A LEAST-SQUARES METHOD FOR AXISYMMETRIC DIV-CURL SYSTEMS

DYLAN M. COPELAND AND JOSEPH E. PASCIAK

Dedicated to Professor Richard E. Ewing on the occasion of his sixtieth birthday.

Abstract. We present a negative-norm least-squares method for axisymmetric div-curl systems arising from Maxwell's equations for electrostatics and magnetostatics in three dimensions. The method approximates the solution in a two-dimensional meridian plane. To achieve this dimension reduction, we must work with weighted spaces in cylindrical coordinates. In this setting, a stable pair of approximation spaces is developed and analyzed. We also report the results of some numerical experiments, which demonstrate a quasi-optimal convergence rate and a robustness with respect to the domain and coefficients.

1. Introduction

Many applications in electromagnetics involve axial symmetry. Considering the high cost of solving Maxwell's equations in three dimensions, there is a need for numerical methods which exploit axial symmetry by achieving a dimension reduction. The strategy of dimension reduction in the presence of rotational symmetry has proved to be effective in finite element methods for the axisymmetric Laplace and Stokes equations (e.g. [1, 2, 10]). In this paper, we present a negative-norm least-squares method for the div-curl systems arising from Maxwell's equations with axisymmetric data and domain Ω in $\mathbb{R}^3$. In this case, the solution is axisymmetric, as proved in [2]. Moreover, the three-dimensional problem may be solved in a two-dimensional meridian domain, reducing the computational cost significantly.

To fix ideas, we consider as a model problem the Maxwell system for electrostatics, defined as follows. Let Ω be a bounded, simply-connected domain in $\mathbb{R}^3$ symmetric with respect to the z-axis, with a connected polyhedral boundary. We assume that the intersection of Ω with the z-axis, denoted $\Gamma_0 \equiv \Omega \cap (\mathbf{0} \times \mathbf{0} \times \mathbb{R})$, is an interval of positive length. If Ω does not intersect the z-axis, then the method presented in this paper is still as effective, but the analysis becomes much simpler and uninteresting. The static Maxwell system for the electric field $\mathbf{e} \in L^2(\Omega)^3$ is

$$\begin{cases} \nabla \times \mathbf{e} & = \mathbf{f} \text{ in } \Omega, \\ \nabla \cdot (\epsilon \mathbf{e}) & = g \text{ in } \Omega, \\ \mathbf{e} \times \mathbf{n} & = \mathbf{0} \text{ on } \partial\Omega. \end{cases} \tag{1.1}$$

Key words and phrases. least squares, axisymmetry, Maxwell's equations.

This work was supported in part by the National Science Foundation through grant No. 0311902.

We assume that the electric permittivity ϵ is piecewise constant and positive and well behaved enough so that solutions to

$$\begin{aligned} -\nabla \cdot \epsilon \nabla u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

are in the Sobolev space $H^{1+s}(\Omega)$ for some $s > 0$.

Further, we assume that the data $\mathbf{f} \in L^2(\Omega)^3$, $g \in L^2(\Omega)$, and ϵ are axisymmetric [2]. By this we mean that (in cylindrical coordinates)

$$\begin{aligned} &g = g(r, z), \qquad \epsilon = \epsilon(r, z), \\ &\mathbf{f} = f_r(r, z)\mathbf{e}_r + f_\theta(r, z)\mathbf{e}_\theta + f_z(r, z)\mathbf{e}_z. \end{aligned}$$

Here $\mathbf{e}_r, \mathbf{e}_\theta$ and $\mathbf{e}_z$ denote the unit cylindrical coordinate vectors. In the literature, scalar functions constant with respect to θ are usually said to be invariant by rotation, but we just use the term axisymmetric for simplicity. We further assume that $\nabla \cdot \mathbf{f} = 0$.

The boundary condition $\mathbf{e} \times \mathbf{n} = \mathbf{0}$ corresponds to a perfect conductor [11]. Inhomogeneous boundary conditions can be handled by simply modifying the data in the weak formulation considered below. Furthermore, the theory and computational methods of this paper can be easily applied to the magnetostatic system, which differs from (1.1) only in the boundary condition $\mathbf{h} \cdot \mathbf{n} = 0$.

Without the assumption of axisymmetry, this three-dimensional problem has been approximated using mixed methods, discretized with the $\mathbf{H}(\mathbf{curl}; \Omega)$ conforming Nedelec elements. Such work can be found, for example, in [4, 7, 8, 9, 12]. The use of mixed methods for the dimension-reduced axisymmetric equations requires an analysis of analogous finite elements in the weighted spaces on the meridian domain. Such methods are currently being studied by the authors. Another successful approach to the general three-dimensional problem is the least-squares method of Bramble and Pasciak [5]. It is in the general framework of that paper that we derive the present least-squares method for the axisymmetric equations. Our method leads to a well conditioned discrete system which can be efficiently solved by a simple conjugate gradient iteration.

There has been some work focused on the efficient solution of axisymmetric Maxwell's equations. For example, geometric multigrid methods have been successfully applied to axisymmetric equations in certain restrictive cases. In [3], Börm and Hiptmair present a multigrid method for the $\mathbf{H}(\mathbf{curl}; \Omega)$-elliptic variational problem arising from discretization of the time-dependent Maxwell equations. Their analysis requires the rather restrictive assumptions that the coefficients of the bilinear form fit a tensor product structure and have constant ratios. Moreover, their method requires mesh structure, as it performs r-line smoothings and semicoarsening in the z-direction.

Gopalakrishnan and Pasciak [10] demonstrated that line smoothing and semicoarsening are unnecessary for the convergence of geometric multigrid for the axisymmetric Laplace and Maxwell equations (azimuthal component only), in the case of constant coefficients. However, the methods of [10] require structured meshes, and the degradation of convergence with variable coefficients is not addressed. It should also be noted that in [10], only the azimuthal component of the vector field is solved for (see equation (1.4)). In this paper, we solve for the meridian components (equation (1.2)).

Under the assumption of axisymmetric data, transforming the system (1.1) to cylindrical coordinates (r, θ, z) yields (see [2]) two decoupled systems in the two-dimensional domain $D = \{(r, z) : (r, 0, z) \in \Omega\}$:

$$\text{(1.2)} \qquad \begin{cases} \nabla \times (e_r, e_z) \equiv \frac{\partial e_r}{\partial z} - \frac{\partial e_z}{\partial r} &= f_\theta \quad \text{in } D, \\ \nabla_r \cdot (\epsilon(e_r, e_z)) \equiv \frac{1}{r}\frac{\partial}{\partial r}(r\epsilon e_r) + \frac{\partial}{\partial z}(\epsilon e_z) &= g \quad \text{in } D, \\ (e_r, e_z) \cdot (-n_z, n_r) &= 0 \quad \text{on } \Gamma_1, \end{cases}$$

and

$$\text{(1.3)} \qquad \begin{cases} -\frac{\partial e_\theta}{\partial z} &= f_r \quad \text{in } D, \\ \frac{1}{r}\frac{\partial}{\partial r}(re_\theta) &= f_z \quad \text{in } D, \\ e_\theta &= 0 \quad \text{on } \Gamma_1, \end{cases}$$

where $\Gamma_1 = \{(r, z) \in \partial D : r > 0\}$. The assumptions on Ω imply that D is a bounded domain in $\mathbb{R}^2$ with a polygonal boundary. Note that D may be nonconvex.

Let $L^2_\alpha(D)$ denotes the weighted Lebesgue space of measurable functions v on D bounded in the norm $\|v\|_{L^2_\alpha(D)} = (\int_D r^\alpha v^2 \, dr \, dz)^{1/2}$. We denote by $H^k_\alpha(D)$ the weighted Sobolev space of functions in $L^2_\alpha(D)$ whose weak derivatives up to order k are in $L^2_\alpha(D)$. The seminorm on $H^k_\alpha(D)$ is denoted $|\cdot|_{H^k_\alpha(D)}$. Further, set $H^1_{1,\diamond}(D) = \{w \in H^1_1(D) : w = 0 \text{ on } \Gamma_1\}$ and $H^1_-(D) = H^1_1(D) \cap L^2_{-1}(D)$. The dual of a Hilbert space H is denoted H'.

