MATHEMATICS OF COMPUTATION
VOLUME 64, NUMBER 212
OCTOBER 1995, PAGES 1397–1419

ON THE CONVERGENCE RATE OF THE CELL DISCRETIZATION ALGORITHM FOR SOLVING ELLIPTIC PROBLEMS

MARIA CAYCO, LESLIE FOSTER, AND HOWARD SWANN

ABSTRACT. Error estimates for the cell discretization algorithm are obtained for polynomial bases used to approximate both $H^k(\Omega)$ and analytic solutions to selfadjoint elliptic problems. The polynomial implementation of this algorithm can be viewed as a nonconforming version of the *h-p* finite element method that also can produce the continuous approximations of the *h-p* method. The examples provided by our experiments provide discontinuous approximations that have errors similar to the finite element results.

INTRODUCTION

This paper concerns a nonconforming version of the finite element method for approximating solutions of elliptic partial differential equations, where the requirement that an approximation be continuous is weakened. We discuss the *cell discretization* algorithm (abbreviated as CDA) formulated by Greenstadt [9, 12, 15–18]; the domain of a problem is partitioned into *cells*, approximations are made on each cell, and the approximations are forced to be weakly continuous across the boundaries of each cell using a method called *moment collocation.* Convergence of the Greenstadt method occurs in quite general situations [24, 25]. The cells do not necessarily diminish in size, and approximations to the solution on each cell can be constructed using any suitably smooth basis. Babuška uses a method similar to moment collocation to make finite element approximations match the boundary data in elliptic problems [2]. See also [8]. Although our error estimates are somewhat different, we give a polynomial implementation that is essentially the *primal hybrid* finite element method of Raviart and Thomas [21], who use finite element approximations that may be discontinuous, yet they converge to a solution as the size of the mesh of the finite element grid becomes small.

In §1 we extend the general results for selfadjoint problems presented in [25] in several ways that are of use later in the paper. Section 2 presents a polynomial implementation of the algorithm for domains in $\mathbb{R}^2$ and $\mathbb{R}^3$. This implementation contains a version of the *h-p* finite element method [3–7, 20] as a special case. We give error estimates that are expressed in terms of the degree of a poly-

Received by the editor June 15, 1993.
1991 *Mathematics Subject Classification.* Primary 65N30; Secondary 65N35, 65N15.
Key words and phrases. Elliptic equations, finite element methods, hybrid methods, nonconforming methods, Lagrange multipliers, domain decomposition, cell discretization.

nomial approximation on each cell, the number of moment collocations used to enforce weak continuity, and the maximum diameter of a cell. In §3 we discuss the linear algebra used in our algorithm and present an example that illustrates the theoretical results. We show that we can relax the requirement that solutions be continuous across cell interfaces and still obtain errors similar to continuous approximations. Our results provide useful information concerning the selections of an appropriate number of moment collocations, basis functions, and cell size in the cell discretiation method.

1. Description of the problem and convergence results

For completeness, we give the following definitions and results from [24, 25].

Let Ω be a bounded domain in $\mathbb{R}^k$ with boundary Γ. We approximate the solution of an elliptic selfadjoint problem of the form

$$\mathbf{E}u = f, \tag{1.1a}$$

$$u = g \quad \text{on } \Gamma, \tag{1.1b}$$

where the operator $\mathbf{E}$ is defined by

$$\mathbf{E}u = -\sum_{i,j}^{K} D_i(A_{ij}(x)D_j u) + A_0(x)u,$$

with D_i representing partial differentiation with respect to x_i.

We use the following variational form of (1.1a) and (1.1b):

Define

$$a(u, v) = \int_\Omega \left(\sum_{i,j}^{K} A_{ij}(x) D_i u D_j v + A_0(x) uv \right) dx.$$

We wish to approximate $u \in H^1(\Omega)$ satisfying $u|_\Gamma = g$ such that

$$a(u, v) = (f, v)$$

for all $v \in H_0^1(\Omega)$, the subspace of functions in $H^1(\Omega)$ whose traces are equal to zero on Γ. The L_2 inner product over Ω is denoted by $(\cdot, \cdot)$, with norm denoted $\|\cdot\|_0$.

The cell discretization algorithm supposes that the domain Ω can be partitioned into subdomains with Lipschitz continuous boundaries that are piecewise C^1 (LPC^1); such subdomains are called *cells*. Suppose there are N cells Ω_i, with $\Omega_i \cap \Omega_j = \varnothing$ if $i \neq j$ and $\overline{\Omega} = \bigcup_{i=1}^{N} \overline{\Omega}_i$. The exterior is $\Omega_0 \equiv \mathbb{R}^K \backslash \overline{\Omega}$.

The Hilbert spaces we use are the following:

$$H^1(\Omega_i) \equiv \{u : \Omega_i \to \mathbb{R} : u \in L_2(\Omega_i);\ D_j u \in L_2(\Omega_i) \text{ for } j = 1, \ldots, K\},$$

where partial derivatives $D_j u$ are distribution derivatives with respect to x_j. The space $H^1(\Omega_i)$ has inner product

$$(u, v)_{1,i} = \sum_{j=1}^{K} (D_j u, D_j v)_i + (u, v)_i,$$

where $(\cdot, \cdot)_i$ represents the $L_2(\Omega_i)$ inner product, with the norm expressed as $\|\cdot\|_{0,i}$. The norm on $H^1(\Omega_i)$ is denoted $\|\cdot\|_{1,i}$. The H^1 inner product and norm over Ω rather than Ω_i are represented by $(\cdot, \cdot)_{1,\Omega}$ and $\|\cdot\|_{1,\Omega}$.

Approximations are in space

$$H \equiv \{u \in L_2(\Omega) : u|_{\Omega_i} \in H^1(\Omega_i);\ i = 1, \ldots, N\}.$$

The Hilbert space H has inner product

$$(u, v)_H \equiv \sum_{i=1}^{N} (u, v)_{1,i},$$

with norm represented by $\|\cdot\|_H$.

Let $\Gamma_{ij} = \overline{\Omega}_i \cap \overline{\Omega}_j$. Assume that Γ_{ij} is the finite union of C^1 patches. To simplify notation, we refer to all such patches as Γ_{ij}, although there may be more than one C^1 component. Γ_{i0} is a boundary segment between Ω_i and Ω_0. (See [25] for a precise definition of these terms.) The inner product for $L_2(\Gamma_{ij})$ is denoted by $\langle \cdot, \cdot \rangle_{ij}$, with norm represented as $\|\cdot\|_{ij}$.

We denote by γ_{ij} the trace operator restricting $u|_{\Omega_i}$ to its values on Γ_{ij}. We regard it as a bounded linear operator from $H^1(\Omega_i)$ to $L_2(\Gamma_{ij})$ [19]; there are constants C_{ij} such that for any $w \in H$, $\|\gamma_{ij}(w)\|_{ij} \leq C_{ij}\|w\|_{1,i}$. Since we are concerned with estimates in terms of $\|\gamma_{ij}(w)\|_{ij}$ rather than the $H^{1/2}(\Gamma_{ij})$ norm of $\gamma_{ij}(w)$ required by full use of the trace theorem [19], the constants C_{ij} can be explicitly obtained for Ω_i with simple kinds of boundaries, and we describe some such C_{ij} in this paper.

For each Γ_{ij}, choose $\{\omega_q^{ij}\}_{q=1}^{\infty}$ to be functions in $H^{1/2}(\Gamma_{ij})$ that are a Schauder basis for $L_2(\Gamma_{ij})$. For any $h \in L_2(\Gamma_{ij})$, there are coefficients d_k such that $h = \sum_{k=1}^{\infty} d_k \omega_k^{ij}$. For any n, let $\mathcal{T}_n^{ij}(h) \equiv \sum_{k=n+1}^{\infty} d_k \omega_k^{ij}$. For any $\varepsilon > 0$, there is some $N(h, \varepsilon)$ such that $n > N(h, \varepsilon)$ implies $\|\mathcal{T}_n^{ij}(h)\|_{ij} < \varepsilon$.

Weak continuity of approximations in H across interfaces Γ_{ij} is enforced by Greenstadt's moment collocation method:

For $u \in H$, we define the kth moment of u on Γ_{ij} to be

$$M_k^{ij}(u) \equiv \langle \gamma_{ij}(u), \omega_k^{ij} \rangle_{ij}.$$

We require that the moments of an approximation u be equal on interfaces Γ_{ij} in the following way.

Let N_I be the number of interfaces Γ_{ij}. We will denote by $[n]$ a multi-index, an N_I-vector of nonnegative integers $(\ldots, n_{ij}, \ldots)$. A partial order is $[n'] \geq [n]$ if and only if for any ij, $n'_{ij} \geq n_{ij}$. We say that $[n^k] \to [\infty]$ if $[n^k] \leq [n^{k+1}]$ and $\inf_{ij}\{n_{ij}^k\} \to \infty$ as $k \to \infty$.

Set

$$G[n] \equiv \{u \in H :\ \text{for any } ij,\ ij = 1, \ldots, N_I \text{ and for any } k \leq n_{ij},\ \text{we have } M_k^{ij}(u) = M_k^{ji}(u)\}.$$

In this case, $[n]$ is the multi-index described above, with all $n_{i0} = 0$, where the n_{i0} refer to the Γ_{i0}. Thus, $G[n]$ is the set of functions u in H such that on any internal interface Γ_{ij}, $\gamma_{ij}(u) - \gamma_{ji}(u)$ is $L_2(\Gamma_{ij})$-orthogonal to ω_k^{ij}, $k = 1, \ldots, n_{ij}$. This gives a notion of weak continuity across interfaces called *moment collocation*: Define

$$G_0[n] = \{u \in G[n] :\ \text{for any } i,\ \text{for any } k \leq n_{i0},\ M_k^{i0}(u) = 0\}.$$

Thus, $G_0[n]$ is the set of functions in $G[n]$ that are weakly 0 on the external interfaces Γ_{i0} making up Γ.

Our approximations of solutions for problems with Dirichlet boundary data g are in

$$D[n] \equiv \{u \in G[n]: \text{ for any } \Gamma_{i0} \neq \varnothing \text{ and } k \leq n_{i0},\ M_k^{i0}(u) = \langle g, \omega_k^{i0}\rangle_{i0}\}.$$

For each ith cell, choose any Schauder basis $\{B_k^i\}$ for $H^1(\Omega_i)$. For any v in $H^1(\Omega_i)$, there are b_k^i such that $\sum_{k=1}^{\infty} b_k^i B_k^i = v$; let $v_{\cdot,m} = \sum_{k=1}^{m} b_k^i B_k^i$. Let $\mathscr{Q}_m^i(v)$ denote the orthogonal projection (in the $H^1(\Omega_i)$ inner product) of v onto the $H^1(\Omega_i)$-orthogonal complement of the span of $\{B_1^i, B_2^i, \ldots, B_m^i\}$. Thus,

$$\mathscr{Q}_m^i(v_{\cdot,m}) = 0; \qquad \mathscr{Q}_m^i(v) = \mathscr{Q}_m^i(v - v_{\cdot,m})$$

and

$$\|\mathscr{Q}_m^i(v)\|_{1,i} \leq \|v - v_{\cdot,m}\|_{1,i} = \left\| \sum_{k=m+1}^{\infty} b_k^i B_k^i \right\|_{1,i}, \qquad \operatorname*{Lim}_{m\to\infty} \|\mathscr{Q}_m^i(v)\|_{1,i} = 0.$$

These properties of $\mathscr{Q}_m^i$ are independent of $[n]$.

