SIAM J. NUMER. ANAL.
Vol. 43, No. 1, pp. 386–408


# ANALYSIS OF FIRST-ORDER SYSTEM LEAST SQUARES (FOSLS) FOR ELLIPTIC PROBLEMS WITH DISCONTINUOUS COEFFICIENTS: PART I*

MARKUS BERNDT†, THOMAS A. MANTEUFFEL‡, STEPHEN F. MCCORMICK‡, AND GERHARD STARKE§

**Abstract.** First-order system least squares (FOSLS) is a recently developed methodology for solving partial differential equations. Among its advantages are that the finite element spaces are not restricted by the inf-sup condition imposed, for example, on mixed methods and that the least-squares functional itself serves as an appropriate error measure. This paper studies the FOSLS approach for scalar second-order elliptic boundary value problems with discontinuous coefficients, irregular boundaries, and mixed boundary conditions. A least-squares functional is defined, and ellipticity is established in a natural norm of an appropriately scaled least-squares bilinear form. For some geometries, this ellipticity is independent of the size of the jumps in the coefficients. The occurrence of singularities at interface corners, cross points, reentrant corners, and irregular boundary points is discussed, and a basis of singular functions with local support around singular points is established. A companion paper shows that the singular basis functions can be added at little extra cost and lead to optimal performance of standard finite element discretization and multilevel solver techniques, also independent of the size of coefficient jumps for some geometries.

**Key words.** least-squares discretization, second-order elliptic problems, finite elements, multilevel methods

**AMS subject classifications.** 65N55, 65N30, 65F10

**DOI.** 10.1137/S0036142903427688

## 1. Introduction.

The purpose of this paper is to apply first-order system least squares (FOSLS; cf. [11] and [12]) to scalar second-order elliptic boundary value problems in two dimensions with discontinuous coefficients, irregular boundaries, and mixed boundary conditions. Such problems arise in various applications, including flow in heterogeneous porous media [29], neutron transport [1], and biophysics [20]. In many physical applications, one is interested not only in accurate approximation of the physical quantity that satisfies the scalar equation, but also in certain of its derivatives. For example, fluid flow in a porous medium can be modeled by the equation

$$-\nabla \cdot (a\nabla p) = f \tag{1.1}$$

*Received by the editors May 7, 2003; accepted for publication (in revised form) August 31, 2004; published electronically June 14, 2005. This work was performed by an employee of the U.S. Government or under U.S. Government contract. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights.

http://www.siam.org/journals/sinum/43-1/42768.html

†Los Alamos National Laboratory, T-7, Mail Stop B284, Los Alamos, NM 87545 (berndt@lanl.gov). The research of this author was supported by the Department of Energy, under contract W-7405-ENG-36, LA-UR-01-2711.

‡Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526 (tmanteuf@boulder.colorado.edu, stevem@boulder.colorado.edu). The research of these authors was supported by the National Science Foundation under grant number DMS-8704169 and by the Department of Energy, Applied Math Program grant DE-FG03-94ER25217.

§Universität Hannover, Institut für Angewandte Mathematik, Welfengarten 1, 30167 Hannover, Germany (starke@ifam.uni-hannover.de).

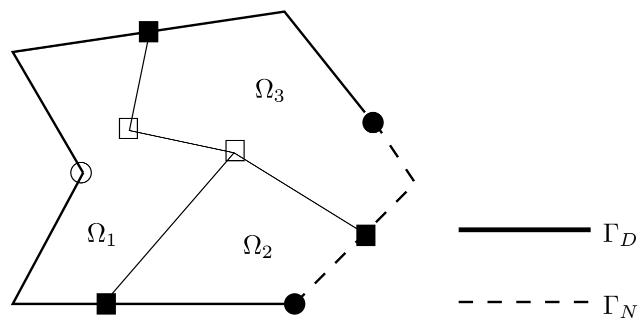


FIG. 1.1. *Polygonal domain $\Omega$ with subdomains $\Omega_i$, $i = 1, 2, 3$, and two cross points.*

for the pressure $p$, where the scalar function $a$ may have large jump discontinuities across interfaces. Of particular interest here is accurate approximation of the flux,

$$\mathbf{u} = a\nabla p. \tag{1.2}$$

For the purposes of discussion, consider problem (1.1) posed on a domain, $\Omega$, composed of a union of polygonal subdomains, $\Omega_i$, in which the coefficient $a$ is constant on each subdomain (see Figure 1.1). In general, the flux, $\mathbf{u}$, will be infinite at certain points, which we will call singular points (see, for example, Strang and Fix [30, Chapter 8]). Singular points can be of several types:

*Cross points:* corner points of the boundary of $\Omega_i$ that lie in the interior of $\Omega$ (□ in Figure 1.1);

*Boundary cross points:* corner points of $\Omega_i$ on the boundary of $\Omega$ that touch another subdomain, $\Omega_j$ (■ in Figure 1.1);

*Reentrant corners:* reentrant corners of $\Omega$ (◯ in Figure 1.1);

*Irregular boundary points:* points on the boundary of $\Omega$ that separate the Dirichlet boundary, $\Gamma_D$, from the Neumann boundary, $\Gamma_N$, for which the interior angle is greater than $\pi/2$ (● in Figure 1.1).

The solution, $p$, can be expressed as the sum of a finite number of singular functions plus a function that is locally smooth, that is, in $H^2(\Omega_i)$ for each $i$. Each singular function is associated with a singular point and, near the singular point, has the form $r^\alpha \Phi(\theta)$, where $(r, \theta)$ are polar coordinates about the singular point and $0 < \alpha < 1$. The character of a singular function depends only on local information near the singular point and is not difficult to compute (see section 5 and [3] for details).

There are many finite element methods for approximating the solution of (1.1). Some yield an approximate solution without specific knowledge of the singular functions, while others use the singular functions either implicitly or explicitly. Below we describe the major approaches.

*Standard Galerkin method.* The standard Galerkin method (cf. Strang and Fix [30]) establishes a weak form and seeks the approximation of $p$ in $H^1(\Omega)$. Convergence deteriorates near the singular points. Early work using $H^1$ singular basis functions can be found in the monograph by Strang and Fix [30, section 8.2]. There, $H^1$ singular basis functions for $p$ were introduced to eliminate the deteriorating finite element approximation near singular points. (See also Cox and Fix [16] and Grisvard [19, section 8.4.2].) A multilevel approach for simultaneously finding the approximate solution and determining the coefficients of the singular basis functions is developed by Brenner and Sung [9]. In [10], Cai and Kim describe a method that is equivalent to a Petrov/Galerkin method in which the singular basis functions are added to the trial space and the dual singular basis functions are added to the test space.

*Mixed methods.* In mixed finite element methods (see, e.g., [8, Chapter 10]), $p$ and $\mathbf{u}$ are usually approximated by different finite element spaces, and, roughly speaking, a Galerkin condition is imposed on the first-order system resulting from (1.1) and (1.2). Normally, the pressure, $p$, is approximated in $L^2$ and the flux, $\mathbf{u}$, is approximated in $H(\text{div})$. Only the integral of the flux is computed along edges of elements, and the pointwise resolution of singularities in the flux is poor.

The least squares methodology for systems of first order is by now several decades old and had its first application in continuum mechanics (see, for example, [21, 31, 22, 26, 15, 23]). Only fairly recently has it produced $H^1$ equivalent forms to which optimal multigrid solvers have been applied (see, for example, [12]). For a thorough review of the least-squares methodology, see [5] and the references therein. The following is an overview of specific least-squares methods and their applicability to the problem at hand.

*Least-squares in* $H(\text{div})$. A similar approach is based on the FOSLS approach developed and analyzed, e.g., in [11, 12, 27, 28]. This methodology replaces the Galerkin condition by the minimization of a least-squares functional associated with a first-order system derived from (1.1) and (1.2). Assuming that $f \in L^2(\Omega)$, the least-squares functional can be defined using the $L^2(\Omega)$-norm. Even in the presence of discontinuities, this translates to ellipticity with respect to the $H^1$-norm for the pressure, $p$, and the $H(\text{div})$-norm in the flux variable, $\mathbf{u}$. This approach, like the mixed method approach, computes only the integral flux and again does not resolve the singularity in the flux variable.