The azimuthal component e_θ can be solved for separately in the following scalar equation, obtained by taking the curl of equation (1.3):

$$\text{(1.4)} \qquad \begin{cases} -\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(re_\theta)\right) - \frac{\partial^2 e_\theta}{\partial z^2} &= \nabla \times (f_r, f_z) \quad \text{in } D \\ e_\theta &= 0 \quad \text{on } \Gamma_1. \end{cases}$$

Observe that integration by parts yields a variational formulation for (1.4) that is coercive on the space $\{\phi \in H^1_-(D) : \phi = 0 \text{ on } \partial D\}$, where $H^1_-(D)$ has the norm $\|\phi\|^2_{H^1_-(D)} = \int_D r^{-1}\phi^2 + r\frac{\partial\phi}{\partial r}^2 + r\frac{\partial\phi}{\partial z}^2 \, dr \, dz$. (This boundary condition is equivalent to the one in (1.4), as all functions in $\phi \in H^1_-(D)$ satisfy $\phi = 0$ on $\Gamma_0 = \partial D \setminus \Gamma_1$.) Multigrid methods can be used to numerically solve (1.4), (cf. [3, 10]).

We are interested in solving (1.2) for the meridian components $\mathbf{e} = (e_r, e_z) \in L^2_1(D)^2$. Now the least-squares problem we shall solve is derived (see Section 3) from the following weak formulation: Find $\mathbf{e} \in L^2_1(D)^2$ satisfying

$$\text{(1.5)} \qquad b(\mathbf{e}, (\phi, q)) \equiv (\mathbf{e}, \nabla_r \times \phi)_{L^2_1(D)} + (\epsilon\mathbf{e}, \nabla q)_{L^2_1(D)} = (f_\theta, \phi)_{L^2_1(D)} - (g, q)_{L^2_1(D)},$$

for all $(\phi, q) \in H^1_-(D) \times H^1_{1,\diamond}(D)$. The differential operator $\nabla_r \times$ is defined below in (2.2). Some of the advantages of the method we shall present are evident from this weak formulation. For example, the solution is sought in $L^2_1(D)^2$, and derivatives and boundary conditions are imposed weakly.

We consider a negative-norm least-squares approximation of (1.5). Since it is based on (1.5), we can simply use piecewise constant vector fields for the finite element space. In addition, the spaces $H^1_-(D)$ and $H^1_{1,\diamond}(D)$ can be approximated by linear Lagrange finite elements, enriched by edge and element bubble functions for stability. As we shall see in Sections 4 and 5, this leads to a stable and accurate approximation method with a linear system which is symmetric and positive definite, and can be solved efficiently by the conjugate gradient method preconditioned by a diagonal matrix.

The remainder of the paper is outlined as follows. Preliminaries pertaining to the r-weighted spaces are covered first in Section 2. In Section 3, the least-squares problem is introduced and analyzed, establishing existence and uniqueness of the solution in $L^2_1(D)^2$. In Section 4, we introduce discrete subspaces approximating $L^2_1(D)^2$, $H^1_-(D)$, and $H^1_{1,\diamond}(D)$. Stability of the resulting discrete least-squares system is proved for this choice of subspaces. The implementation of the discrete system is discussed in Section 5. Finally, the results of numerical experiments are reported in Section 6, demonstrating a first order convergence rate and a robustness with respect to the domain and coefficient.

2. Preliminaries on Weighted Spaces

In cylindrical coordinates with the basis representation $\mathbf{v} = v_r\mathbf{e}_r + v_\theta\mathbf{e}_\theta + v_z\mathbf{e}_z$, the differential operators curl, divergence, and gradient have the expressions

$$\nabla\times\mathbf{v} = \left(\frac{1}{r}\frac{\partial v_z}{\partial\theta} - \frac{\partial v_\theta}{\partial z}\right)\mathbf{e}_r + \left(\frac{\partial v_r}{\partial z} - \frac{\partial v_z}{\partial r}\right)\mathbf{e}_\theta + \frac{1}{r}\left(\frac{\partial}{\partial r}(rv_\theta) - \frac{\partial v_r}{\partial\theta}\right)\mathbf{e}_z,$$

$$\nabla\cdot\mathbf{v} = \frac{1}{r}\frac{\partial}{\partial r}(rv_r) + \frac{1}{r}\frac{\partial v_\theta}{\partial\theta} + \frac{\partial v_z}{\partial z},$$

$$\nabla\phi = \frac{\partial\phi}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial\phi}{\partial\theta}\mathbf{e}_\theta + \frac{\partial\phi}{\partial z}\mathbf{e}_z. \tag{2.1}$$

Assuming axisymmetry, all derivatives with respect to θ equal zero. The differential operators curl, divergence, and gradient give rise to r-dependent operators in the meridian domain (the r-z domain). We shall denote such operators with a subscript r when they differ from the standard operators, i.e.,

$$\nabla_r\times\phi = \left(-\frac{\partial\phi}{\partial z}, \frac{1}{r}\left(\frac{\partial}{\partial r}(r\phi)\right)\right) \text{ and } \nabla_r\cdot\mathbf{v} = \frac{1}{r}\frac{\partial}{\partial r}(rv_r) + \frac{\partial v_z}{\partial z}. \tag{2.2}$$

We use the notation $\Gamma_0 = \partial D\setminus\Gamma_1$, $(\cdot,\cdot)_r = (\cdot,\cdot)_{L^2_1(D)}$, $(\cdot,\cdot)_{r,\tau} = (\cdot,\cdot)_{L^2_1(\tau)}$, and $<\cdot,\cdot>_{r,\Gamma_1} = (\cdot,\cdot)_{L^2_1(\Gamma_1)}$. This notation is also used for the inner products on the spaces $L^2_1(D)^2$ and $L^2_1(\tau)^2$ of vector-valued functions. The meaning will be clear from the context. For vectors $\mathbf{v}\in\mathbb{R}^2$, we denote the Euclidean norm $|\mathbf{v}| = (v_r^2 + v_z^2)^{1/2}$.

The Green's formulas stated in the following lemma will be used repeatedly.

Lemma 1. *Let $\mathbf{n} = (n_r, n_z)$ denote the outward unit normal and $\mathbf{t} = (-n_z, n_r)$ denote the unit tangent vector (oriented counterclockwise). Then for all $\mathbf{v}\in H^1_1(D)^2$ and $\phi\in H^1_-(D)$, we have*

$$(\nabla\times\mathbf{v},\phi)_r = (\mathbf{v},\nabla_r\times\phi)_r - <\mathbf{v}\cdot\mathbf{t},\phi>_{r,\Gamma_1}. \tag{2.3}$$

Also, for all $\mathbf{v}\in H^1_-(D)\times H^1_1(D)$ and $\phi\in H^1_1(D)$, we have

$$(\nabla_r\cdot\mathbf{v},\phi)_r = -(\mathbf{v},\nabla\phi)_r + <\mathbf{v}\cdot\mathbf{n},\phi>_{r,\Gamma_1}. \tag{2.4}$$

Proof. Clearly (2.3) holds for $\mathbf{v}\in C^\infty(\overline{D})^2$ and $\phi\in C^\infty(\overline{D})$, with ϕ vanishing in a neighborhood of Γ_0. By Lemma 3.1 of [10], the subspace of functions in $C^\infty(\overline{D})$ vanishing in a neighborhood of Γ_0 is dense in $H^1_-(D)$. The density of $C^\infty(\overline{D})$ in $H^1_1(D)$ is given by Proposition 2.1(1) of [10]. Thus (2.3) follows by a density argument.

Similarly, (2.4) holds for $\mathbf{v}\in C^\infty(\overline{D})^2$ and $\phi\in C^\infty(\overline{D})$, with v_r vanishing in a neighborhood of Γ_0. By density (see the previous paragraph), (2.4) holds on the spaces stated therein. □

3. The Least-Squares Method

Applying Green's formulas (Lemma 1) to (1.2) yields the weak formulation (1.5) of (1.2). This is the basis of our least-squares approximation.