Let $[m]$ be an N-dimensional multi-index indicating the number of basis functions used in the approximation on each cell; we employ the same notational conventions as those used for the multi-index $[n]$.

We let $H[m]$ be the subspace of H such that for any $v \in H[m]$, $v|_{\Omega_i}$ is in the span of $\{B_1^i, B_2^i, \ldots, B_{m_i}^i\}$.

Given $[m]$ and any function v in H, $\mathscr{Q}_{[m]}(v)$ is defined to be the function in H such that $\mathscr{Q}_{[m]}(v)|_{\Omega_i} = \mathscr{Q}_{m_i}^i(v|_{\Omega_i})$. Thus, $\mathscr{Q}_{[m]}(\cdot)$ is the projection of H onto $H[m]^{\perp}$. $\operatorname{Lim}_{[m]\to[\infty]} \|\mathscr{Q}_{[m]}(v)\|_H = 0$.

Let

$$G[n][m] = \left\{ u \in G[n] : u|_{\Omega_i} = \sum_{k=1}^{m_i} b_k^i B_k^i \right\},$$

which is a finite-dimensional space; the moment collocation requirements are met by requiring that certain linear equations hold among the b_k^i.

The bilinear form $a(u, v)$ can be extended to H; its restriction to Ω_k is represented as $a(u, v)_k$. Variational methods allow us to approximate the solution by obtaining the function u in $D[n][m] \equiv D[n] \cap G[n][m]$ that minimizes

$$a(u, u) - 2(f, u)$$

over all $u \in D[n][m]$.

If we use the Schauder basis on each cell Ω_k, then

$$\begin{aligned} a(u, u) - 2(f, u) &= \sum_{k=1}^{N} [a(u, u)_k - 2(f, u)_k] \\ &= \sum_{k=1}^{N} \left[\sum_{i=1}^{m_k} b_i^k \sum_{j=1}^{m_k} b_j^k a(B_i^k, B_j^k)_k - 2 \sum_{i=1}^{m_k} b_i^k (f, B_i^k)_k \right]. \end{aligned}$$

This quadratic form is to be minimized subject to the moment collocation constraints. This is done by adding terms of the form

$$-\lambda_q^{ij}(\langle \gamma_{ij}(u), \omega_q^{ij}\rangle_{ij} - \langle \gamma_{ji}(u), \omega_q^{ij}\rangle_{ij}), \qquad q = 1, \ldots, n_{ij},$$

and

$$-\lambda_q^{i0}(\langle \gamma_{i0}(u), \omega_q^{i0}\rangle_{i0} - \langle g, \omega_q^{i0}\rangle_{i0}), \qquad q = 1, \ldots, n_{i0},$$

to the quadratic form for each interface Γ_{ij}, where $-\lambda_q^{ij}$ is a Lagrange multiplier. This converts the problem to that of finding the unconstrained minimum of a function $F(b, \lambda)$, which produces a system of linear equations of form

$$\left(\begin{array}{c|c} \mathbf{C} & \mathbf{M}^T \\ \hline \mathbf{M} & 0 \end{array}\right)\begin{pmatrix} \mathbf{b} \\ -\lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}.$$

If the selfadjoint elliptic equation is of Helmholtz type, with $A_0 \geq c > 0$, matrix $\mathbf{C}$ consists of symmetric positive definite blocks along the diagonal and is zero elsewhere. Each block corresponds to a cell, and the number of basis functions used on the cell is the number of rows of a block.

The vector containing the coefficients to be used with the basis functions to obtain the approximation is $\mathbf{b}$.

Entries corresponding to the right-hand side of the elliptic equation $\mathbf{E}u = f$ are represented by $\mathbf{f}$.

The rectangular matrix $\mathbf{M}$, which we call the matrix of *moment collocation rows*, consists of a band of blocks, with zeros below the band; it is sparse above the band. In [25] it is shown that the rows of $\mathbf{M}$ are independent if the total number of basis functions used in the approximation is sufficiently large.

The Lagrange multipliers λ_q^{ij} used to enforce the linear moment collocation requirements, expressed here as $\mathbf{Mb} = \mathbf{g}$, are represented by λ.

The vector $\mathbf{g}$ consists of zeros where λ^{ij} is the Lagrange multiplier; where λ^{i0} is the Lagrange multiplier we have entries dependent on the boundary value g.

Let $D_{\mathbf{n}_{ij}}u$ represent the "co-normal" derivative of u relative to Γ_{ij}. This is defined for sufficiently smooth u as follows: If $\mathbf{n} = (n_1, n_2, \ldots, n_K)$ is the unit normal to Γ_{ij} (pointing outward relative to the interior of Ω_i), then $D_{\mathbf{n}_{ij}}u \equiv \sum_{p,q}^{K} \gamma_{ij}(A_{pq}D_q u)n_p$. Results in [24] show that $D_{\mathbf{n}_{ij}}u$ is approximated by $\sum_{k=1}^{n_{ij}} \lambda_k^{ij}\omega_k^{ij}$.

The estimates establishing convergence for the inhomogeneous Dirichlet problem are based on the following assumptions:

(1.2a) We assume that $A_{ij}(x) \in H^1(\Omega)$ with $D_k A_{ij}(x) \in L_\infty(\Omega)$ and that the $A_{ij}(x)$ are Lipschitz continuous on $\overline{\Omega}$. We suppose that $A_{ij}(x) = A_{ji}(x)$. We assume that there exists $c > 0$ such that $\sum_{i,j}^{K} A_{ij}(x)z_i z_j \geq c\sum_{i=1}^{K} z_i^2$ for x in Ω and any z_i in $\mathbb{R}$. We assume that $A_0(x) \in L_\infty(\Omega)$ and $A_0(x) \geq c > 0$. (We show that this last assumption is not necessary in Lemma 1.2 below.)

(1.2b) We assume that $f \in L_2(\Omega)$ and that the boundary data g is in $H^{3/2}(\Gamma_{i0})$ for each Γ_{i0}.

Under assumptions (1.2a) and (1.2b), the following convergence result is shown in [24] and [25]:

Theorem 1.1. *Suppose that the solution u to* (1.1a), (1.1b) *is in $H^2(\Omega)$. Let n_f denote the largest number of faces Γ_{ij} of any of the N cells. Let M be a constant such that $a(v, v) \leq M\|v\|_{H^1}^2$ and c be a coercivity constant such that $c\|v\|_{H^1}^2 \leq a(v, v)$. Denote by C_T the maximum of the "trace constants" C_{ij}. Let W be an upper bound for the squares of the $L_2(\Gamma_{ij})$-norms of the*

collocation weight functions ω_p^{ij} *used for collocation on* Γ_{ij} *and let* m_c *represent the largest number of collocations used on all the faces of any cell. Suppose that* $u_{n,m}$ *denotes the approximation obtained by solving the linear system described above. Then*

$$c\|u - u_{n,m}\|_H \le n_f\sqrt{N}C_T \max\{\|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}}u)\|_{ij}\} + M\sqrt{1 + 2(1/\mu)C_T^2 W m_c}\,\|\mathscr{C}_{[m]}(u)\|_H . \tag{2}$$

A slight alteration of the proof also provides the following estimate:

$$c\|u - u_{n,m}\|_H \le \sqrt{2n_f}C_T \left\{\sum_{\Gamma_{ij}} \|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}}u)\|_{ij}^2\right\}^{1/2} + M\sqrt{1 + 2(1/\mu)C_T^2 W m_c}\,\|\mathscr{C}_{[m]}(u)\|_H . \tag{3}$$

Lemma 1.3 below describes μ and a method for obtaining its value and shows that Wm_c can be replaced by n_f.

The dependence of the error on the solution u is expressed in the two terms $\|\mathscr{C}_{[m]}(u)\|_H$ and $\|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}}u)\|_{ij}$. The second term represents the $L_2(\Gamma_{ij})$ norm of the residual of the normal derivative of the solution u that is *not* in the span of the first n_{ij} weight functions used for moment collocation on the interface Γ_{ij}. We present estimates of these two errors for a polynomial implementation in the next section.

Our first new result is that we need not require that the operator is of Helmholtz type, so that there is some $c > 0$ such that $A_0(x) \ge c$; $A_0(x)$ can be zero. We show that under mild restrictions on the weight functions ω_k^{ij}, Poincaré's inequality $\|v\|_0 \le C\|\nabla v\|_0$ for some constant C holds over the space $G_0[n]$; the result then follows from the ellipticity assumption in (1.2a).

Lemma 1.2. *For* $u \in H$, *the distribution* ∇u_i *exists in* $L_2(\Omega_i)$ *for each* i, $i = 1, \ldots, N$. (*We use the symbol* ∇u *to denote the function defined in this manner for each cell.*) *If* $[n]$ *is sufficiently large so that for each* Γ_{ij} *there is some* $k \le n_{ij}$ *such that* $\int_{\Gamma_{ij}} \omega_k^{ij}\, d\Gamma \ne 0$, *then there exists some constant* $c > 0$ *such that for all* $u \in G_0[n]$, $\|\nabla u\|_0^2 \ge c[\|\nabla u\|_0^2 + \|u\|_0^2]$.

Proof. If there is no such $c > 0$, then there exists some sequence u_m in $G_0[n]$ with $\|\nabla u_m\|_0^2 + \|u_m\|_0^2 = 1$ such that $\|\nabla u_m\|_0^2 \to 0$. Since, for LPC^1 domains, the embedding $H \hookrightarrow L_2(\Omega)$ is compact [26], for a bounded sequence u_m there is a subsequence $u_{m(i)}$ such that $u_{m(i)}$ converges strongly to some u in $L_2(\Omega)$. Since $\nabla u_{m(i)}$ converges strongly (to zero) in $L_2(\Omega)$, $u_{m(i)}$ converges strongly to u in H (and $\nabla u = 0$ as a distribution). $G_0[n]$ is closed in H by the continuity of the trace operator. Thus, $u \in G_0[n]$, $\nabla u = 0$, and $\|u\|_0^2 = \|\nabla u\|_0^2 + \|u\|_0^2 = 1$. Since $\|\nabla u\|_0 = 0$, u has higher distribution derivatives (equal to zero) of any order, so, by the Sobolev embedding theorem, u can be taken to be continuous on each Ω_i. The derivatives of u are all zero on each cell; hence u must be some constant K_i on each cell Ω_i.