*Weighted least-squares in* $H(\text{div})\cap H(\text{curl})$. Augmenting the basic system with the curl-condition, $\nabla\times(\mathbf{u}/a) = 0$ (see [12, 27]), leads to ellipticity with respect to a scaled version of the $H(\text{div}) \cap H(\text{curl})$ norm in the flux variable. Standard finite element spaces, for example piecewise polynomials with the appropriate jump conditions across interfaces, are not dense in the scaled $H(\text{div}) \cap H(\text{curl})$ norm, and thus convergence cannot be obtained. However, the use of an appropriate weight function near each singular point yields ellipticity in a weighted (and scaled) $H(\text{div})\cap H(\text{curl})$ norm. The piecewise polynomial spaces are dense in this new space. The weighting effectively ignores the singularity while insulating the rest of the region from the presence of the singularity. For the case of reentrant corners, weighted least-squares approaches are presented and analyzed in [17, 16]. Specifically, the method presented in [17] for corner singularities does not rely on the explicit knowledge of the flux singularity at the corner. Its analytic part is computed implicitly. For a weighted least-squares approach in a more general setting, see [25].

*Inverse norm functionals.* Another potentially more general form of the least-squares approach is based on the $H^{-1}(\Omega)$-norm (see [6, 7, 13, 4]). Such schemes based on "inverse" norms can, in principle, be applied when $f \in H^{-1}(\Omega)$, although the theory has so far restricted $f$ to $L^2(\Omega)$. Thus, both the $H^{-1}(\Omega)$ and $L^2(\Omega)$ versions of FOSLS have been developed under the same general assumptions that are usually in force for mixed methods. Standard finite element spaces are dense in $L^2$, and thus convergence is obtained, although only in an $L^2$ sense. This approach uses norms that do not generally take the coefficients of the equation into account and thus have performance that deteriorates for problems with large jumps in the coefficients.

*FOSLL* functionals.* A more recently developed approach, called FOSLL* [14], can be viewed as a least-squares method based on an inverse norm that involves the operator and thus has superior properties in the presence of large jumps in the coefficients. In addition, it handles the more general case, $f \in H^{-1}(\Omega)$.

*Least-squares in* $H(\mathrm{div}) \cap H(\mathrm{curl})$. The current paper is concerned with least-squares functionals using finite element spaces in $H(\mathrm{div})\cap H(\mathrm{curl})$. This paper builds on the theory developed in [2]. Here, and in the companion paper [3], we describe a least-squares approach that includes a curl-condition, $\nabla \times (\mathbf{u}/a) = 0$. While the theory developed in [11] and [12] already allows for discontinuous coefficients, special care must be taken to prove ellipticity, in an appropriate norm, with constants that grow as slowly as possible with respect to the size of the jumps. For this purpose, an appropriate scaling of the least-squares functional that depends on the size of $a$ in different parts of the domain is introduced.

The flux components will, in general, not be in $H^1(\Omega)$, nor will they be in $H^1(\Omega_i)$. Here, we construct singular basis functions for the flux, $\mathbf{u}$, that are in the scaled $H(\mathrm{div}) \cap H(\mathrm{curl})$ but not in $H^1(\Omega_i)$ and have support only near singular points. These are included in our finite element space. As a result, the flux can be computed very accurately near cross points. For standard mixed methods, it would be necessary to make sure that the Ladyzhenskaya–Babuška–Brezzi condition (cf. [8, section 10.5]) is satisfied for the finite element spaces that include the singular function. This is not the case for our first-order system least-squares approach.

In this paper and the companion paper [3], we show that one can add singular basis functions at little additional cost. A singular basis function is composed of a singular function multiplied by a cut-off function that takes the value one in a region around the singularity (the platform) and drops from one to zero in a narrow region around the platform (the fringe). The key is that the singular basis functions satisfy a homogeneous equation of type (1.1) in the platform. Thus, these singular basis functions are orthogonal to any standard basis function that is either supported completely inside the platform or supported completely outside the platform and fringe. Nonzero inner products arise only between singular basis functions and standard basis functions whose support intersects the fringe. As a result, the cost of adding a singular basis function is proportional to the number of grid points in the fringe. In our approach, the fringe has a width of one element, so this additional cost is $O(\sqrt{N})$, where $N$ is the number of grid points.