Define the operators **curl** $: L^2_1(D)^2 \to (H^1_-(D))'$, $\mathrm{div}_\epsilon : L^2_1(D)^2 \to (H^1_{1,\diamond}(D))'$, and $B : L^2_1(D)^2 \to (H^1_-(D) \times H^1_{1,\diamond}(D))'$ by

$$\begin{aligned} <\mathbf{curl}\,\mathbf{v}, \psi> &= (\mathbf{v}, \nabla_r \times \psi)_r \quad \text{for all } \mathbf{v} \in L^2_1(D)^2, \psi \in H^1_-(D), \\ <\mathrm{div}_\epsilon \mathbf{v}, \psi> &= (\epsilon \mathbf{v}, \nabla \psi)_r \quad \text{for all } \mathbf{v} \in L^2_1(D)^2, \psi \in H^1_{1,\diamond}(D), \\ B\mathbf{v} &= (\mathbf{curl}\,\mathbf{v}, \mathrm{div}_\epsilon \mathbf{v}) \quad \text{for all } \mathbf{v} \in L^2_1(D)^2. \end{aligned}$$

Thus B satisfies $(B\mathbf{e}, (\phi, q)) = b(\mathbf{e}, (\phi, q))$ for all $\mathbf{e} \in L^2_1(D)^2$ and $(\phi, q) \in H^1_-(D) \times H^1_{1,\diamond}(D)$ (recall that b was defined in (1.5)). Further define the symmetric bilinear form A on $L^2_1(D)^2 \times L^2_1(D)^2$ and the linear functional F by

$$\begin{aligned} A(\mathbf{u}, \mathbf{v}) &\equiv (B\mathbf{u}, B\mathbf{v})_{(H^1_-(D)\times H^1_{1,\diamond}(D))'} \\ &= (\mathbf{curl}\,\mathbf{u}, \mathbf{curl}\,\mathbf{v})_{(H^1_-(D))'} + (\mathrm{div}_\epsilon \mathbf{u}, \mathrm{div}_\epsilon \mathbf{v})_{(H^1_{1,\diamond}(D))'}, \\ <F, \mathbf{v}> &\equiv (f_\theta, \mathbf{curl}\,\mathbf{v})_{(H^1_-(D))'} - (g, \mathrm{div}_\epsilon \mathbf{v})_{(H^1_{1,\diamond}(D))'}, \end{aligned}$$

for all $\mathbf{u}$ and $\mathbf{v}$ in $L^2_1(D)^2$. Then the dual based least-squares formulation of (1.5) is to find $\mathbf{e} \in L^2_1(D)^2$ satisfying

$$A(\mathbf{e}, \mathbf{v}) = <F, \mathbf{v}> \quad \text{for all } \mathbf{v} \in L^2_1(D)^2. \tag{3.1}$$

It is easy to see that B is a bounded operator on $L^2_1(D)^2$. That its inverse is bounded follows from the orthogonal decomposition for the space $L^2_1(D)^2$ given in the next lemma.

Lemma 2. *For any $\mathbf{v} \in L^2_1(D)^2$, there exist $\phi \in H^1_-(D)$ and $q \in H^1_{1,\diamond}(D)$ satisfying $\mathbf{v} = \nabla_r \times \phi + \epsilon \nabla q$. This decomposition is orthogonal with respect to the inner-product $(\epsilon^{-1}\cdot, \cdot)_r$ on $L^2_1(D)^2$.*

Proof. The Green's formula (2.3) and the density of smooth functions yield $(\nabla_r \times \phi, \nabla q)_r = 0$ for any $\phi \in H^1_-(D)$ and $q \in H^1_{1,\diamond}(D)$. In addition, $\|\epsilon^{-1/2}\nabla_r \times \phi\|_r$ and $\|\epsilon^{1/2}\nabla q\|_r$ provide equivalent norms on $H^1_-(D)$ and $H^1_{1,\diamond}(D)$, respectively. Thus, by density, it suffices to prove the result for $\mathbf{v} \in (C^\infty_0(D))^2$. Associated with $\mathbf{v} = (v_r, v_z)$, we have a rotated function

$$\breve{\mathbf{v}} = v_r \mathbf{e}_r + v_z \mathbf{e}_z$$

defined on Ω. By Theorem II.2.6 of [2], $\breve{\mathbf{v}}$ is smooth. Let $\breve{q} \in H^1_0(\Omega)$ solve

$$(\epsilon \nabla \breve{q}, \nabla \theta)_{L^2(\Omega)} = (\breve{\mathbf{v}}, \nabla \theta)_{L^2(\Omega)} = -(\nabla \cdot \breve{\mathbf{v}}, \theta)_{L^2(\Omega)} \quad \text{for all } \theta \text{ in } H^1_0(\Omega),$$

and set

$$\breve{\mathbf{w}} = \breve{\mathbf{v}} - \epsilon \nabla \breve{q}.$$

Since $\breve{q}$ solves an axisymmetric Laplace equation, $\breve{q}$ is axisymmetric and so is $\breve{\mathbf{w}}$. By restricting to $\theta = 0$, we have the decomposition

$$\mathbf{v} = \mathbf{w} + \epsilon \nabla q.$$

The map $\breve{q} \to q$ is an isomorphism (see [2]) from the space of axisymmetric functions in $H^1_0(\Omega)$ to $H^1_{1,\diamond}(D)$, so we have

$$\|q\|_{H^1_1(D)} \le C \|\mathbf{v}\|_{L^2_1(D)^2}.$$

By regularity assumptions on ϵ, we have that $\breve{q}$ is in $H^{1+s}(\Omega)$ for some $s > 0$. This implies that $\breve{\mathbf{w}}$ is in $(H^s(\Omega))^3$. It follows from this and the fact that $\breve{\mathbf{w}}$ is divergence-free in Ω that

$$\int_S \breve{\mathbf{w}} \cdot \mathbf{n} \, ds$$

exists on any surface S generated by rotation of a piecewise smooth curve $\gamma \subset D$ which is nowhere tangent to Γ_0 (the argument is similar to the proof of Lemma 5.38 in [11]). We conclude that

$$\int_S \breve{\mathbf{w}} \cdot \mathbf{n} \, ds = 2\pi \int_\gamma r\mathbf{w} \cdot \mathbf{n} \, ds$$

is finite, where $\mathbf{n}$ denotes the tangent along γ rotated by -90 degrees. If γ is a curve with its endpoints on Γ_0, then

$$2\pi \int_\gamma r\mathbf{w} \cdot \mathbf{n} \, ds = \int_S \breve{\mathbf{w}} \cdot \mathbf{n} \, ds = \int_{\breve{D}_S} \nabla \cdot \breve{\mathbf{w}} \, dx = 0.$$

Here $\breve{D}_S$ is the domain enclosed by S. The same holds true when γ is a closed curve.

For any point $(r_1, z_1) \in D$, let γ be any piecewise smooth path from some point a_0 on Γ_0 to (r_1, z_1). Define

$$\phi(r_1, z_1) = -\frac{1}{r_1} \int_\gamma r\mathbf{w} \cdot \mathbf{n} \, ds.$$

The above considerations show that ϕ is well defined and independent of the path γ and starting point a_0 on Γ_0. We claim that this function ϕ satisfies the requirements of the lemma.

First, we verify that $\nabla_r \times \phi = \mathbf{w}$. For $(r_1, z_1) \in D$ and sufficiently small $h > 0$, consider the straight-line path γ_h from $(r_1, z_1 - h)$ to $(r_1, z_1 + h)$. Then $\mathbf{t} = (0, 1)$ and $\mathbf{n} = (1, 0)$, so

$$-\frac{\partial \phi}{\partial z} = \lim_{h \to 0} \frac{1}{2hr_1} \int_{\gamma_h} r w_r \, ds = \frac{1}{r_1} r_1 w_r = w_r.$$

Now let γ_h be the straight-line path from $(r_1 - h, z_1)$ to $(r_1 + h, z_1)$, so that $\mathbf{t} = (1, 0)$ and $\mathbf{n} = (0, -1)$. Then

$$\frac{1}{r_1} \frac{\partial}{\partial r}(r\phi)(r_1, z_1) = \frac{1}{r_1} \lim_{h \to 0} \frac{1}{2h} \int_{\gamma_h} r w_z \, ds = \frac{1}{r_1} r_1 w_z = w_z, \tag{3.2}$$

and we conclude that $\nabla_r \times \phi = \mathbf{w}$.

Next we show that ϕ is in $L^2_{-1}(D)$ by using the Hardy inequality (see Appendix A.4 of [13])

$$\left(\int_0^\infty \left(\int_0^x f(y) \, dy \right)^p x^{-k-1} \, dx \right)^{1/p} \leq \frac{p}{k} \left(\int_0^\infty (xf(x))^p x^{-k-1} \, dx \right)^{1/p}. \tag{3.3}$$

We illustrate the argument when the strip $[0, R] \times \Gamma_0$ is contained in D for some $R > 0$. Simple modifications of this argument give the general case. For $z \in \Gamma_0$, applying the Hardy inequality gives

$$\begin{aligned} \int_0^R r_1^{-1} \phi(r_1, z)^2 \, dr_1 &= \int_0^R r_1^{-3} \bigg(\int_0^{r_1} r\mathbf{w}_z(r, z) \, dr \bigg)^2 \, dr_1 \\ &\leq \int_0^R r |\mathbf{w}_z(r, z)|^2 \, dr. \end{aligned}$$

Integrating over z gives

$$\|\phi\|^2_{L^2_{-1}([0,R]\times\Gamma_0)} \le \|\mathbf{w}\|^2_{L^2_1([0,R]\times\Gamma_0)^2} \le C\|\mathbf{v}\|^2_{L^2_1([0,R]\times\Gamma_0)^2}.$$

Thus $\|\phi\|_{L^2_{-1}(D)} \le C\|\mathbf{v}\|_{L^2_1(D)^2}$. Using the identity (3.2), we compute

$$\frac{\partial\phi}{\partial r} = -\frac{\phi}{r} + w_z.$$

That $\frac{\partial\phi}{\partial r}$ is in $L^2_1(D)$ follows from the facts that w_z is in $L^2_1(D)$ and ϕ is in $L^2_{-1}(D)$. Finally,

$$\frac{\partial\phi}{\partial z} = -w_r$$

is also in $L^2_1(D)$. This completes the proof of the lemma. □

The main result of this section now follows.