For any cell Ω_i with an external boundary face Γ_{i0}, for the ω_k^{i0} such that $\int_{\Gamma_{i0}} \omega_k^{i0}\, d\Gamma \ne 0$ we must have

$$0 = \langle \gamma_{i0}(u), \omega_k^{i0}\rangle_{i0} = \langle K_i, \omega_k^{i0}\rangle_{i0} = K_i \int_{\Gamma_{i0}} \omega_k^{i0}\, d\Gamma .$$

Hence, $K_i = 0$. Similarly, for a cell Ω_j adjacent to Ω_i,

$$0 = \langle \gamma_{ij}(u) - \gamma_{ji}(u), \omega_k^{ij} \rangle_{ij} = \langle 0 - K_j, \omega_k^{ij} \rangle_{ij} = -K_j \int_{\Gamma_{ij}} \omega_k^{ij} \, d\Gamma,$$

so $K_j = 0$ also. Extending this argument throughout Ω, we see that all K_j are zero; hence $u \equiv 0$. Yet $\|u\|_0^2 = 1$. This contradiction establishes the result. □

The parameter μ appearing in Theorem 1.1 is the smallest eigenvalue of $\mathbf{M}'\mathbf{M}'^T$, where $\mathbf{M}'$ is the array of collocation rows, *assuming that the basis on each cell* Ω_i *is* $H^1(\Omega_i)$*-orthonormal*. It depends on $[n]$ and $[m]$, the domain decomposition and the choice of bases $\{B_k^r\}$ and $\{\omega_k^{ij}\}$; it is independent of the elliptic problem. From [25], for fixed $[n]$, $1/\mu$ is nonincreasing for $[m'] \geq [m]$. The following lemma describes a way to obtain μ that does not assume that $H^1(\Omega_i)$ is orthonormal and improves the estimate in Theorem 1.1.

Lemma 1.3. *Suppose that* $\mathbf{M}$ *and* $\mathbf{C}$ *are the matrix components of the linear system obtained by approximating the solution to the Helmholtz problem* $-\Delta u + u = f$. *Then* μ *is the smallest eigenvalue of* $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$. *The parameter* μ *is independent of a linear change of bases* $\{B_k^r\}$ *and any change of basis* $\{\omega_k^{ij}\}$ *using an orthogonal matrix. If* μ *is calculated relative to any basis* $\{\omega_k^{ij}\}$ *that is* $L_2(\Gamma_{ij})$*-orthonormal, the product* Wm_c *in Theorem* 1.1 *can be replaced by* n_f, *the maximum number of* C^1 *faces of any cell.*

Proof. For any cell Ω_r, given a basis $\{B_k^r\}$, $k = 1, \ldots, m_r$, the Gram-Schmidt process allows us to construct an $H^1(\Omega_r)$-orthonormal basis $\{E_j^r\}$ expressed in terms of suitable coefficients g_{kj}^r as $E_j^r = \sum_{k=1}^j g_{kj}^r B_k^r$. Let the matrix (g_{kj}^r) be denoted by $\mathbf{G}_r$. If $\mathbf{C}_r$ is the matrix $((B_i^r, B_j^r)_{1,r})$, then $\mathbf{G}_r^T \mathbf{C}_r \mathbf{G}_r = ((E_i^r, E_j^r)_{1,r}) = \mathbf{I}$, the identity matrix.

The collocation rows for each cell Ω_r relative to the basis $\{E_j^r\}$ are of the form

$$(\langle \gamma_{ri}(E_1^r), \omega_q^{ri} \rangle, \langle \gamma_{ri}(E_2^r), \omega_q^{ri} \rangle, \ldots, \langle \gamma_{ri}(E_m^r), \omega_q^{ri} \rangle),$$

which we denote by $\mathbf{M}_r'$. In terms of the original basis $\{B_k^r\}$, this row is

$$\left(\left\langle \gamma_{ri} \left(\sum_{k=1}^{1} g_{k1}^r B_k^r \right), \omega_q^{ri} \right\rangle, \left\langle \gamma_{ri} \left(\sum_{k=1}^{2} g_{k2}^r B_k^r \right) w_q^{ri} \right\rangle, \ldots, \left\langle \gamma_{ri} \left(\sum_{k=1}^{m} g_{km}^r B_k^r \right), \omega_q^{ri} \right\rangle \right)$$

$$= (\langle \gamma_{ri}(B_1^r), \omega_q^{ri} \rangle, \langle \gamma_{ri}(B_2^r), \omega_q^{ri} \rangle, \ldots, \langle \gamma_{ri}(B_m^r), \omega_q^{ri} \rangle) \mathbf{G}_r.$$

If $\mathbf{M}_r$ has rows $(\langle \gamma_{ri}(B_1), \omega_q^{ri} \rangle, \langle \gamma_{ri}(B_2), \omega_q^{ri} \rangle, \ldots, \langle \gamma_{ri}(B_m), \omega_q^{ri} \rangle)$, this is expressed as a matrix equation as $\mathbf{M}_r' = \mathbf{M}_r \mathbf{G}_r$.

The proof generalizes the two-cell case, where $\mathbf{M}'$ has the form

$$\begin{pmatrix} \mathbf{M}_{10}' & 0 \\ \mathbf{M}_{12}' & \mathbf{M}_{21}' \\ 0 & \mathbf{M}_{20}' \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{10}\mathbf{G}_1 & 0 \\ \mathbf{M}_{12}\mathbf{G}_1 & \mathbf{M}_{21}\mathbf{G}_2 \\ 0 & \mathbf{M}_{20}G_2 \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{10} & 0 \\ \mathbf{M}_{12} & \mathbf{M}_{21} \\ 0 & \mathbf{M}_{20} \end{pmatrix} \begin{pmatrix} \mathbf{G}_1 & 0 \\ 0 & \mathbf{G}_2 \end{pmatrix}.$$

Denote the first matrix by $\mathbf{M}$ and the second matrix by $\mathbf{G}$. Thus $\mathbf{M}' = \mathbf{M}\mathbf{G}$.

Now μ is the smallest eigenvalue of $\mathbf{M}'\mathbf{M}'^T = [\mathbf{MG}][\mathbf{G}^T\mathbf{M}^T] = \mathbf{M}[\mathbf{GG}^T]\mathbf{M}^T$. Further, $\mathbf{G}^T\mathbf{CG} = \mathbf{I}$, so $\mathbf{C} = (\mathbf{G}^T)^{-1}\mathbf{G}^{-1}$, and $\mathbf{C}^{-1} = \mathbf{GG}^T$. Thus, $\mathbf{M}'\mathbf{M}'^T = \mathbf{M}[\mathbf{GG}^T]\mathbf{M}^T = \mathbf{MC}^{-1}\mathbf{M}^T$.

A straightforward argument establishes that the matrix $\mathbf{MC}^{-1}\mathbf{M}^T$ is invariant under a linear change of basis $\{B_k^r\}$. Any change of basis $\{\omega_k^{ij}\}$ under an orthogonal matrix $\mathbf{K}$ expresses $\mathbf{MC}^{-1}\mathbf{M}^T$ as $(\mathbf{KM}')\mathbf{C}^{-1}(\mathbf{KM}')^T = \mathbf{KM}'\mathbf{C}^{-1}\mathbf{M}'^T\mathbf{K}^T$, where $\mathbf{M}'$ is the array of collocation rows relative to the altered collocation weight functions replacing $\{\omega_k^{ij}\}$. The matrix $\mathbf{M}'\mathbf{C}^{-1}\mathbf{M}'^T$ has the same eigenvalues as $\mathbf{MC}^{-1}\mathbf{M}^T$ since $\mathbf{K}^{-1} = \mathbf{K}^T$.

To complete the proof, we can extend the argument for one cell Ω with one C^1 boundary Γ, where we approximate a solution to the homogeneous boundary value problem. Let $\mathscr{P}_m^n(v)$ denote the H-orthogonal projection of $v \in G_0[n]$ onto $G_0[n][m] \equiv G_0[n] \cap H[m]$. The argument in [25] requires that we majorize $\|v - \mathscr{P}_m^n(v)\|_1$ by some constant times $\|\mathscr{C}_{[m]}(v)\|_H$. We assume that $\{B_k\}$ is $H^1(\Omega)$-orthonormal. From [25], for the one cell, one boundary case, $\|v - \mathscr{P}_m^n(v)\|_1^2 = \|\mathscr{C}_{[m]}(v)\|_H^2 + \mathbf{aA}^{-1}\mathbf{a}^T$, where $\mathbf{A} = \mathbf{MM}^T$, and

$$a = -(\langle \gamma(\mathscr{C}_{[m]}(v)), \omega_1^{10}\rangle, \ldots, \langle \gamma(\mathscr{C}_{[m]}(v)), \omega_n^{10}\rangle).$$

The parameter μ is the smallest eigenvalue of $\mathbf{A}$, so $\mathbf{aA}^{-1}\mathbf{a}^T \le (1/\mu)\mathbf{aa}^T$. We assume that $\{\omega_k^{10}\}$ is $L_2(\Gamma)$-orthonormal, so $-\mathbf{a}$ is the vector of the first n Fourier coefficients for $\mathscr{C}_{[m]}(v)$ and $\mathbf{aa}^T$ is the square of the magnitude of the projection $\mathscr{R}$ of $\gamma(\mathscr{C}_{[m]}(v))$ onto the span of ω_k^{10}, $k = 1, \ldots, n$.

Thus,

$$\begin{aligned}\|v - \mathscr{P}_m^n(v)\|_1^2 &\le \|\mathscr{C}_{[m]}(v)\|_H^2 + (1/\mu)\|\mathscr{R}\|_{10}^2 \\ &\le \|\mathscr{C}_{[m]}(v)\|_H^2 + (1/\mu)\|\gamma(\mathscr{C}_{[m]}(v))\|_{10}^2 \\ &\le \|\mathscr{C}_{[m]}(v)\|_H^2 + (1/\mu)C_{10}^2\|\mathscr{C}_{[m]}(v)\|_H^2,\end{aligned}$$

where C_{10} is the trace constant.

The argument for a general cell decomposition concludes as in [25]. □

It is proved in [25] that $\mathbf{M}$ is of full rank if the total number of basis functions is sufficiently large. Our polynomial implementation presented in the next section can produce approximations that are continuous. However, in this case $\mathbf{M}$ is often not of full rank, for we use the same number of collocations on each Γ_{ij} and the system of equations may be *overdetermined*; more collocations are used than the minimum necessary to force continuity. Theoretically, we should eliminate just enough rows of $\mathbf{M}$ so that $\mathbf{M}$ is of full rank, yet continuity is still enforced, allowing us to obtain the approximation or compute μ for this case. This is awkward to do during the construction of the linear system; we have set up procedures for eliminating redundant rows of $\mathbf{M}$ during the final solution of the linear equations.

We obtain explicit theoretical bounds for all the terms in (2) and (3) for the polynomial implementation in the next section, except for $1/\mu$. To estimate a bound for $1/\mu$, we combine theoretical results with experimental data. A useful general result about $1/\mu$ is the following:

Lemma 1.4. *Suppose that matrices* $\mathbf{C}$ *and* $\mathbf{M}$ *are constructed to approximate solutions to the Helmholtz equation* $-\Delta u + u = f$. *The matrices* $\mathbf{C}$ *and* $\mathbf{C}^{-1}$ *are positive definite blocks along the diagonal, one for each cell. We confine the*

Helmholtz operator to each cell Ω_i *and, using the same collocations as those in the general problem, obtain* μ_i *for each cell. Then* $\mu \geq \min\{\mu_i\}$.