In this paper, we introduce the problem in section 2; then, in section 3, we construct a scaled FOSLS functional for $p$ and $\mathbf{u}$ and show that this functional is continuous and coercive in a scaled $H^1 \times H(\mathrm{div}) \cap H(\mathrm{curl})$-norm. The coercivity and continuity constants are shown to depend on the coefficient $a$ in a complicated way that involves the geometry of the partition of $\Omega$. We then introduce a *flux-only* functional for $\mathbf{u}$ alone and show that it is continuous and coercive in the scaled version of $H(\mathrm{div}) \cap H(\mathrm{curl})$. In section 4, we introduce the div-curl operator associated with the flux-only functional and discuss its properties. Then, in section 5, we show that the solution, $\mathbf{u}$, can be decomposed as

$$\mathbf{u} = \mathbf{u}_0 + \sum_{m=1}^{M} \sum_{n=1}^{N_m} b_{m,n} \mathbf{s}_{m,n},$$

where $\mathbf{s}_{m,n}$ are a finite number of singular basis functions associated with singular points $\mathbf{x}_m$, $m = 1, \dots, M$, and $\mathbf{u}_0 \in H^1(\Omega_i)$ for every $i$. Thus, $\mathbf{u}_0$ can be approximated by standard finite elements within each domain, provided that they posses the proper jumps across domain interfaces.

In the companion paper [3], we show how to compute approximate singular basis functions, and then we construct a finite element basis using them. We develop error estimates by way of new results for nonconforming spaces in the FOSLS context. We prove that the accuracy of singular basis functions need only be $O(h^p)$, $p > 1/2$. Finally, we develop a multilevel algorithm that includes singular basis functions on all coarser levels and provide numerical results that illustrate its performance.

Our restriction to two-dimensional problems is mainly for the purpose of exposition. However, technical complications arise in higher dimensions. For example, two different types of singularities, associated with edges and with corners or cross points, arise in three dimensions. We do not consider these additional complications in the present paper.

**2. Problem statement and preliminaries.** Consider the following prototype problem on a bounded domain $\Omega \subset \Re^2$:

$$\begin{aligned} -\nabla \cdot (a\nabla p) &= f \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \Gamma_D, \\ \mathbf{n} \cdot a\nabla p &= 0 \quad \text{on } \Gamma_N, \end{aligned} \tag{2.1}$$

where $\mathbf{n}$ denotes the outward unit vector normal to the boundary, $f \in L^2(\Omega)$, and $a(x_1, x_2)$ is a scalar function that is uniformly positive and bounded in $\Omega$ a.e. but may have large jumps across interfaces. Suppose that $\Gamma_D$ has positive measure, so that the Poincaré–Friedrichs inequality

$$\|p\|_{0,\Omega} \le \gamma \|\nabla p\|_{0,\Omega} \tag{2.2}$$

holds for all functions satisfying the boundary conditions in (2.1). Then (2.1) has a unique solution in $H^1(\Omega)$.

Following [12], we rewrite (2.1) as a first-order system by introducing the flux variable, $\mathbf{u} = \sqrt{a}\nabla p$:

$$\begin{aligned} \mathbf{u} - \sqrt{a}\nabla p &= \mathbf{0} \quad \text{in } \Omega, \\ -\nabla \cdot \sqrt{a}\mathbf{u} &= f \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \Gamma_D, \\ \mathbf{n} \cdot \sqrt{a}\mathbf{u} &= 0 \quad \text{on } \Gamma_N. \end{aligned} \tag{2.3}$$

Since $\mathbf{u}/\sqrt{a} = \nabla p$ with $p \in H^1(\Omega)$, we then have (cf. [18, Theorem 2.9])

$$\nabla \times \left(\frac{\mathbf{u}}{\sqrt{a}}\right) := \partial_1 \left(\frac{u_2}{\sqrt{a}}\right) - \partial_2 \left(\frac{u_1}{\sqrt{a}}\right) = 0 \quad \text{in } \Omega\,.$$

(By the term $\partial_k$, we mean $\partial/\partial x_k, k = 1, 2$.) Moreover, the homogeneous Dirichlet boundary condition on $\Gamma_D$ implies the tangential flux condition

$$\boldsymbol{\tau} \cdot \left(\frac{\mathbf{u}}{\sqrt{a}}\right) := \frac{n_1 u_2 - n_2 u_1}{\sqrt{a}} = 0 \quad \text{on } \Gamma_D.$$

(Here, $\boldsymbol{\tau}$ is the counterclockwise unit tangent vector.)