Theorem 1. *For all* $\mathbf{v} \in L^2_1(D)^2$,

$$(\epsilon\mathbf{v},\mathbf{v})_r = \sup_{(\phi,q)\in H^1_-(D)\times H^1_{1,\diamond}(D)} \frac{(\mathbf{v},\nabla_r\times\phi)^2_r}{\|\epsilon^{-1/2}\nabla_r\times\phi\|^2_r} + \frac{(\epsilon\mathbf{v},\nabla q)^2_r}{\|\epsilon^{1/2}\nabla q\|^2_r}.$$

Proof. Applying Lemma 2 to $\epsilon\mathbf{v} \in L^2_1(D)^2$ yields $\phi \in H^1_-(D)$ and $q \in H^1_{1,\diamond}(D)$ satisfying $\mathbf{v} = \epsilon^{-1}\nabla_r\times\phi + \nabla q$. This decomposition is orthogonal in the inner product $(\epsilon\cdot,\cdot)_r$. Thus,

$$(\epsilon\mathbf{v},\mathbf{v})_r = (\epsilon\mathbf{v},\epsilon^{-1}\nabla_r\times\phi + \nabla q)_r = \frac{(\mathbf{v},\nabla_r\times\phi)^2_r}{\|\epsilon^{-1/2}\nabla_r\times\phi\|^2_r} + \frac{(\epsilon\mathbf{v},\nabla q)^2_r}{\|\epsilon^{1/2}\nabla q\|^2_r}.$$

It follows from the Schwarz inequality and the $(\epsilon\cdot,\cdot)_r$-orthogonality of the decomposition that taking the supremum here preserves the equality. □

Remark 1. *The norm* $(\epsilon^{-1}\nabla_r\times\phi,\nabla_r\times\phi)^{1/2}_r$ *provides an equivalent norm on* $H^1_-(D)$ *while* $(\epsilon\nabla q,\nabla q)^{1/2}_r$ *defines an equivalent norm on* $H^1_{1,\diamond}(D)$. *If we use these norms to define the dual spaces then the above theorem can be restated*

$$(\epsilon\mathbf{v},\mathbf{v})_r = \|\mathbf{curl}\,\mathbf{v}\|^2_{(H^1_-(D))'} + \|div_\epsilon\mathbf{v}\|^2_{(H^1_{1,\diamond}(D))'}.$$

This immediately implies that the existence and uniqueness of the solution to the least-squares problem (3.1).

Corollary 1. *Using the norm* $(\epsilon\cdot,\cdot)^{1/2}$ *on* $(L^2_1(D))^2$ *and the above norms on the dual spaces, the operator* $B : (L^2_1(D))^2 \to (H^1_-(D))' \times (H^1_{1,\diamond}(D))'$ *is an isometry. Moreover,* $B^{-1}(f_\theta, g) = \mathbf{e}$ *where* $\mathbf{e}$ *is the unique solution of* (3.1).

Proof. By Remark 1 and the generalized Lax-Milgram lemma (see, e.g. [5]), B is an isometry onto its image in $(H^1_-(D)\times H^1_{1,\diamond}(D))'$. We need only check that it is onto $(H^1_-(D))'\times(H^1_{1,\diamond}(D))'$. It suffices to show that the only pair of functions $\phi \in H^1_-(D)$ and $q \in H^1_{1,\diamond}(D)$ satisfying

$$b(\mathbf{v},(\phi,q)) = 0 \quad \text{for all } \mathbf{v} \in (L^2_1(D))^2$$

is $(\phi,q) = (0,0)$. This is immediate since setting $\mathbf{v} = \epsilon^{-1}\nabla_r\times\phi + \nabla q$ gives

$$b(\mathbf{v},(\phi,q)) = (\epsilon^{-1}\nabla_r\times\phi,\nabla_r\times\phi)_r + (\epsilon\nabla q,\nabla q)_r.$$

That $B^{-1}(f_\theta,g) = \mathbf{e}$ is immediate from the uniqueness of solutions to (3.1). □

4. Stable Approximation

In this section, we describe a stable pair of approximation spaces for the least-squares method (3.1). Let $\mathcal{T}_h$ be a quasi-uniform triangulation of D aligned with the discontinuities of ϵ. The diameter of a triangle τ in $\mathcal{T}_h$ is denoted by h_τ and the length of an edge e in $\mathcal{T}_h$ by h_e. Denote by $\mathbf{X}_h$ the space of piecewise constant vector fields in $L^2_1(D)^2$ and by S^h_- the space of piecewise linear functions in $H^1_-(D)$. By Lemma 3 of [1], S^h_- consists of the continuous piecewise linear functions which vanish on Γ_0. Let $S^h_\diamond$ be the space of piecewise linear functions in $H^1_{1,\diamond}(D)$.

For each edge $e \in \mathcal{T}_h$, let τ_e be a triangle having e as an edge and denote by $\lambda_i(\mathbf{x})$, $i = 1, 2, 3$, the barycentric coordinate for $\mathbf{x} \in \tau$, where $\lambda_3(\mathbf{x})$ corresponds to the vertex not in e. Define B_e as the edge bubble space spanned by $\lambda_1(\mathbf{x})\lambda_2(\mathbf{x})$ and set $H^h_{e,-} \equiv \oplus_{e\not\subseteq\Gamma_0} B_e$, $H^h_{e,\diamond} \equiv \oplus_{e\not\subseteq\Gamma_1} B_e$. For each $\tau \in \mathcal{T}_h$, define B_τ as the element bubble space spanned by $\lambda_1(\mathbf{x})\lambda_2(\mathbf{x})\lambda_3(\mathbf{x})$. Denote $H^h_\tau \equiv \oplus_{\tau\in\mathcal{T}_h} B_\tau$. We then define $H^h_- \equiv S^h_- \oplus H^h_{e,-}$ and $H^h_\diamond \equiv S^h_\diamond \oplus H^h_{e,\diamond} \oplus H^h_\tau$.

The main result in this section is that the above subspaces provide stable approximation pairs. This is given in the following theorem. Its proof will be developed in the remainder of this section as a series of lemmas. The inf-sup condition given in this theorem yields existence and uniqueness of solutions to the discrete least-squares problem (see Section 5).

Theorem 2. *There exists a constant $C > 0$ (not depending on ϵ) such that*

$$\|\mathbf{v}\|_{L^2_1(D)^2} \le C \left(\sup_{(\phi,q)\in H^h_-\times H^h_\diamond} \frac{(\mathbf{v}, \nabla_r \times \phi)_r}{\|\phi\|_{H^1_-(D)}} + \frac{(\epsilon\mathbf{v}, \nabla q)_r}{\|q\|_{H^1_1(D)}} \right) \text{ for all } \mathbf{v} \in \mathbf{X}_h.$$

In the remainder of the paper, C represents a generic positive constant which may depend on the minimal angle of the mesh $\mathcal{T}_h$ but does not depend on the sizes h_τ or h_e.

Lemma 3. *There exists a constant $C > 0$ such that if b_e is a bubble function associated with an edge e in $\mathcal{T}_h$ not contained in Γ_0,*

$$\|b_e\|^2_{L^2_1(\tau)} \le C r_\tau^{-1} \left(\int_e r b_e \, ds \right)^2$$

Here τ is either triangle having e as an edge and r_τ is the maximum value of r on τ.