Proof. The result for $N = 2$ generalizes. When there are two cells, $\mathbf{C}^{-1}$ has the form

$$\left(\begin{array}{c|c} \mathbf{B}_1 & 0 \\ \hline 0 & \mathbf{B}_2 \end{array}\right)$$

and $\mathbf{M}$ the form

$$\left(\begin{array}{c|c} \mathbf{M}_1 & 0 \\ \hline 0 & \mathbf{M}_2 \end{array}\right).$$

The blocks $\mathbf{B}_1$ and $\mathbf{B}_2$ are symmetric and positive definite. Suppose that x is an eigenvector corresponding to the least eigenvalue μ of $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$. Represent x by $(x_1 \vdots x_2 \vdots x_3)^T$, where $(x_1 \vdots x_2)$ has length equal to the number of rows in $\mathbf{M}_1$ and $(x_2 \vdots x_3)$ has length equal to the number of rows in $\mathbf{M}_2$. Owing to the block structure of $\mathbf{C}^{-1}$,

$$\begin{aligned} \mu x^t x &= x^T \mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T x \\ &= (x_1 \vdots x_2)\mathbf{M}_1\mathbf{B}_1\mathbf{M}_1^T(x_1 \vdots x_2)^T + (x_2 \vdots x_3)\mathbf{M}_2\mathbf{B}_2\mathbf{M}_2^T(x_2 \vdots x_3)^T. \end{aligned}$$

Since

$$(x_1 \vdots x_2)\mathbf{M}_1\mathbf{B}_1\mathbf{M}_1^T(x_1 \vdots x_2)^T \geq \mu_1(x_1 \vdots x_2)(x_1 \vdots x_2)^T = \mu_1(x_1x_1^T + x_2x_2^T)$$

and

$$(x_2 \vdots x_3)\mathbf{M}_2\mathbf{B}_2\mathbf{M}_2^T(x_2 \vdots x_3)^T \geq \mu_2(x_2 \vdots x_3)(x_2 \vdots x_3)^T = \mu_2(x_2x_2^T + x_3x_3^T),$$

we have

$$\begin{aligned} \mu(x_1x_1^T + x_2x_2^T + x_3x_3^T) = \mu x^T x &= x^T \mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T x \\ &= (x_1 \vdots x_2)\mathbf{M}_1\mathbf{B}_1\mathbf{M}_1^T(x_1 \vdots x_2)^T + (x_2 \vdots x_3)\mathbf{M}_2\mathbf{B}_2\mathbf{M}_2^T(x_2 \vdots x_3)^T \\ &\geq \mu_1(x_1x_1^T + x_2x_2^T) + \mu_2(x_2x_2^T + x_3x_3^T) \\ &\geq \min\{\mu_1, \mu_2\}(x_1x_1^T + 2x_2x_2^T + x_3x_3^T). \end{aligned}$$

Thus,

$$\mu \geq \frac{(x_1x_1^T + 2x_2x_2^T + x_3x_3^T)}{(x_1x_1^T + x_2x_2^T + x_3x_3^T)} \min\{\mu_1, \mu_2\} \geq \min\{\mu_1, \mu_2\}. \quad \square$$

2. Error estimates for polynomial implementations in $\mathbb{R}^2$ and $\mathbb{R}^3$

We have written programs that produce approximations to solutions of problems with domains in $\mathbb{R}^2$ [25]. We accommodate four types of cells. Cells can be parallelograms or triangles in any orientation. Two kinds of cells with one curved (external) boundary segment are accepted; the first has one straight side and one curved side; the second has two straight sides and one curved side.

Legendre polynomials are used to generate an L_2-orthonormal basis for a square, which provides a basis for any parallelogram by the use of affine transformations. An L_2-orthonormal polynomial basis has been contrived for triangles. Our software currently generates up to 66 basis functions for any cell,

giving a full tenth-degree polynomial basis. These two bases are adapted for use in the two types of cells with a curved boundary segment. We use Legendre polynomials for the weight functions ω_k^{ij}. Our error estimates for domains in $\mathbb{R}^2$ assume that the cells are triangles or parallelograms.

For domains in $\mathbb{R}^3$, we confine our attention to tetrahedral or parallelepiped cells. In [24] we describe a method for using Legendre polynomials to construct an L_2-orthonormal basis for a standard cube and propose a method for constructing an L_2-orthonormal basis for a standard 3-simplex; these basis functions are currently being computed. Affine transformations can carry such bases to any parallelepiped or tetrahedron in $\mathbb{R}^3$. The bases for triangles or parallelograms can be used to provide L_2-orthonormal collocation weight functions for the faces of such cells.

Our polynomial implementation of the algorithm includes a version of the *h-p* finite element method [4, 6] as a special case. For example, suppose our elements are in $\mathbb{R}^2$. If we use polynomials of degree less than or equal to p for the basis in each cell and choose the first $p+1$ Legendre polynomials for collocation weight functions on each interface Γ_{ij}, our approximation is continuous throughout Ω, since the difference of the traces of the approximation on either side of any Γ_{ij}, if nonzero, is a polynomial of degree at most p, yet the difference must be orthogonal to the Legendre polynomial weight functions ω_k^{ij} for $k \leq p+1$. Since our variational principle is the same as that of the finite element method, our approximation is exactly that of the finite element method as described in [6], with boundary data obtained by the $L_2(\Gamma_{i0})$ projection onto the span of the pth-order collocation weight functions. The *h-p* method described in [4] can be implemented with a small modification of our algorithm. (This requires that we replace two moment collocations on each boundary edge with point collocations at the end points of the edge.) For parallelepiped or tetrahedral cells in $\mathbb{R}^3$ we would need $(p+1)(p+2)/2$ collocations on each interface to force continuity of pth-order approximations.

Our method is more general than the usual finite element p-method, for we can choose the number of moment collocations, say, in $\mathbb{R}^2$, to be less than $p+1$. For domains in $\mathbb{R}^2$ we use the same number $q+1$ of moment collocations on each interface, corresponding to moments involving polynomials of degree q or less.

To obtain error estimates in terms of p and q and the cell diameter h, we use the values for the trace constants C_{ij} that are given in [24]. The results are the following:

If the cell is a parallelogram, $C_{ij} \leq \sqrt{1/\sin\theta + 1/l}$, where θ is the angle made by two adjacent sides of the parallelogram and, if the base is Γ_{ij}, l is the height of the parallelogram. If the cell is a triangle, $C_{ij} \leq \sqrt{2/\sin\theta + 4/l}$, where θ is the smallest angle in the triangle and, if the base of the triangle is Γ_{ij}, l is the height of the triangle. Let h represent the diameter of a triangle or parallelogram. For later estimates, for any particular problem, we define K_1 to be $h/\min\{l\}$, where the minimum is taken over all altitudes of triangular cells and all heights of parallelogram cells. Thus $l \geq h/K_1$, so that $1/l \leq K_1/h$.

If the cell is a parallelepiped, $C_{ij} \leq \sqrt{1/\det \mathbf{N} + 1/l}$, $\det \mathbf{N} = |\det(\mathbf{n}_1 \vdots \mathbf{n}_2 \vdots \mathbf{n}_3)|$, where $\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$ are unit normals to the sides of the parallelepiped, and

l is the smallest height of the parallelepiped relative to any base. If the cell is a tetrahedron, $C_{ij} \leq \sqrt{7/\det \mathbf{N} + 14/l}$, $\det \mathbf{N} = \min |\det(\mathbf{n}_1 \vdots \mathbf{n}_2 \vdots \mathbf{n}_3)|$, where $\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$ are any three of the unit normals to the sides of the tetrahedron, and l is the smallest height of the tetrahedron relative to any base.

We can be more specific about the dependence of μ on the size of a cell with the use of the following lemmas:

Lemma 2.1. *For the bases and types of cells described above, the values for μ for a single cell are invariant under translation or rotation of the cell.*

The proof is straightforward and can be found in [24].

If μ is obtained for a cell, and the cell is scaled by a factor h $(h < 1)$, so that a side of length l becomes a side of length hl, we expect that the associated μ_h is generally greater than or equal to $h\mu$. For triangles, parallelograms, tetrahedra, or parallelepipeds, this is the content of Lemma 2.2.

Lemma 2.2. *Let bases for triangles, parallelograms, tetrahedra, or parallelepipeds be obtained from bases for the standard simplex, square, 3-simplex, or cube using an affine transformation of the form $y \mapsto x_0 + \mathbf{T}y$. Assume that a standard basis $\{B_i\}$ is L_2-orthonormal. Let $\mathbf{D}$ denote the matrix with $i-jth$ entry equal to the integral of $(\nabla B_i)^T T^{-1}(T^{-1})^T(\nabla B_j)$ over a standard cell. The (Helmholtz problem) matrix $\mathbf{C}$ for the image of a standard cell is of the form $\mathbf{I}+\mathbf{D}$. Suppose that μ_1 is the smallest eigenvalue for $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$. When all sides of a cell are scaled by a factor h, $h < 1$, the matrix $\mathbf{C}_h$ is $\mathbf{I} + (1/h^2)\mathbf{D}$; the matrix of collocation rows is $\mathbf{M}_h \equiv (1/\sqrt{h})\mathbf{M}$, where $\mathbf{M}$ is the collocation row matrix when $h = 1$. Suppose that μ_h is the smallest eigenvalue of $\mathbf{M}_h\mathbf{C}_h^{-1}\mathbf{M}_h^T$. Then $\mu_h \geq h\mu_1$.*

Proof. Since $\mathbf{C} = (\mathbf{D} + \mathbf{I}) = (\mathbf{D} + h^2\mathbf{I}) + (1 - h^2)\mathbf{I}$, we have

$$(\mathbf{D} + h^2\mathbf{I})^{-1} = \mathbf{C}^{-1} + (1 - h^2)(\mathbf{C}^{-1})(\mathbf{D} + h^2\mathbf{I})^{-1},$$

so

$$\begin{aligned} x^T\mathbf{M}_h(\mathbf{C}_h)^{-1}\mathbf{M}_h^T x &= x^T(1/\sqrt{h})\mathbf{M}((1/h^2)(\mathbf{D} + h^2\mathbf{I}))^{-1}(1/\sqrt{h})\mathbf{M}^T x \\ &= hx^T\mathbf{M}(\mathbf{D} + h^2\mathbf{I})^{-1}\mathbf{M}^T x \\ &= hx^T\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T x + hx^T\mathbf{M}(1 - h^2)(\mathbf{C}^{-1})(\mathbf{D} + h^2\mathbf{I})^{-1}\mathbf{M}^T x, \end{aligned}$$

$$\begin{aligned} (\mathbf{C}^{-1})(\mathbf{D} + h^2\mathbf{I})^{-1} &= (\mathbf{D} + \mathbf{I})^{-1}(\mathbf{D} + h^2\mathbf{I})^{-1} = ((\mathbf{D} + h^2\mathbf{I})(\mathbf{D} + \mathbf{I}))^{-1} \\ &= (h^2\mathbf{I} + (1 + h^2)\mathbf{D} + \mathbf{D}^2)^{-1}. \end{aligned}$$

The matrix $h^2\mathbf{I} + (1 + h^2)\mathbf{D} + \mathbf{D}^2$ is positive definite, so

$$x^T\mathbf{M}(1 - h^2)(\mathbf{C}^{-1})(\mathbf{D} + h^2\mathbf{I})^{-1}\mathbf{M}^T x$$

is a positive expression. Thus,

$$\mu_h = \inf x^T\mathbf{M}_h(\mathbf{C}_h)^{-1}\mathbf{M}_h^T x \geq h \inf x^T\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T x = h\mu_1,$$

where the infimum is taken over all x of norm 1. □

We first apply the methods used by Babuška et al. [3, 4] to provide estimates for cells in $\mathbb{R}^2$ that are triangles or parallelograms. Lemma 2.3 is one of the estimates used in their arguments.