Adding these equations to first-order system (2.3) yields the augmented, but consistent, system

$$
\begin{aligned}
\mathbf{u}-\sqrt{a}\nabla p &= \mathbf{0} && \text{in } \Omega,\\
-\nabla\cdot\sqrt{a}\mathbf{u} &= f && \text{in } \Omega,\\
\nabla\times\left(\frac{\mathbf{u}}{\sqrt{a}}\right) &= 0 && \text{in } \Omega,\\
p &= 0 && \text{on } \Gamma_D,\\
\mathbf{n}\cdot\sqrt{a}\mathbf{u} &= 0 && \text{on } \Gamma_N,\\
\boldsymbol{\tau}\cdot\left(\frac{\mathbf{u}}{\sqrt{a}}\right) &= 0 && \text{on } \Gamma_D.
\end{aligned}
\tag{2.4}
$$

Problems (2.1) and (2.4) are equivalent in that their unique solutions are in correspondence ($p$ solves (2.1) if and only if $p$ and $\mathbf{u}=\sqrt{a}\nabla p$ solve (2.4)). If $\Gamma_N$ is not connected, then we add the constraint

$$
\int_{\Gamma_{N_i}}\boldsymbol{\tau}\cdot\left(\frac{\mathbf{u}}{\sqrt{a}}\right)=0 \tag{2.5}
$$

for every disjoint piece, $\Gamma_{N_i}$, of $\Gamma_N$. This constraint is necessary to ensure that the *flux-only* functional described below (see (3.17)) has a unique solution.

For both scalar and vector quantities, denote the standard Sobolev spaces as $L^2(\Omega)$ and $H^k(\Omega)$, with respective norms $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{k,\Omega}$. We also define the spaces

$$
\begin{aligned}
H(\operatorname{div} a;\Omega) &:= \left\{\mathbf{v}\in L^2(\Omega)^2 : \nabla\cdot\sqrt{a}\mathbf{v}\in L^2(\Omega)\right\},\\
H(\operatorname{curl} a;\Omega) &:= \left\{\mathbf{v}\in L^2(\Omega)^2 : \nabla\times\left(\frac{\mathbf{v}}{\sqrt{a}}\right)\in L^2(\Omega)\right\},\\
V &:= \left\{q\in H^1(\Omega) : q=0 \text{ on } \Gamma_D\right\},\\
\mathbf{W} &:= \Bigg\{\mathbf{v}\in H(\operatorname{div} a;\Omega)\cap H(\operatorname{curl} a;\Omega) : \mathbf{n}\cdot\sqrt{a}\mathbf{v}=0 \text{ on } \Gamma_N,\\
&\qquad \boldsymbol{\tau}\cdot\left(\frac{\mathbf{v}}{\sqrt{a}}\right)=0 \text{ on } \Gamma_D, \int_{\Gamma_{N_i}}\boldsymbol{\tau}\cdot\left(\frac{\mathbf{u}}{\sqrt{a}}\right)=0\Bigg\}.
\end{aligned}
$$

Denote the respective seminorm and norm on $\mathbf{W}$ by

$$
\begin{aligned}
|\mathbf{v}|_{\mathbf{W}}^2 &:= \left\|\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{v}\right\|_{0,\Omega}^2+\left\|\sqrt{a}\nabla\times\frac{1}{\sqrt{a}}\mathbf{v}\right\|_{0,\Omega}^2,\\
\|\mathbf{v}\|_{\mathbf{W}}^2 &:= |\mathbf{v}|_{\mathbf{W}}^2+\|\mathbf{v}\|_{0,\Omega}^2.
\end{aligned}
\tag{2.6}
$$

We show in Lemma 3.3 below that this seminorm is in fact a norm on $\mathbf{W}$ by establishing a Poincaré–Friedrichs-type inequality.

Note that $\mathbf{v}\in\mathbf{W}$ is characterized by the fact that, across any curve in $\Omega$ with normal $\mathbf{n}$ and tangent $\boldsymbol{\tau}$, both $\mathbf{n}\cdot\sqrt{a}\mathbf{v}$ and $\boldsymbol{\tau}\cdot\frac{1}{\sqrt{a}}\mathbf{v}$ are continuous (a.e.). (For the first condition see, for example, [32, Chapter 6.2]. The second condition can be derived analogously.) We refer to the continuity of these two terms at lines of discontinuity of $a$ as *interface conditions* for $\mathbf{u}\in\mathbf{W}$. Clearly, for the solution of (2.1), we have $p\in V$ and $\mathbf{u}\in\mathbf{W}$, so it is appropriate to pose (2.4) on these spaces.