Proof. Let $\widehat{B}_e$ denote the edge bubble space on a reference edge of unit length. For edges not intersecting Γ_0, a standard scaling argument and the equivalence of norms on the one-dimensional space $\widehat{B}_e$ gives that for $b_e \in B_e$,

$$\|b_e\|^2_{L^2_1(\tau)} \le r_\tau \|b_e\|^2_{L^2(\tau)} \le C r_\tau \left(\int_e b_e \, ds \right)^2.$$

Since e does not intersect Γ_0, r_τ can be bounded by a constant times the minimum value of r on e, so

$$\left(\int_e b_e \, ds \right)^2 \le C r_\tau^{-2} \left(\int_e r b_e \, ds \right)^2.$$

We next consider an edge e which intersects Γ_0 at a point which we denote by $\mathbf{a}_1$. There are two cases. If τ intersects Γ_0 only at $\mathbf{a}_1$ then there are constants C_0, C_1 depending on quasi-uniformity such that

$$C_0 r \le h_\tau(\lambda_2+\lambda_3) \le C_1 r \quad \text{for all } (r,z)\in\tau.$$

Here λ_2 and λ_3 are the barycentric coordinates of (r,z) associated with the two vertices of τ not on Γ_0. Scaling and again using equivalence of norms on $\widehat{B}_e$ gives that for $b_e \in B_e$,

$$\begin{aligned}\|b_e\|^2_{L^2_1(\tau)} &\le Ch_\tau \|(\lambda_1+\lambda_2)^{1/2} b_e\|^2_{L^2(\tau)}\\ &\le Ch_\tau^3\bigg(\int_{\hat e}(\widehat{\lambda}_2+\widehat{\lambda}_3)\hat b_e\,ds\bigg)^2 \le Ch_\tau^{-1}\bigg(\int_e r b_e\,ds\bigg)^2.\end{aligned} \tag{4.1}$$

The remaining case is when τ intersects Γ_0 along the edge with endpoints $\mathbf{a}_1$ and $\mathbf{a}_2$. Then on τ, $r=\alpha_\tau h_\tau \lambda_3$ where, because of quasi-uniformity, $0<C_0\le\alpha_\tau\le C_1$. Replacing $\lambda_2+\lambda_3$ by λ_3 in (4.1) completes the proof. □

Lemma 4. *Let τ be a triangle in $\mathcal{T}_h$ and e be an edge of τ not on Γ_0. There exists a constant $C>0$ such that for all $u\in H^1_1(\tau)$*

$$\|u\|^2_{L^2_1(e)} \le C(h_e^{-1}\|u\|^2_{L^2_1(\tau)} + h_e\|u\|^2_{H^1_1(\tau)}).$$

Proof. By Proposition 2.1 of [10], $C^\infty(\overline{\tau})$ is dense in $H^1_1(\tau)$ and the trace operator is continuous from $H^1_1(\tau)$ to $L^2_1(e)$. Therefore, it suffices to prove the result for $u\in C^\infty(\overline{\tau})$. When τ does not intersect Γ_0, the result easily follows from the standard (unweighted) estimate.

We next consider the case when $e\cap\Gamma_0$ is a point. For each point $\mathbf{x}=(x_r,x_z)\in e$, we let $\eta_{\mathbf{x}}$ denote a unit vector pointing from $\mathbf{x}$ to the vertex $\mathbf{a}_3$ of τ not on e. There is a positive number α independent of h such that $\mathbf{x}+\alpha h\eta_{\mathbf{x}}$ is in τ and is of distance greater than Ch from $\mathbf{a}_3$. It follows that the value of r on the line from $\mathbf{x}$ to $\mathbf{x}+\alpha h\eta_{\mathbf{x}}$ is bounded above and below by a constant (independent of h) multiple of x_r. We write

$$u(\mathbf{x})^2 = -\int_0^t \frac{\partial}{\partial y}(u^2)(\mathbf{x}+y\eta_{\mathbf{x}})\,dy + u(\mathbf{x}+t\eta_{\mathbf{x}})^2 \text{ for all } 0<t<\alpha h.$$

Multiplying the above equation by $r=x_r$, using the above equivalence and integrating over e gives

$$\begin{aligned}\|u\|^2_{L^2_1(e)} \le C\int_e\int_0^t (\mathbf{x}(s)+y\eta_{\mathbf{x}})_r\left|\frac{\partial}{\partial y}(u^2)(\mathbf{x}(s)+y\eta_{\mathbf{x}})\right|\,dy\,ds\\ + C\int_e (\mathbf{x}(s)+t\eta_{\mathbf{x}})_r u(\mathbf{x}(s)+t\eta_{\mathbf{x}})^2\,ds.\end{aligned}$$

By quasi-uniformity, the angle between e and $\eta_{\mathbf{x}}$ does not degenerate, so changing the integration variable and applying the Schwarz inequality along with integration over t in $(0,\alpha h)$ gives

$$\alpha h\|u\|^2_{L^2_1(e)} \le C(h\|u\|_{L^2_1(\tau)}\|\nabla u\|_{L^2_1(\tau)} + \|u\|^2_{L^2_1(\tau)}),$$

from which the desired bound immediately follows.

The same argument handles the remaining case as well. When τ intersects Γ_0 and e does not, r behaves like h_τ on e. Fortunately, since $\mathbf{x}+\alpha h_\tau\eta_{\mathbf{x}}$ stays away from $\mathbf{a}_3\in\Gamma_0$, $(\mathbf{x}+y\eta_{\mathbf{x}})_r$ behaves like h_τ for $0\le y\le \alpha h_\tau$. The above argument gives the desired result. This completes the proof of the lemma. □

Lemma 5. *There exists a constant $C > 0$ independent of h such that*

$$\sup_{\phi \in H^1_-(D)} \frac{(\mathbf{v}, \nabla_r \times \phi)_r}{\|\phi\|_{H^1_-(D)}} \le C \sup_{\phi \in H^h_-} \frac{(\mathbf{v}, \nabla_r \times \phi)_r}{\|\phi\|_{H^1_-(D)}} \text{ for all } \mathbf{v} \in \mathbf{X}_h.$$

Proof. Using $L^2_1(\tau_i)$-orthogonal projectors, it is possible to construct a Clément-like projector $\Pi^h : H^1_-(D) \to S^h_-$ satisfying (cf. [1])

$$h_\tau^{-2}\|\phi - \Pi^h\phi\|^2_{L^2_1(\tau)} + \|\phi - \Pi^h\phi\|^2_{H^1_-(\tau)} \le C\|\phi\|^2_{H^1_-(\Delta_\tau)}, \tag{4.2}$$

where Δ_τ denotes the union of all triangles in $\mathcal{T}_h$ sharing a common vertex with τ.

Let $\mathbf{v} \in \mathbf{X}_h$ and $\phi \in H^1_-(D)$ be given and set $\psi = \phi - \Pi^h\phi$. For each edge e in $\mathcal{T}_h$ not contained in Γ_0, define $w_e \in B_e$ by $\int_e r w_e\, ds = \int_e r\psi\, ds$. Set $q_e = \sum_{e \not\subseteq \Gamma_0} w_e$. The function q_e is constructed so that $\phi_h = \Pi^h\phi + q_e$ satisfies

$$(\mathbf{v}, \nabla_r \times \phi)_r = (\mathbf{v}, \nabla_r \times \phi_h)_r. \tag{4.3}$$

Indeed, since $\mathbf{v}$ is piecewise constant, $\mathbf{v}\cdot\mathbf{t}$ is constant on each edge and (2.3) gives

$$\begin{aligned}(\mathbf{v}, \nabla_r \times \psi)_{r,\tau} &= < \mathbf{v}\cdot\mathbf{t}, \psi >_{r,\partial\tau\setminus\Gamma_0} \\ &= < \mathbf{v}\cdot\mathbf{t}, q_e >_{r,\partial\tau\setminus\Gamma_0} = (\mathbf{v}, \nabla_r \times q_e)_{r,\tau}.\end{aligned}$$

Summing over all $\tau \in \mathcal{T}_h$ yields (4.3). Thus, the lemma will follow if we show that

$$\|\phi_h\|_{H^1_-(D)} \le C\|\phi\|_{H^1_-(D)}.$$

By (4.2), it suffices to show

$$\|q_e\|_{H^1_-(D)} \le C\|\phi\|_{H^1_-(D)}. \tag{4.4}$$

By Lemma 3 of [1], for any polynomial $f \in P_k(\tau)$ vanishing on Γ_0 when $\tau \cap \Gamma_0 \neq \emptyset$,

$$\|f\|_{L^2_{-1}(\tau)} \le C h_\tau^{-1}\|f\|_{L^2_1(\tau)}. \tag{4.5}$$

Thus,

$$\|q_e\|^2_{L^2_{-1}(D)} \le C \sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} \|w_e\|^2_{L^2_{-1}(\tau)} \le C \sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} h_\tau^{-2}\|w_e\|^2_{L^2_1(\tau)}.$$