Lemma 2.3. *Let* $I = (-s, s)$, $v \in H^m(I)$, $m \geq 1$. *Then there exists a polynomial* z_q *of degree* q *and a constant* $C(m)$ *independent of* s, q *and* v *such that*

$$\|v - z_q\|_{L_2[I]} \leq C(m) s^{\min(m, q+1)} q^{-m} \|v\|_{H^m[I]}.$$

Lemma 2.4. *In the polynomial implementation of the CDA applied to triangular or parallelogram cells in* $\mathbb{R}^2$ *with diameter* h *there is a constant* $K_{ij}(m)$, *depending on* m, *but not on the solution* u *or* h, *such that, if* $D_{\mathbf{n}_{ij}} u$ *is in* $H^m(\Gamma_{ij})$, *then, with* $n_{ij} \equiv q + 1$ *collocations enforced on* Γ_{ij}, *we have*

$$\|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}} u)\|_{ij} \leq K_{ij}(m)(h/2)^{\min(m, q+1)} q^{-m} \|D_{\mathbf{n}_{ij}} u\|_{H^m(\Gamma_{ij})}.$$

Proof. Legendre polynomials are used to provide collocation weight functions on the interfaces in the implementation described above. Hence $\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}} u)$ is the $L_2(\Gamma_{ij})$-orthogonal complement to the projection $\mathscr{P}$ of $D_{\mathbf{n}_{ij}} u$ onto the span of the weight functions ω_{ij}^k, $k = 1, \ldots, n_{ij}$. Hence $\mathscr{P}$ is the w that minimizes $\|D_{\mathbf{n}_{ij}} u - w\|_{ij}$ over all polynomials w of degree $n_{ij} - 1$ or less. Thus, in this case, $\|D_{\mathbf{n}_{ij}} u - \mathscr{P}\|_{ij} = \|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}} u)\|_{ij}$. The interfaces Γ_{ij} are straight line segments, so, using Lemma 2.3, we have $\|D_{\mathbf{n}_{ij}} u - \mathscr{P}\|_{ij} \leq \|D_{\mathbf{n}_{ij}} u - z_q\|_{ij}$ for the z_q supplied by the lemma (with $q = n_{ij} - 1$), and the result follows. □

We obtain a more global estimate of the first error term in the constant-coefficient case.

Lemma 2.5. *In the polynomial implementation of the CDA applied to triangular or parallelogram cells in* $\mathbb{R}^2$, *if the coefficients* A_{ij} *are constant, there is a constant* K_2 *depending on* A_{ij}, k, *the unit normals to* Γ_{ij} *and the ratio of the sides of any parallelograms, but not on* h *or the solution* u, *such that, if* u *is in* $H^k(\Omega)$, *then, with* $q + 1$ *collocations enforced on each* Γ_{ij}, *we have*

$$\left(\sum_{\Gamma_{ij}} \|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}} u)\|_{ij}^2 \right)^{1/2} \leq K_2 \sqrt{n_f} C_T (h/2)^{\min(k-2, q+1)} q^{-(k-2)} \|u\|_{H^k(\Omega)}.$$

Proof. We let $m = k - 2$. We express $D_{\mathbf{n}_{ij}} u$ in terms of the traces of the first derivatives of u, the constants A_{ij}, and the unit normals to Γ_{ij}. Let c_i denote various constants dependent on A_{ij}, the unit normals to the Γ_{ij} and the multi-index α. Then

$$\begin{aligned}
\|D_{\mathbf{n}_{ij}} u\|_{H^m(\Gamma_{ij})}^2 &= \|c_1 \gamma_{ij}(D_1 u) + c_2 \gamma_{ij}(D_2 u)\|_{H^m(\Gamma_{ij})}^2 \\
&\leq \sum_{|\alpha| \leq m} \|c_3 \gamma_{ij}(D^\alpha D_1 u) + c_4 \gamma_{ij}(D^\alpha D_2 u)\|_{ij}^2 \\
&\leq C_T^2 \sum_{|\alpha| \leq m} \|c_3 D^\alpha D_1 u + c_4 D^\alpha D_2 u\|_{1,i}^2 \leq C_T^2 c_5 \|u\|_{H^{m+2}(\Omega_i)}^2.
\end{aligned}$$

Using Lemma 2.4, we obtain

$$\sum_{\Gamma_{ij}} \|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}}u)\|_{ij}^2 \le (h/2)^{2\min(m,q+1)} q^{-2m} \sum_{\Gamma_{ij}} K_{ij}^2 \|D_{\mathbf{n}_{ij}}u\|_{H^m(\Gamma_{ij})}^2$$

$$\le (h/2)^{2\min(m,q+1)} q^{-2m} C_T^2 c_6 \sum_{\Gamma_{ij}} \|u\|_{H^{m+2}(\Omega_i)}^2$$

$$\le (h/2)^{2\min(m,q+1)} q^{-2m} C_T^2 c_6 n_f \sum_{i=1}^{N} \|u\|_{H^{m+2}(\Omega_i)}^2$$

$$= (h/2)^{2\min(m,q+1)} q^{-2m} C_T^2 c_6 n_f \|u\|_{H^{m+2}(\Omega)}^2 . \quad \square$$

If we apply results of Babuška et al. [3, 4] in the style of Lemma 2.4, we obtain an estimate of the second error term $\|\mathscr{E}_{[m]}(u)\|_H$ that holds for the domains in $\mathbb{R}^2$ described above and similar domains in $\mathbb{R}^3$:

Lemma 2.6. *Suppose that the solution u to the Dirichlet problem is in $H^k(\Omega)$. Suppose that the cells partitioning Ω are affine maps of the unit square, standard triangle, unit cube or unit simplex, with h representing the maximum diameter. Suppose that $[m]$ is large enough so that the basis functions used on each cell can generate any pth-degree polynomial. Then there exists a constant K_3 depending on k and the unit normals to the sides of the cells, but independent of p, h, and u, such that*

$$\|\mathscr{E}_{[m]}(u)\|_H \le K_3 h^{\min(k-1,p)} p^{-(k-1)} \|u\|_{H^k(\Omega)} .$$

The estimates given in the previous lemma combine with the second estimate of Theorem 1.1 to yield the following result:

Theorem 2.7. *Suppose $\Omega \subset \mathbb{R}^2$ is partitioned into N triangles or parallelograms (or both) of diameter h or less with smallest angle between the sides denoted by θ. Let $K_1 = h/\min\{l\}$, where l is any altitude of a triangle or any height of a parallelogram (relative to any side). Assume that the A_{ij} and A_0 are constant. Suppose that $q+1$ collocations are used on each interface (corresponding to collocations with polynomials with degree $\le q$) and the number of basis functions m used on any cell is $(p+1)(p+2)/2$, corresponding to a full pth-order basis; $p \ge q$. Suppose that the solution $u \in H^k(\Omega)$, $k > 2$. Then there are constants C_i, $i = 1, \ldots, 4$, depending on the A_{ij}, A_0 and the angles of the cells and k, and there is a parameter μ_1 depending on p, q and the angles of the cells (and the ratios of the sides of any parallelograms) such that the approximation $u_{q,p} \equiv u_{n,m}$ satisfies the following estimate:*

$$\|u - u_{q,p}\|_H$$

$$\le [C_1(1/\sin\theta + 2K_1/h)(h/2)^{\min(k-2,q+1)} q^{-(k-2)}$$

$$+ C_2\sqrt{1 + 16(1/(h\mu_1))(1/\sin\theta + 2K_1/h)}\, h^{\min(k-1,p)} p^{-(k-1)}]\|u\|_{H^k(\Omega)} .$$

A more succinct estimate is

$$\|u - u_{q,p}\|_H \le [C_3(h/2)^{\min(k-3,q)} q^{-(k-2)} + C_4\sqrt{1/\mu_1}\, h^{\min(k-2,p-1)} p^{-(k-1)}]\|u\|_{H^k} .$$

A similar estimate can be obtained for domains in $\mathbb{R}^3$:

Theorem 2.8. *Suppose* $\Omega \subset \mathbb{R}^3$ *is partitioned into* N *tetrahedra or parallelepipeds (or both) of diameter* h *or less with* $\det N = \min_k |\det(\mathbf{n}_1^k : \mathbf{n}_2^k : \mathbf{n}_3^k)|$, *where* $\mathbf{n}_1^k$, $\mathbf{n}_2^k$, *and* $\mathbf{n}_3^k$ *are unit normals to the sides of any tetrahedron or parallelepiped cell* Ω_k. *Let* $K_1 = h/\min\{l\}$, *where* l *is any altitude of a tetrahedron or parallelepiped (relative to any face). Assume that the* A_{ij} *are constant. Suppose that* $n_{ij} \equiv (q+1)(q+2)/2$ *collocations are used on each interface (corresponding to collocations with polynomials with degree* $\leq q$*) and the number of basis functions* m *used on any cell is* $(p+1)(p+2)(p+3)/6$, *corresponding to a full* p*th-order basis* $(p \geq q)$. *Suppose that the solution* $u \in H^k(\Omega)$, $k > 2$. *Then there are constants* C_i *depending on the* A_{ij}, A_0 *and the normals to the sides of the cells and* k, *and there is a parameter* μ_1 *depending on* p, q, *the normals to sides of the cells and the ratios of the areas of the sides of any parallelepiped cells, but independent of* h *and* u, *such that the approximation* $u_{q,p} \equiv u_{n,m}$ *satisfies the following estimate:*

$$\|u - u_{q,p}\|_H \leq [C_1(7/\det \mathbf{N} + 14K_1/h)h^{\min(k-2,q+1)}q^{-(k-2)} + C_2\sqrt{1 + 12(1/h\mu_1)}(7/\det \mathbf{N} + 14K_1/h)h^{\min(k-1,p)}p^{-(k-1)}]\|u\|_{H^k(\Omega)}.$$

A shorter estimate is

$$\|u - u_{q,p}\|_H \leq [C_3 h^{\min(k-3,q)}q^{-(k-2)} + C_4\sqrt{1/\mu_1}h^{\min(k-2,p-1)}p^{-(k-1)}]\|u\|_{H^k}.$$

As mentioned earlier, sufficient collocations can be used to force an approximation to be continuous with $\mathbf{M}$ of full rank. It follows from the proof of Theorem 1.1 [25] that the expression $\|\mathcal{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}}u)\|_{ij}$ is eliminated from the error estimate for a homogeneous Dirichlet problem with a polygonal domain. The estimates of Theorems 2.7 and 2.8 then give

$$\|u - u_{p,p}\|_H \leq C_4\sqrt{1/\mu_1}h^{\min(k-2,p-1)}p^{-(k-1)}\|u\|_{H^k}.$$

For the h-p method, the estimates of [5] give

$$\|u - u_p\|_H \leq Ch^{\min(k-1,p)}p^{-(k-1)}\|u\|_{H^k}$$

for the two-dimensional case. Our estimate contains $1/\mu_1$, which depends on p and can be large, so our current error estimate does not contain the h-p error estimate as a limiting case, and it may be possible to improve our results for general p and q when a basis consists of polynomials.

