1986), 379–412.

10. ____, *Two preconditioners based on the multi-level splitting of finite element spaces*, manuscript, Universität Dortmund, 1990.

INSTITUT FÜR ANGEWANDTE MATHEMATIK, E.T.H.-ZÜRICH, RÄMISTRASSE 74, CH-8001 ZÜRICH, SWITZERLAND