Applying Lemma 3 gives

$$\begin{aligned}\sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} h_\tau^{-2}\|w_e\|^2_{L^2_1(\tau)} &\le C \sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} r_\tau^{-1}h_\tau^{-2}\left(\int_e r w_e\, ds\right)^2 \\ &= C \sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} r_\tau^{-1}h_\tau^{-2}\left(\int_e r\psi\, ds\right)^2 \\ &\le C \sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} h_\tau^{-1}\|\psi\|^2_{L^2_1(e)}.\end{aligned} \tag{4.6}$$

Combining the above inequalities with Lemma 4 and (4.2) gives

$$\sum_{\tau\in\mathcal{T}_h} \sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} h_\tau^{-2}\|w_e\|^2_{L^2_1(\tau)} \le C(h_\tau^{-2}\|\psi\|^2_{L^2_1(D)} + \|\psi\|^2_{H^1_1(D)}) \le C\|\phi\|^2_{H^1_-(D)}, \tag{4.7}$$

from which it follows that

$$\|q_e\|_{L^2_{-1}(D)} \le C\|\phi\|_{H^1_-(D)}.$$

By Lemma 4 of [1], for any polynomial $f \in P_k(\tau)$ we have

$$\|f\|_{H^1_1(\tau)} \le Ch_\tau^{-1}\|f\|_{L^2_1(\tau)}. \tag{4.8}$$

Thus,

$$\|q_e\|^2_{H^1_1(D)} \le C\sum_{\tau\in\mathcal{T}_h}\sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}}\|w_e\|^2_{H^1_1(\tau)} \le C\sum_{\tau\in\mathcal{T}_h}\sum_{\substack{e\subset\overline{\tau}\\ e\not\subseteq\Gamma_0}} h_\tau^{-2}\|w_e\|^2_{L^2_1(\tau)}.$$

Applying (4.7) shows

$$\|q_e\|^2_{H^1_1(D)} \le C\|\phi\|_{H^1_-(D)}.$$

Thus we have proved (4.4), which completes the proof of the Lemma. □

The analogous result for the second term in the discrete inf-sup condition (Theorem 2) is contained in the following lemma. Its proof is given later.

Lemma 6. *There exists a constant $C > 0$ independent of h such that*

$$\sup_{q\in H^1_{1,\diamond}(D)}\frac{(\epsilon\mathbf{v},\nabla q)_r}{\|q\|_{H^1_1(D)}} \le C\sup_{q\in H^h_\diamond}\frac{(\epsilon\mathbf{v},\nabla q)_r}{\|q\|_{H^1_1(D)}} \text{ for all } \mathbf{v}\in\mathbf{X}_h.$$

In the proof of Lemma 5, we applied the Green's formula (2.3) with $\mathbf{v}$ piecewise constant. In this case, $\mathbf{v}$ was in the appropriate space, namely $H^1_1(\tau)^2$. Similarly, the proof of Lemma 6 will require the Green's formula for the divergence, but (2.4) does not apply since constant vectors $\mathbf{v}$ do not satisfy $v_r \in H^1_-(\tau)$ for triangles τ with nodes on Γ_0. However, the Green's formula for the divergence does hold in the specific cases we require for proving Lemma 6. This is stated precisely in the following lemma.

Lemma 7. *If τ is a triangle with no edges contained in Γ_0 (i.e., $\overline{\tau}\cap\Gamma_0$ is empty or a vertex), then*

$$(\nabla_r\cdot\mathbf{v},\phi)_{r,\tau} = -(\mathbf{v},\nabla\phi)_{r,\tau} + <\mathbf{v}\cdot\mathbf{n},\phi>_{r,\partial\tau} \tag{4.9}$$

for all $\phi\in H^1_1(D)$ and constant $\mathbf{v}\in\mathbb{R}^2$.

Proof. The case when τ does not intersect Γ_0 is already contained in Lemma 1. Suppose that τ intersects Γ_0 at a point. When ϕ is $C^1(\tau)$, the above formula follows immediately from the divergence theorem, i.e.,

$$\int_\tau \nabla\cdot(r\phi\mathbf{v})\,dx = \int_{\partial\tau} r\phi\mathbf{v}\cdot\mathbf{n}\,ds.$$

Thus, by density, it suffices to show that each term in (4.9) is bounded for $\phi\in H^1_1(\tau)$. Applying the Schwarz inequality gives

$$|(\mathbf{v},\nabla\phi)_{r,\tau}| \le \|\mathbf{v}\|_{L^2_1(\tau)}\|\nabla\phi\|_{L^2_1(\tau)}.$$

It follows from Lemma 4 that

$$\begin{aligned}|<\mathbf{v}\cdot\mathbf{n},\phi>_{r,\partial\tau}| &\le \|\mathbf{v}\cdot\mathbf{n}\|_{L^2_1(\partial\tau)}\|\phi\|_{L^2_1(\partial\tau)}\\ &\le C\|\mathbf{v}\cdot\mathbf{n}\|_{L^2_1(\partial\tau)}(h_\tau^{-1}\|\phi\|^2_{L^2_1(\tau)} + h_\tau\|\phi\|^2_{H^1_1(\tau)})^{1/2}.\end{aligned}$$

Since $\mathbf{v}$ is a constant and the width of τ in the z-direction is bounded by Cr, it follows from the Schwarz inequality that

$$|(\nabla_r \cdot \mathbf{v}, \phi)_{r,\tau}| = \left| \int_\tau v_r \phi \, dx \right| \le C(\tau)|v_r| \|\phi\|_{L^2_1(\tau)}. \tag{4.10}$$

Thus all of the terms in (4.9) are bounded. This completes the proof of the lemma. □

Proof of Lemma 6. Given $\phi \in H^1_{1,\diamond}(D)$ and $\mathbf{v} \in \mathbf{X}_h$, we shall construct $\phi_h \in H^h_\diamond$ satisfying

$$\begin{aligned} (\mathbf{v}, \nabla\phi)_r &= (\mathbf{v}, \nabla\phi_h)_r \\ \|\phi_h\|_{H^1_1(D)} &\le C\|\phi\|_{H^1_1(D)}. \end{aligned} \tag{4.11}$$

Again, we use a Clément-like operator $\Pi^h_\diamond : H^1_{1,\diamond}(D) \to S^h_\diamond$ satisfying (see [1])

$$h_\tau^{-2}\|\phi - \Pi^h_\diamond\phi\|^2_{L^2_1(\tau)} + \|\phi - \Pi^h_\diamond\phi\|^2_{H^1_1(\tau)} \le C\|\phi\|^2_{H^1_1(\Delta_\tau)}. \tag{4.12}$$

Set $\psi = \phi - \Pi^h_\diamond\phi$. As in Lemma 5, for each edge e in $\mathcal{T}_h$ not contained in ∂D, define $w_e \in B_e$ by $\int_e r w_e \, ds = \int_e r\psi \, ds$ and set $q_e = \sum_e w_e$. Lemma 7 gives that for triangles without edges on Γ_0,

$$\begin{aligned} (\mathbf{v}, \nabla\psi)_{r,\tau} &= -v_r \int_\tau \psi \, dx + < \mathbf{v}\cdot\mathbf{n}, \psi >_{r,\partial\tau} \\ &= -v_r \int_\tau \psi \, dx + < \mathbf{v}\cdot\mathbf{n}, q_e >_{r,\partial\tau} . \end{aligned}$$

Next define $w_\tau \in B_\tau$ by

$$\int_\tau w_\tau \, dx = \int_\tau (\psi - q_e) \, dx.$$

Then

$$(\mathbf{v}, \nabla\psi)_{r,\tau} = -v_r \int_\tau (w_\tau + q_e) \, dx + < \mathbf{v}\cdot\mathbf{n}, w_\tau + q_e >_{r,\partial\tau} = (\mathbf{v}, \nabla(w_\tau + q_e))_{r,\tau}.$$

Accordingly, we define $\phi_h = \Pi^h_\diamond\phi + q_e + w_\tau$ on τ when τ does not have an edge on Γ_0. Then we have

$$(\mathbf{v}, \nabla\phi)_{r,\tau} = (\mathbf{v}, \nabla\phi_h)_{r,\tau}. \tag{4.13}$$