In the experimental results of §3 below, the known solution is analytic, and we get approximately the same error using our implementation of the h-p method and the cell method when continuity of an approximation is not enforced and q is about $p-2$ or $p-3$. We give additional polynomial approximation error results for such smooth solutions based on Taylor's series [24] that, except for μ_1, are quite specific.

Lemma 2.9. *If* $v \in C^{q+1}([a,b])$, *let* v^{q+1} *denote the* $(q+1)$*st-order derivative of* v *and* v_q *denote the* q*th-order Taylor series approximation to* v *around the point* $d \equiv (a+b)/2$; *then*

$$\|v - v_q\|_0 \leq \frac{(b-a)^{q+1}}{2^{q+3/2}(q+1)!}\|v^{q+1}\|_0,$$

where the subscript 0 *of* $\|\cdot\|_0$ *denotes the* $L_2[a,b]$ *norm.*

If we apply this result in the manner of Lemma 2.4 and use the density results of smooth functions in $H^k(\Omega)$ [26], we get

Corollary 2.10. *In the polynomial implementation of the CDA applied to triangular or parallelogram cells in* $\mathbb{R}^2$ *in* §2, *if the functions* A_{ij} *in the elliptic problem are constant and* h_{ij} *is the length of* Γ_{ij}, *then, if* u *is in* $H^{q+5/2}(\Omega)$, *with* $n_{ij} = q+1$ *for each* Γ_{ij}, *we have*

$$\|\mathscr{T}_{n_{ij}}^{ij}(D_{\mathbf{n}_{ij}}u)\|_{ij} \leq [h_{ij}^{q+1}/[2^{q+3/2}(q+1)!\,]]\|(D_{\mathbf{n}_{ij}}u)^{q+1}\|_{ij},$$

where $(D_{\mathbf{n}_{ij}}u)^{q+1}$ *represents the* $(q+1)$*st tangential derivative of* $(D_{\mathbf{n}_{ij}}u)$ *on* Γ_{ij}.

From [24], $H^1(\Omega)$ error estimates for the Taylor series approximation on any convex or star-shaped domain in $\mathbb{R}^2$ or $\mathbb{R}^3$ are the following:

Lemma 2.11. *Suppose* Ω *is any convex domain (with* x_0 *at the center of a largest diameter) or* Ω *is a star-shaped domain (with respect to* x_0*) in* $\mathbb{R}^2$ *or* $\mathbb{R}^3$. *Let* h *represent the diameter of* Ω; *suppose* $R = h/2$ *if* Ω *is convex or* $R = h$ *if* Ω *is star-shaped. Suppose that* $v \in C^{p+2}(\Omega)$ *and* v_p *is the Taylor series expansion of* v *of degree* p *around* x_0. *Let* $|v|_{H^k}$ *denote the seminorm* $[\sum_{|\alpha|=k} \|D^\alpha v\|_0^2]^{1/2}$. *Then*

(i) *If* $\Omega \subset \mathbb{R}^2$,

$$\|v - v_p\|_{H^1(\Omega)}^2 \leq \frac{R^{2p}2^{p-1}}{p((p-1)!\,)^2}\left(\frac{R^2}{(p+1)(2p-1)}+1\right) \cdot \left[[|v|_{H^{p+1}}]^2 + \left(\frac{R^2}{(p-1)(p+1)}\right)[|v|_{H^{p+2}}]^2\right].$$

(ii) *If* $\Omega \subset \mathbb{R}^3$,

$$\|v - v_p\|_{H^1(\Omega)}^2 \leq \frac{R^{2p}3^{p-1}}{2((p-1)!\,)^2}\left(\frac{3R^2}{p^2}+1\right)\left[[|v|_{H^{p+1}}]^2 + \frac{3R^2}{p^2}[|v|_{H^{p+2}}]^2\right].$$

Since the space of polynomials of degree p is finite-dimensional, we can use the density of smooth functions in $H^{p+2}(\Omega)$ and a compactness argument to establish the existence of a polynomial v_p of degree p for any $v \in H^{p+2}(\Omega)$ that satisfies the estimates of Lemma 2.11.

We assemble these results to obtain an error estimate for smooth solutions for the case where a doman in $\mathbb{R}^2$ is partitioned into N triangles or parallelograms (or both). Using Stirling's formula and the method of Lemma 2.4, since we can take $R = h/2$, we obtain the following:

Theorem 2.12. *Suppose* $\Omega \subset \mathbb{R}^2$ *is partitioned into* N *triangles or parallelograms (or both) of diameter* h *or less with smallest angle between the sides denoted by* θ. *Let* $K_1 = h/\min\{l\}$, *where* l *is any altitude of a triangle or any height of a parallelogram (relative to any side). Assume that the* A_{ij} *are constant. Suppose that* $q+1$ *collocations are used on each interface (corresponding to collocations with polynomials up to* q*th degree) and the number of basis functions* m *used on any cell is* $(p+1)(p+2)/2$, *corresponding to a full* p*th-order*

basis. Suppose that the solution $u \in H^{p+2}(\Omega)$, $p \geq 1$. *Then*

$$c\|u - u_{n,m}\|_H \leq 2.2\sqrt{N(1/\sin\theta + 2K_1/h)}\mathscr{E}_1(u, h, q) + .29M\sqrt{1 + 16(1/(h\mu_1))(1/\sin\theta + 2K_1/h)}\mathscr{E}_2(u, h, p),$$

where $\mathscr{E}_1(u, h, q) = h^{q+1}(.73(q+2))^{-(q+3/2)} \max \|(D_{\mathbf{n}_{ij}} u)^{q+1}\|_{ij}$ *and the maximum is taken over all* Γ_{ij}, *and*

$$\mathscr{E}_2(u, h, p) = h^p(.52p)^{-p}\left((h^2/(8p^2)+1)\left[[|u|_{H^{p+1}}]^2 + \left(\frac{h^2}{4(p-1)(p+1)}\right)[|u|_{H^{p+2}}]^2\right]\right)^{1/2},$$

where $|u|_{H^k}$ *represents the seminorm taken over the entire domain* Ω.

The parameter μ_1 has the properties described in Theorem 2.7.

Note that if we are subdividing the unit square into cells of side h, the number of cells $N \cong 1/h^2$. Owing to the decrease in the size of Γ_{ij}, we might expect $\|(D_{\mathbf{n}_{ij}} u)^{q+1}\|_{ij}$ to decrease by a factor $h^{1/2}$. Then the h dependency of the first error term is $Ch^{-3/2}h^{1/2}h^{q+1} = Ch^q$. This can be made rigorous for any polygonal domain by the methods of Lemma 2.5. The h dependency of the second error term is Ch^{p-1}.

For parallelepiped or tetrahedral cells in $\mathbb{R}^3$, results from [24] concerning L_2 estimates derived from Taylor's series applied to the interfaces of such cells establish the following lemma:

Lemma 2.13. *Suppose that a domain in* $\mathbb{R}^3$ *is partitioned into parallelepipeds or tetrahedrons and the collocation weight functions on the interfaces are polynomials of degree* q *or less, and the diameter of* Γ_{ij} *is* h. *Let* $n_{ij} = (q+1)(q+2)/2$. *If* $u \in H^{q+3+1/2}(\Omega)$, *then*

$$\|\mathscr{T}^{ij}_{n_{ij}}(D_{\mathbf{n}_{ij}} u)\|^2_{ij} \leq \frac{h^{2q+2}}{2^{q+3}(q+1)(q!)^2}[[\|(D_{\mathbf{n}_{ij}} u)^{q+1}\|_{ij}]^2 + h^2/(4q^2)[\|(D_{\mathbf{n}_{ij}} u)^{q+2}\|_{ij}]^2],$$

where $[\|v^k\|_{ij}]^2$ *denotes the sum of the* L_2 *norm (squared) of the* k*th-order tangential derivatives of* v *on the 2-dimensional interface* Γ_{ij}.

We assemble the previous estimates and the estimates of the trace constants to obtain the following result for domains in $\mathbb{R}^3$ with the help of Stirling's formula:

Theorem 2.14. *Suppose* $\Omega \subset \mathbb{R}^3$ *is partitioned into* N *tetrahedra or parallelepipeds (or both) of diameter* h *or less with* $\det N = \min_k |\det(\mathbf{n}_1^k : \mathbf{n}_2^k : \mathbf{n}_3^k)|$, *where* $\mathbf{n}_1^k$, $\mathbf{n}_2^k$, *and* $\mathbf{n}_3^k$ *are unit normals to the sides of any tetrahedron or parallelepiped cell* Ω_k. *Let* $K_1 = h/\min\{l\}$, *where* l *is any altitude of a tetrahedron or parallelepiped (relative to any face). Assume that the* A_{ij} *are constant. Suppose that* $n_{ij} \equiv (q+1)(q+2)/2$ *collocations are used on each interface (corresponding to collocations with polynomials up to* q*th degree) and the number of basis functions* m *used on any cell is* $(p+1)(p+2)(p+3)/6$, *corresponding to*

a full pth-order basis $(p \geq q)$. *Suppose that the solution* $u \in H^{p+2}(\Omega)$. *Then*

$$c\|u - u_{n,m}\|_H \leq .63\sqrt{N(7/\det \mathbf{N} + 14K_1/h)}\mathscr{E}_1(u, h, q) + .29M\sqrt{1 + 12(1/(h\mu_1))(7/\det \mathbf{N} + 14K_1/h)}\mathscr{E}_2(u, h, p),$$

where

$$\mathscr{E}_1(u, h, q) = h^{(q+1)}(.52(q+1))^{-(q+1)} \max[[\|(D_{\mathbf{n}_{ij}}u)^{q+1}\|_{ij}]^2 + h^2/(4q^2)[\|(D_{\mathbf{n}_{ij}}u)^{q+2}\|_{ij}]^2]^{1/2},$$

the maximum being taken over all Γ_{ij} *and*

$$\mathscr{E}_2(u, h, p) = h^p(.42p)^{-(p-1/2)}\left\{\left(\frac{h^2}{p^2} + 1\right)\left[[|u|_{H^{p+1}}]^2 + \frac{h^2}{p^2}[|u|_{H^{p+2}}]^2\right]\right\}^{1/2},$$

where $|u|_{H^k}$ *represents the seminorm taken over the entire domain* Ω.

The parameter μ_1 has the properties described in Theorem 2.8.