To deal with the remaining triangles, we have to avoid integration by parts in the r direction with terms involving ψ. Specifically, for a triangle τ with an edge e_0 on Γ_0, we choose the bubble function $w_{e_0} \in B_{e_0}$ so that

$$(\mathbf{v}, \nabla\phi)_{r,\tau} = (\mathbf{v}, \nabla(\Pi^h_\diamond\phi + q_e + w_{e_0}))_{r,\tau} \tag{4.14}$$

and set $\phi_h = \Pi^h_\diamond\phi + q_e + w_{e_0} \equiv \tilde{\phi}_h + w_{e_0}$. We shall see that (4.14) is indeed possible. We have

$$\begin{aligned} (1, \frac{\partial\phi}{\partial z})_{r,\tau} &= (1, \frac{\partial\phi_h}{\partial z})_{r,\tau} + \int_{\partial\tau\setminus\Gamma_0} n_z r(\phi - \tilde{\phi}_h) \, ds, \\ (1, \frac{\partial w_{e_0}}{\partial r})_{r,\tau} &= -\int_\tau w_{e_0} \, dx. \end{aligned}$$

Here n_z denotes the z-component of $\mathbf{n}$. Using the first equality above and the definition of q_e gives

$$\begin{aligned}\left(v_z, \frac{\partial(\phi-\phi_h)}{\partial z}\right)_{r,\tau} &= \int_{\partial\tau\setminus\Gamma_0} v_z n_z r(\phi-\tilde{\phi}_h)\,ds \\ &= \int_{\partial\tau\setminus\Gamma_0} v_z n_z r(\psi - q_e)\,ds = 0.\end{aligned}$$

Thus

$$(\mathbf{v}, \nabla\phi)_{r,\tau} = (\mathbf{v}, \nabla\phi_h)_{r,\tau} + v_r\bigg(\int_\tau r\frac{\partial}{\partial r}(\phi-\tilde{\phi}_h)\,dx + \int_\tau w_{e_0}\,dx\bigg). \tag{4.15}$$

Choosing $w_{e_0} \in B_{e_0}$ so that

$$\int_\tau w_{e_0}\,dx = \int_\tau r\frac{\partial}{\partial r}(\tilde{\phi}_h - \phi)\,dx \tag{4.16}$$

gives (4.14) on triangles τ with an edge on Γ_0. Combining (4.13) and (4.14) gives the equality of (4.11).

To complete the proof, we need only show that the inequality of (4.11) holds. By (4.12) and the triangle inequality, we need only bound norms of the functions w_τ, q_e and w_{e_0}. The arguments in the proof of Lemma 5 with (4.12) give

$$\|q_e\|_{H^1_1(\tau)} \le C\|\phi\|_{H^1_1(\Delta_\tau)} \qquad \text{and} \qquad \|q_e\|_{L^2_1(\tau)} \le Ch_\tau\|\phi\|_{H^1_1(\Delta_\tau)}. \tag{4.17}$$

For the remaining terms, there are two cases.

First, we consider τ such that τ does not intersect Γ_0 on an edge. Applying a scaling argument and (4.8) gives

$$\begin{aligned}\|w_\tau\|^2_{H^1_1(\tau)} &\le Ch_\tau^{-2}\|w_\tau\|^2_{L^2_1(\tau)} \le Cr_\tau h_\tau^{-4}\bigg(\int_\tau w_\tau\,dx\bigg)^2 \\ &= Cr_\tau h_\tau^{-4}\bigg(\int_\tau(\psi-q_e)\,dx\bigg)^2\end{aligned} \tag{4.18}$$

If τ intersects Γ_0, we use the fact that the width of τ in the z-direction is bounded by Cr to obtain

$$\bigg(\int_\tau(\psi-q_e)\,dx\bigg)^2 \le Cr_\tau\|\psi-q_e\|^2_{L^2_1(\tau)} \le Cr_\tau h_\tau^2\|\phi\|^2_{H^1_1(\Delta_\tau)}. \tag{4.19}$$

To get the last inequality above, we used (4.6), (4.12), and the triangle inequality. It follows that

$$\|w_\tau\|_{H^1_1(\tau)} \le C\|\phi\|_{H^1_1(\Delta_\tau)} \tag{4.20}$$

in this case. Otherwise, if $\tau \cap \Gamma_0 = \phi$,

$$\bigg(\int_\tau(\psi-q_e)\,dx\bigg)^2 \le Cr_\tau^{-1}h_\tau^2\|\psi-q_e\|^2_{L^2_1(\tau)} \le Cr_\tau^{-1}h_\tau^4\|\phi\|^2_{H^1_1(\Delta_\tau)}$$

and (4.20) immediately follows.

We finally consider triangles τ with edges on Γ_0. As in (4.18),

$$\|w_{e_0}\|^2_{H^1_1(\tau)} \le Ch_\tau^{-5}\bigg(\int_\tau w_{e_0}\,dx\bigg)^2 \le Ch_\tau^{-2}\|\psi-q_e\|^2_{L^2_1(\tau)}.$$

The term on the right hand side is bounded as in (4.19) and we obtain

$$\|w_{e_0}\|_{H_1^1(\tau)} \leq C\|\phi\|_{H_1^1(\Delta_\tau)}. \tag{4.21}$$

Combining (4.12), (4.17), (4.20), and (4.21) proves the inequality in (4.11). This completes the proof of the lemma. □

Theorem 2 immediately follows from Lemmas 5 and 6 and Remark 1.

5. Discrete Least-Squares System

Let $\mathbf{X}_h$ and $Y_h = H_-^h \times H_\diamond^h$ be the stable approximation pair introduced in the previous section and define the operator $B_h : \mathbf{X}_h \to Y_h'$ by

$$< B_h\mathbf{x}, y >= b(\mathbf{x}, y) \text{ for all } \mathbf{x} \in \mathbf{X}_h, y \in Y_h.$$

Further define $T_{Y_h} : Y_h' \to Y_h$ by $(T_{Y_h} f, y)_{Y_h} =< f, y >$ for all $y \in Y_h$. Then the discrete least-squares problem is to find $\mathbf{x}_h \in \mathbf{X}_h$ such that

$$A_h(\mathbf{x}_h, \mathbf{x}) \equiv< B_h\mathbf{x}_h, T_{Y_h} B_h\mathbf{x} >=< F, T_{Y_h} B_h\mathbf{x} > \text{ for all } \mathbf{x} \in \mathbf{X}_h. \tag{5.1}$$

Using Theorem 2, it is easy to see that

$$C_0 A_h(\mathbf{x}, \mathbf{x}) \leq (\epsilon\mathbf{x}, \mathbf{x})_r \leq C_1 A_h(\mathbf{x}, \mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbf{X}_h \tag{5.2}$$

with constants C_0, C_1 independent of h. Accordingly, the discrete system corresponding to (5.1) can be uniformly preconditioned by the matrix corresponding to $(\epsilon\cdot, \cdot)_r$ on $\mathbf{X}_h$. The following theorem is an immediate consequence of Theorem 2 (see [5]).

Theorem 3. *The problem* (5.1) *has a unique solution which satisfies*

$$\|\mathbf{e} - \mathbf{x}_h\|_r \leq C \inf_{\mathbf{v}\in\mathbf{X}_h} \|\mathbf{e} - \mathbf{v}\|_r.$$

Here $\mathbf{e}$ *is the unique solution of* (1.5).

Remark 2. *The above theorem immediately implies that when the solution is in* $H_1^s(D)^2$, *for* $0 < s \leq 1$, *the convergence rate will be at least of order* s.

Remark 3. *In the case when* ϵ *has large jumps away from* Γ_0, *Clément operators can be constructed satisfying* (4.2) *and* (4.12) *with* C *replaced by* $C\ln(h^{-1})^2$ *(see,* [6]*), where* C *is independent of the magnitude of the jumps in* ϵ. *In this case,* (5.2) *holds with* $C_1/C_0 \leq C\ln(h^{-1})^2$.