Note that if we are subdividing the unit cube into cells of side h, the number of cells $N \cong 1/h^3$. Owing to the decrease in the size of Γ_{ij}, we expect $\|(D_{\mathbf{n}_{ij}}u)^{q+1}\|_{ij}$ to decrease by a factor h. Then the h dependency of the first error term is $Ch^{-2}h^1h^{q+1} = Ch^q$, and this can again be made rigorous for any polyhedral domain. The h dependency of the second error term is Ch^{p-1}.

We return to a consideration of the parameter $1/\mu$. Recall that when sufficient collocations are used so that an approximation is continuous (and $\mathbf{M}$ is still of full rank), the first error term can be deleted, and the estimate of Theorem 1.1 for a homogeneous problem with polyhedral domains is

$$c\|u - u_{n,m}\|_H \leq \mathbf{M}\sqrt{1 + 2(1/\mu)C_T^2 n_f}\|\mathscr{C}_{[m]}(u)\|_H.$$

Since $\|\mathscr{C}_{[m]}(u)\|_H$ makes no reference to the continuity of the solution across interfaces, estimates of the parameter $1/\mu$ provide an indication of the effect of enforcing continuity in this case.

We confine our discussion to triangular or parallelogram cells. For a fixed number $q+1$ of collocations, it is shown in [25] that $1/\mu$ is nonincreasing as the number m of basis functions used on each cell is increased. Lemma 1.4 shows that knowledge of a value for $1/\mu$ for low values of m corresponding to approximations of order close to q for *one* cell of various types would give us effective upper bounds for $1/\mu$. Lemmas 2.1 and 2.2 show that, for triangles or parallelograms, it suffices to estimate $1/\mu$ if a cell is rotated and scaled so that a largest side is on the unit interval on the x-axis. The μ obtained when a largest side is of length 1 is denoted by μ_1 and we present some sample empirical estimates for $1/\mu_1$ below.

Our software has the option of computing $\|\gamma_{ij}(u_{n,m}) - \gamma_{ji}(u_{n,m})\|_{ij}$ and $1/\mu$ in any test, yielding the following results. A domain decomposed into a triangle on top of a square requires that $q = p$ to force internal interface continuity. However, computations for just one square or triangle show that continuity of an approximation with zero boundary data is enforced when $q < p$. For example, in all tests with $p \leq 20$, for squares and even p, collocations of degree $q = p - 2$ on three sides and $q = p - 1$ on the fourth suffice to force continuity. For odd p, $q = p - 2$ on three sides and $q = p - 3$ on a fourth

q	values for p						
	4	5	6	7	8	9	10
1	12.93	12.93	12.88	12.88	12.85	12.85	12.84
2	99.19	85.80	22.35	22.35	22.07	22.07	21.91
3			67.43	67.43	34.65	34.65	33.84
4			500.3	340.4	71.25	71.25	49.78
q	values for p						
	8	9	10	11	12	13	14
5	232.2	202.8	83.42	83.42	67.57	67.57	65.20
6	1438.	880.7	179.9	174.4	101.0	101.0	87.92
7			566.7	462.4	174.4	174.4	122.7
8			3154.	1823.	389.0	356.9	187.7
q	values for p						
	12	13	14	15	16	17	18
9	1137.	892.6	333.4	324.8	209.0	209.0	175.9
10	5895.	3287.	727.9	645.6	322.7	322.7	235.9
11			2013.	1542.	585.2	553.0	336.7
12			9913.	5393.	1232.	1069.	530.0

FIGURE 2.1. Values for $1/\mu_1$ with $q+1$ collocations and basis order p

force continuity. The largest value for $1/\mu_1$ for squares and $p \leq 10$ is $6,583$; $p \leq 20$ has maximum $1/\mu_1 > 65,000$. For triangles and odd $p \leq 9$, $q = p - 1$ forces continuity; p even and less than or equal to 10 requires $q = p - 1$ on two sides and $q = p - 2$ on a third. Some experimental results for squares are shown in Figure 2.1.

For any choice of q, values of $1/\mu_1$ are initially high and then drop fairly rapidly as p increases. Likewise, a glance at the columns of Figure 2.1 shows that, for any choice of p, values of $1/\mu_1$ drop rapidly as q decreases. Based on a data set that extends the results in Figure 2.1 and various simple power and exponential multiple regression models to approximate the dependence of $1/\mu_1$ on p and q, the best suggests that

$$1/\mu_1 = 38.5q^{2.14}(p - q - 1)^{-1.58}$$

for even q from 2 to 12 and $p = q + 2, \ldots, q + 7$. (The coefficient of determination R^2 for the logarithmic relationship is $R^2 > .953$.)

When we compare the values of $1/\mu_1$ for q with $p = q + 2$, the logarithms of q and $1/\mu_1$ are highly correlated in the computations for even q from 2 to 20 ($r > .999$). The estimate is $1/\mu_1 = 13.3(p - 2)^{2.67}$. A similar result is obtained for odd p, with q set to $p - 3$.

Computations for a general parallelogram with base 1 give the same sort of empirical estimates; $1/\mu_1$ is proportional to $1/\sin\theta$, where θ is the acute angle of the parallelogram. For the standard simplex, empirical estimates of $1/\mu_1$ are dominated by $21(p - 1)^2$. Although the first error term is eliminated when q is sufficiently close to p so that the approximation is continuous, the results in the next section suggest that an approximation of similar accuracy can be achieved if we relax the requirement of continuity to some extent by decreasing q (which decreases the size of the system of linear equations needed to obtain the approximation). This results in a strong decrease in $1/\mu$, which is balanced

by the growth of the first error term; the best empirical results occur when q is $p-1$ or $p-2$ (for triangles) and $p-2$ or $p-3$ (when cells are rectangles). Note that in Theorem 2.12 the q error dependency is $C(.73(q+2))^{-(q+3/2)}$ and the p error dependency is $C(.52p)^{-p}$ (disregarding $1/\mu$); these are about the same if $q = p-2$.

3. Methods for solving the linear system and experimental results

We first briefly describe the linear algebra used to solve

$$\left(\begin{array}{c|c} \mathbf{C} & \mathbf{M}^T \\ \hline \mathbf{M} & 0 \end{array}\right)\begin{pmatrix} \mathbf{b} \\ -\lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}.$$

The basic algorithm is to use the Schur complement of $\mathbf{C}$; first solve

$$\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T\lambda = -\mathbf{M}\mathbf{C}^{-1}\mathbf{f} + \mathbf{g}$$

for λ, and then solve $\mathbf{C}\mathbf{b} = \mathbf{f} + \mathbf{M}^T\lambda$ for $\mathbf{b}$. (See also [25].) Each of these systems is symmetric semidefinite. Furthermore, $\mathbf{C}$ is block diagonal with block sizes at most 66×66 for $p \le 10$, so it is easy to form $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$, using Linpack [10] to calculate the Cholesky factors of each block of $\mathbf{C}$.

There are several potential difficulties with this approach. First, if there are many cells, then $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$ can be large. However, $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$ is sparse since the only nonzeros in $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$ correspond to Lagrange multipliers associated with interfaces of adjacent cells. For example, for 256 rectangles with $p = 10$ and $q = 10$, $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$ is a 5984×5984 matrix with 1.2% of its entries nonzero. In our implementation we used Sparsepak [13] to solve $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T\lambda = -\mathbf{M}\mathbf{C}^{-1}\mathbf{f} + \mathbf{g}$. In the p version of the finite element method, the matrix that results after static condensation [23] has a structure similar to $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$. If $q < p-1$, then our system will be smaller than the corresponding finite element system.

Following Lemma 1.3, we discussed a second potential difficulty. If q is almost p, then some of the moment constraints may be redundant and $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$ is singular, so we must delete some rows of $\mathbf{M}$. When we detect that a row of $\mathbf{M}$ is numerically dependent on other rows, it is easy to show that an equivalent procedure is to set the corresponding component of λ to zero. This is done with a minor modification of Sparsepak: if a diagonal entry in the LDL^T factorization [13] of $\mathbf{M}\mathbf{C}^{-1}\mathbf{M}^T$ is sufficiently close to zero, we set the relevant component of λ to zero.

A third potential difficulty occurs when we treat Poisson's equation. Lemma 1.2 shows that our methods can be used in this case, but the diagonal blocks comprising $\mathbf{C}$ are singular. As discussed in [25], iterative refinement [14] can be used to overcome this difficulty. In all the cases that we tried, four steps of iterative refinement sufficed to provide solutions to Poisson's equation that are as accurate as the solution to the Helmholtz equation. In most cases one or two steps were sufficient. Iterative refinement is also useful if $\mathbf{C}$ is poorly conditioned but not exactly singular; for example, when $A_0(x)$ in (1.1) is small but not identically zero.

Our test problem is adapted from sample problem 53 of ELLPACK [22]. We seek an approximate solution to the Dirichlet problem

$$-\Delta u + u = f$$

q	n = number of cells: values of h. n=1:h=1.	n=2:h=1.	n=8:h=1.	n=8:h=.5
1	16.61	13.27	10.76	21.42
2	26.28	19.89	17.03	33.97
3	43.18	33.99	28.92	57.77
4	59.67	45.26	40.54	81.03
5	83.37	65.42	57.52	115.0
6	106.7	81.06	76.99	153.9
7	143.4	119.6	111.8	223.5
8	197.7	193.2	188.6	377.1

FIGURE 3.1. Values for $1/\mu_h$ when $p = 10$

on the unit square, where the boundary conditions and f are obtained from the intended true solution

$$u(x, y) = \exp(xy)\cos(\pi y)\sin(\pi(x - y)).$$

We use uniform meshes, with as many as 128 triangles similar to the standard simplex and 256 squares.

Our first tests concern values for $1/\mu$, which only depends on the decomposition of the domain and is independent of the problem. Lemma 1.4 shows that the largest value for $1/\mu$ for a single cell gives an upper bound for $1/\mu$ for multi-cell meshes. This result is demonstrated in Figure 3.1, where the first three columns give decreasing values for $1/\mu$ for various values of q using one triangle, then two and eight triangles partitioning a square. The order of the basis is $p = 10$, and side $h = 1$ in these tests. Lemma 2.2 proves that if μ_h is the value for a cell with sides scaled by a factor of h, then $\mu_h \geq h\mu_1$, or $h(1/\mu_h) \leq (1/\mu_1)$. A test of this result is shown in the last two columns of Figure 3.1, where we show results for eight similar triangular cells when h is 1 and then .5. The entries in the fourth column are almost exactly twice the entries in the third, suggesting that the estimate of Lemma 2.2 is quite tight.

Tests were made to obtain approximation errors for various values of p, q and h. The difference between the true solution and the approximation was calculated on a uniform 41×41 grid; the squares of the "L_2" ("H^1") errors are evaluated using ELLPACK's technique of using the average of the squares of the differences (and the squares of the differences of the derivatives).

We show three sample error computations. The first, in Figure 3.2, relates the logarithm of various errors and p for various values of q. We use two domain decompositions: 32 congruent triangles and 16 congruent squares. We note two results: first, for any fixed value of q, accuracy is not improved by increasing p beyond a certain point, and second, for fixed q, optimal accuracy appears to occur when $p = q + 2$ for regular triangular cells and $p = q + 3$ for regular square cells.