In our computations, instead of T_{Y_h} we use a spectrally equivalent preconditioner $T^h : Y_h' \to Y_h$. In this case, the system remains well conditioned and the quasi-optimal estimate of Theorem 3 still holds. Specifically, we define $T^h f = (T_-^h f, T_+^h f)$ where the action of T_-^h and T_+^h involve smoothing on the bubble spaces and a multigrid algorithm with point smoothing and geometric refinement (cf., [10]) on the spaces S_-^h and $S_\diamond^h$. To implement this, we assemble matrices A_- and A_+ in the bases of H_-^h and $H_\diamond^h$ according to the partition

$$A_- = \begin{bmatrix} A_-^{bb} & A_-^{bl} \\ A_-^{lb} & A_-^{ll} \end{bmatrix}, \; A_+ = \begin{bmatrix} A_+^{bb} & A_+^{bl} \\ A_+^{lb} & A_+^{ll} \end{bmatrix},$$

where b indicates the space of bubble functions and l the space of piecewise linear functions. The bilinear forms defining A_- and A_+ are, respectively,

$$a_-(u,v) = \int_D \frac{1}{r\epsilon}\partial_r(ru)\partial_r(rv)\,dr\,dz + \int_D \frac{r}{\epsilon}\partial_z u\partial_z v\,dr\,dz,$$
$$a_+(u,v) = \int_D r\epsilon(\partial_r u\partial_r v + \partial_z u\partial_z v)\,dr\,dz.$$

It is shown in [10] that these bilinear forms are continuous and coercive on their respective spaces. The multigrid algorithm is used to provide approximate inverses M_-^{ll} and M_+^{ll} for A_-^{ll} and A_+^{ll}, respectively. Forward and backward point Gauss-Seidel smoothing provides smoothers in the multigrid algorithm resulting in a symmetric M_-^{ll} (cf. [10]). The vector $T_-^h f = x = (x_b, x_l)^t \in Y_h$ is computed via the algorithm:

1. $x_b \leftarrow$ Forward Gauss-Seidel iteration on A_-^{bb} with data f_b
2. $x_l = M_-^{ll}(f_l - A_-^{lb}x_b)$
3. $x_b \leftarrow x_b +$ Backward Gauss-Seidel iteration on A_-^{bb} with data $f_b - A_-^{bl}x_l$.

The vector $T_+^h f$ is computed similarly.

In the case of constant ϵ, using the results of [10], (4.4), (4.17), (4.20) and (4.21), it is not difficult to show that the resulting preconditioned systems, $T_-^h A_-$ and $T_+^h A_+$ are uniformly well conditioned independently of h. This gives rise to an operator T^h which is spectrally equivalent (independently of h) to T_{Y_h}.

The weights in $a_-(\cdot,\cdot)$ and $a_+(\cdot,\cdot)$ are chosen in accordance with Theorem 1. Although a complete analysis of the multigrid algorithm in the case of jumping coefficients is not available, it appears that multigrid results in a uniformly conditioned system independent of the magnitude of the jumps provided that the coarsest grid aligns with the jumps. Our computational results are in agreement with this conjecture. In fact, we observe that preconditioning (5.1) with the $(\epsilon\cdot,\cdot)_r$ mass matrix results in an algorithm whose iterative convergence rate is similar to that for the case when ϵ is identically one. Note that the mass matrix for $(\epsilon\cdot,\cdot)$ is diagonal as there is no interaction between basis functions of the piecewise constant space $\mathbf{X}_h$ corresponding to different triangles.

6. Numerical Experiments

The linear system representing (5.1) is symmetric and positive definite. We solve the system using the preconditioned conjugate gradient method (PCG), with the $(\epsilon\cdot,\cdot)_r$ mass matrix as the preconditioner. The relative tolerance is 10^{-12}. In the first two experiments, reported in Tables 6.1 and 6.2, D is the unit square and the mesh is uniform, with square or triangular elements. Table 6.1 lists the numerical results for the model problem (1.2) with constant coefficient $\epsilon = 1$ and data

$$f_\theta = \pi(r-1)\cos\pi r\cos\pi z$$
$$g = 2\cos\pi r\sin\pi z - \pi(r+1)\sin\pi r\sin\pi z.$$

The exact solution is $(e_r, e_z) = (r\cos\pi r\sin\pi z, \sin\pi r\cos\pi z)$. These results demonstrate the theoretically predicted first order convergence rate. The number of iterations shows only a modest increase with finer meshes, consistent with the theory.

	Square mesh				Triangular mesh	
h	L^2 error	Ratio	PCG	Unknowns	PCG	Unknowns
1/8	0.0723809	1.96296	7	128	8	256
1/16	0.0363771	1.98974	7	512	9	1024
1/32	0.0182127	1.99735	8	2048	10	4096
1/64	0.00910943	1.99933	9	8192	10	16384
1/128	0.0045551	1.99983	9	32768	11	65536
1/256	0.0022776	1.99996	9	131072	11	262144
1/512	0.00113881	1.99999	10	524288	12	1048576

TABLE 6.1. Numerical results for a constant coefficient ($\epsilon = 1$), with the square domain.

	Square mesh		Triangular mesh	
h	PCG Iterations	Unknowns	PCG Iterations	Unknowns
1/8	7	128	9	256
1/16	7	512	9	1024
1/32	8	2048	10	4096
1/64	9	8192	10	16384
1/128	9	32768	11	65536
1/256	9	131072	11	262144
1/512	10	524288	12	1048576

TABLE 6.2. Numerical results for a coefficient with jumps, with the square domain.

Table 6.2 gives the PCG iteration counts when the coefficient has a jump,

$$\epsilon = \begin{cases} 10^4 & \text{if } r, z > 1/2, \\ 1 & \text{otherwise} \end{cases}.$$

The number of iterations observed for the various mesh sizes in this case are essentially identical to those reported in the constant coefficient case (Table 6.1).

To further investigate the iterative convergence behavior of the method, we report three other experiments without tabulating the details of the results. Figure 6 illustrates three nonconvex computational domains used in these experiments. For the problems which we ran, the exact solution is unknown and so we cannot confirm convergence rates. However the computations do demonstrate that the number of PCG iterations does not grow when the domain has reentrant corners and the boundary is not rectangular. Indeed, the number of PCG iterations does not exceed 17 for $1/512 \leq h \leq 1$ and any of the domains.

REFERENCES

[1] Z. Belhachmi, C. Bernardi, and S. Deparis. Weighted Clement operator and application to the finite element discretization of the axisymmetric Stokes problem. *Numer. Math. (to appear).*

[2] C. Bernardi, M. Dauge, and Y. Maday. *Spectral methods for axisymmetric domains*, volume 3 of *Series in Applied Mathematics (Paris).* Gauthier-Villars, Éditions Scientifiques et Médicales Elsevier, Paris, 1999.

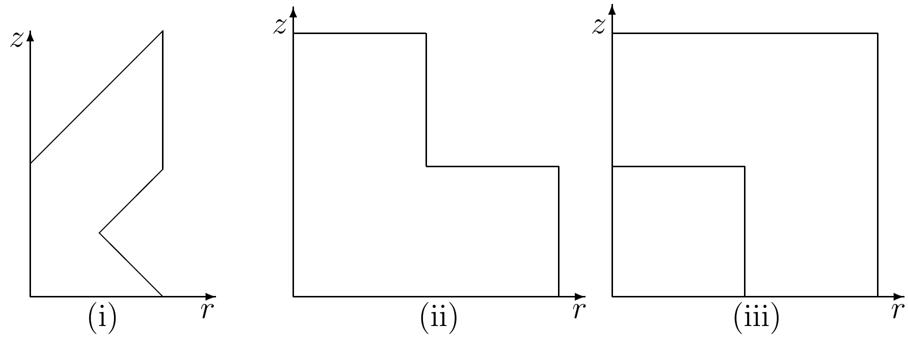

FIGURE 1. Computational domains

[3] S. Börm and R. Hiptmair. Multigrid computation of axisymmetric electromagnetic fields. *Adv. Comput. Math.*, 16(4):331–356, 2002.
[4] A. Bossavit. *Computational electromagnetism*. Electromagnetism. Academic Press Inc., San Diego, CA, 1998. Variational formulations, complementarity, edge elements.
[5] J. H. Bramble and J. E. Pasciak. A new approximation technique for div-curl systems. *Math. Comp.*, 73(248):1739–1762 (electronic), 2004.
[6] J. H. Bramble and J. Xu. Some estimates for a weighted L^2 projection. *Math. Comp.*, 53(194):463–476, 1991.
[7] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.
[8] L. Demkowicz and L. Vardapetyan. Modeling of electromagnetic absorption/scattering problems using *hp*-adaptive finite elements. *Comput. Methods Appl. Mech. Engrg.*, 152(1-2):103–124, 1998. Symposium on Advances in Computational Mechanics, Vol. 5 (Austin, TX, 1997).
[9] V. Girault and P.-A. Raviart. *Finite element approximation of the Navier-Stokes equations*, volume 749 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1979.
[10] J. Gopalakrishnan and J. Pasciak. The convergence of V-cycle multigrid algorithms for axisymmetric Laplace and Maxwell equations. *Preprint*.
[11] P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003.
[12] I. Perugia, D. Schötzau, and P. Monk. Stabilized interior penalty methods for the time-harmonic Maxwell equations. *Comput. Methods Appl. Mech. Engrg.*, 191(41-42):4675–4697, 2002.
[13] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368.

E-mail address: copeland@math.tamu.edu

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368.

E-mail address: pasciak@math.tamu.edu