The true solution is analytic; we test the error estimates for such solutions given in Theorem 2.12. Disregarding $1/\mu_1$, the theoretical p dependency of the "H^1" error is of the form $C(.52p)^{-p}$, and the q dependency is $C(.73(q + 2))^{-(q+3/2)}$, with the C's depending on various seminorms of the solution. Thus, we plot the logarithm of various errors against $p \log p$, giving the results shown in Figure 3.3 for various values of q. We again use 32

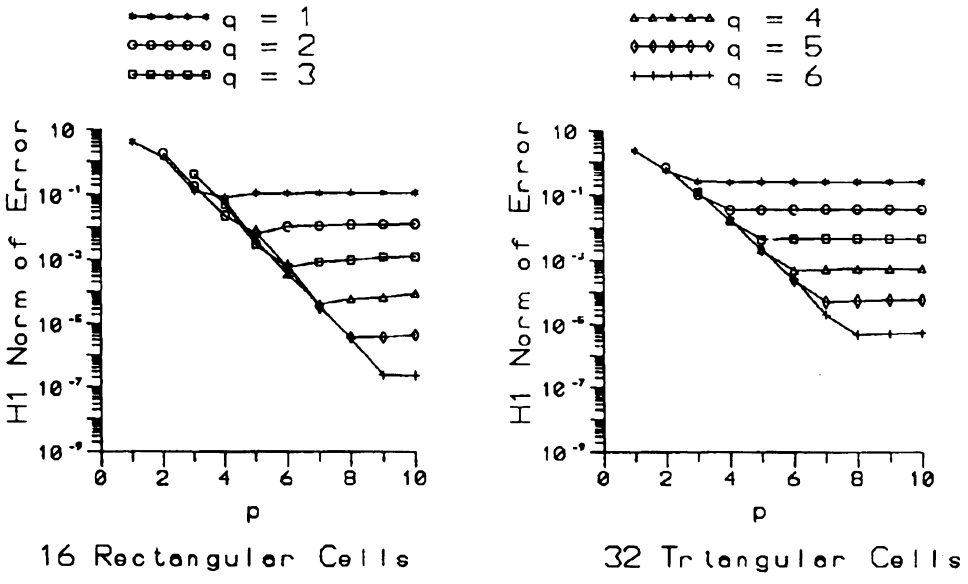

FIGURE 3.2. A comparison of errors and p for various values of q

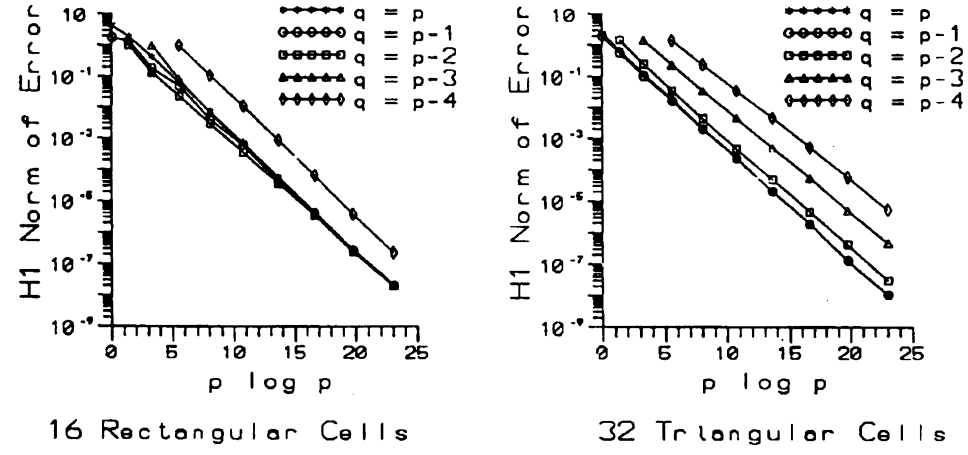

FIGURE 3.3. Log of errors vs. $p \ln p$ for various values of q

triangular cells and 16 square cells, so h is the same for both examples. For triangular cells, when $q = p - 1$, the empirical relationship is

$$(H^1\text{-error}) = .81(.63p)^{-p},$$

with correlation $r > .99$ between values of the log of the error and the log of the approximation. For square cells, when $q = p - 2$, the relationship is

$$(H^1\text{-error}) = .84(.59p)^{-p},$$

with $r > .99$.

A third test is concerned with the h dependency of the approximation. We collect error evaluations for various decompositions into square cells, ranging from one cell to 256 cells. Theorem 2.12 suggests that the h dependency of the

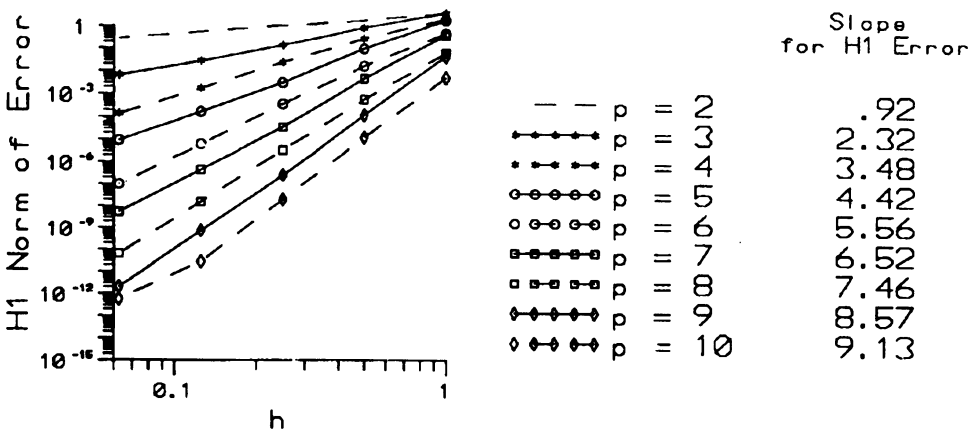

FIGURE 3.4. Log of errors vs. log of h for various values of p

H^1 errors is of the form $C_1h^q + C_2h^{p-1}$. We compare the log of the errors and log of h for various values of p, with $q = p - 2$, in Figure 3.4.

The (approximate) slope of each line gives the exponent for h; this is given in Figure 3.4 beside each value of p. When $q = p - 2$, the theoretically dominant term should be $C_1h^q = C_1h^{p-2}$, yet the empirical results give slopes close to $p - .5$, suggesting that, for parallelogram cells, and $q = p - 2$, we may be able to improve the error bound. For $p = 10$, the graphs are not straight for small h, owing to the effect of computer arithmetic, and the slope listed omits the smallest h. However, we do calculate approximations with maximum errors as small as 10^{-14} with relative machine precision $= 2 \times 10^{-16}$, which suggests that our algorithms are quite robust.

In summary, our polynomial implementation of the cell discretization algorithm has resulted in an alternative method for implementing the p or h-p finite element method, with the option of relaxing the requirement that approximations be continuous across cell interfaces. In our experiments, some discontinuous approximations to smooth solutions have errors similar to continuous approximations.

BIBLIOGRAPHY

1. E. Anderson et al., *LAPACK users' guide*, SIAM, Philadelphia, PA, 1991.
2. I. Babuška, *The finite element method with Lagrangian multipliers*, Numer. Math. **20** (1973), 179–192.
3. I. Babuška and M. R. Dorr, *Error estimates for the combined h and p versions of the finite element method*, Numer. Math. **37** (1981), 257–277.
4. I. Babuška and M. Suri, *The optimal convergence rate of the p-version of the finite element method*, SIAM J. Numer. Anal. **24** (1987), 750–776.
5. ———, *The h-p version of the finite element method with quasiuniform meshes*, RAIRO Modél. Math. Anal. Numér. **21** (1987), 198–238.
6. ———, *The treatment of nonhomogeneous Dirichlet boundary conditions by the p-version of the finite element method*, Numer. Math. **49** (1989), 97–121.
7. ———, *The p and h-p versions of the finite element method, an overview*, Comput. Methods Appl. Mech. Engrg. **80** (1990), 5–26.

8. J. H. Bramble, *The Lagrange multiplier method for Dirichlet's problem*, Math. Comp. **37** (1981), 1–11.
9. M. W. Coffey, J. Greenstadt, and A. Karp, *The application of cell discretization to a "circle in the square" model problem*, SIAM J. Sci. Statist. Comput. **7** (1986), 917–939.
10. J. Dongarra et al., *LINPACK users' guide*, SIAM, Philadelphia, PA, 1979.
11. M. R. Dorr, *The approximation of solutions of elliptic boundary-value problems via the p-version of the finite element method*, SIAM J. Numer. Anal. **23** (1986), 58–76.
12. ———, *On the discretization of interdomain coupling in elliptic boundary-value problems*, Domain Decomposition Methods (T. F. Chan, R. Glowinski, J. Periaux, and O. B. Widlund. eds.), SIAM, Philadelphia, PA, 1989, pp. 17–37.
13. A. George and J. Liu, *Computer solution of large sparse positive definite systems*, Prentice-Hall, Englewood Cliffs, NJ, 1981.
14. W. Govaerts and J. D. Pryce, *Mixed block elimination for linear systems with wider borders*, IMA J. Numer. Anal. **13** (1993), 161–180.
15. J. Greenstadt, *Cell discretization*, Conference on Applications of Numerical Analysis (J. H. Morris, ed.), Lecture Notes in Math., vol. 228, Springer-Verlag, New York, 1971, pp. 70–82.
16. ———, *The cell discretization algorithm for elliptic partial differential equations*, SIAM J. Sci. Statist. Comput. **3** (1982), 261–288.
17. ———, *The application of cell discretization to nuclear reactor problems*, Nuclear Sci. and Engrg. **82** (1982), 78–95.
18. ———, *Cell discretization of nonselfadjoint linear elliptic partial differential equations*, SIAM J. Sci. Statist. Comput. **12** (1991), 1074–1108.
19. P. Grisvard, *Elliptic problems in non-smooth domains*, Pitman, New York, 1985.
20. B. Guo and I. Babuška, *The h-p version of the finite element method. Part* 1: *The basic approximation results*; *Part* 2: *General results and applications*, Comput. Mech. **1** (1986), 21–41, 203–226.
21. P. A. Raviart and J. M. Thomas, *Primal hybrid finite element methods for second order elliptic equations*, Math. Comp. **31** (1977), 391–413.
22. J. R. Rice and R. Boisvert, *Solving elliptic problems using ELLPACK*, Springer-Verlag, Berlin and New York, 1985.
23. G. Strang and G. J. Fix, *An analysis of the finite element method*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
24. H. Swann and M. Nishimura, *Implementation of the cell discretization algorithm for solving elliptic partial differential equations*, CAM-16-88, Center for Applied Mathematics and Computer Science, San José, CA, 1988; revision, 1990.
25. H. Swann, *On the use of Lagrange multipliers in domain decomposition for solving elliptic problems*, Math. Comp. **60** (1993), 49–78.
26. J. Wloka, *Partial differential equations*, Cambridge Univ. Press, Cambridge, 1987.

Department of Mathematics and Computer Science, San José State University, San José, California 95192-0103
E-mail address, H. Swann: swann@sjsumcs.sjsu.edu