As mentioned above, our main interest is in the solution of (2.1) when $a(x_1, x_2)$ has large jumps. For this purpose, we assume that

$$\overline{\Omega} = \bigcup_{i=1}^{J} \overline{\Omega}_i, \tag{2.7}$$

where $\Omega_i$ are mutually disjoint, open, simply connected, polygonal regions (see Figure 1.1). Assume also that the restriction of $a(x_1, x_2)$ to $\Omega_i$ is in $C^{1,1}(\Omega_i)$ and that

$$c_1\omega_i \le a(x_1, x_2) \le c_2\omega_i \quad \text{for all } (x_1, x_2) \in \Omega_i, \tag{2.8}$$

with order one constants $c_1, c_2$ and arbitrary positive constants $\omega_i$. In other words, $a(x_1, x_2)$ is assumed to be of approximate size $\omega_i$ throughout $\Omega_i$ for each $i$, but $\omega_i$ is allowed to have large variations over $i$. In the bounds derived below, we separate the dependence on the variation in $\{\omega_i\}$ from the variation within each $\Omega_i$, that is, on $c_1$, $c_2$, and

$$c_3 := \max_{1\le i\le J} \|\nabla a\|_{0,\Omega_i} < \infty. \tag{2.9}$$

Given this decomposition of $\Omega$, define the *split* seminorms and norms, respectively, as follows:

$$|\mathbf{v}|^2_{k,S} := \sum_{i=1}^{J} |\mathbf{v}|^2_{k,\Omega_i} \tag{2.10}$$

and

$$\|\mathbf{v}\|^2_{k,S} := \|\mathbf{v}\|^2_{0,\Omega} + \sum_{j=1}^{k} |\mathbf{v}|^2_{j,S}. \tag{2.11}$$

Let $H^k_S(\Omega)$ denote the closure of $C^\infty(\overline{\Omega})$ in the split norm, and define

$$\mathbf{W}^1_S := H^1_S(\Omega) \cap \mathbf{W}. \tag{2.12}$$

We now show that if $a$ is piecewise constant ($c_1 = c_2$ in (2.8)) with respect to the decomposition, then

$$\|\mathbf{v}\|_{1,S} = \|\mathbf{v}\|_{\mathbf{W}} \quad \text{for every } \ \mathbf{v} \in H^1_S(\Omega). \tag{2.13}$$

We first need to establish two lemmas. For the first lemma, consider one polygonal, simply connected subdomain, $\Omega_i$, of $\Omega$, with vertices labeled $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K$ in counterclockwise order. Letting $\mathbf{x}_{K+1} = \mathbf{x}_1$, denote by $\Gamma_j$ the side connecting $\mathbf{x}_j$ and $\mathbf{x}_{j+1}$. If $\Gamma_j$ makes angle $\theta_j$ with the positive $x_1$-axis, then $\mathbf{n}_j = (\sin(\theta_j), -\cos(\theta_j))^t$ and $\boldsymbol{\tau}_j = (\cos(\theta_j), \sin(\theta_j))^t$ are the outward unit normal and counterclockwise unit tangent to $\Gamma_j$, respectively.

LEMMA 2.1. *Assume that $\Omega_i$ is a polygonal domain and that $\mathbf{u} = (u_1, u_2)^t \in (H^2(\Omega_i))^2$; then*

$$\begin{aligned}\int\!\!\int_{\Omega_i} \partial_1 u_1 \partial_2 u_2 dz \;&= \int\!\!\int_{\Omega_i} \partial_2 u_1 \partial_1 u_2 dz - \int_{\partial\Omega_i} (\boldsymbol{\tau}\cdot\mathbf{u}) d(\mathbf{n}\cdot\mathbf{u}) \\ &\quad + \frac{1}{2}\sum_{j=1}^{K} \left( (\boldsymbol{\tau}_j\cdot\mathbf{u})(\mathbf{n}_j\cdot\mathbf{u})|_{\mathbf{x}_j} - (\boldsymbol{\tau}_{j-1}\cdot\mathbf{u})(\mathbf{n}_{j-1}\cdot\mathbf{u})|_{\mathbf{x}_j}\right).\end{aligned} \tag{2.14}$$

*Proof.* First, assume that $\Omega$ is simply connected. For $\mathbf{u} \in H^2(\Omega_i)$, Green's identity yields

$$\int\int_{\Omega_i} \partial_1 u_1 \partial_2 u_2 dz = \int\int_{\Omega_i} \partial_2 u_1 \partial_1 u_2 dz + \int_{\partial\Omega_i} u_1 du_2.$$

The definition of $\mathbf{n}_i$ and $\boldsymbol{\tau}_i$ and a bit of algebra yield

$$\int_{\Gamma_j} (\boldsymbol{\tau}_j \cdot \mathbf{u}) d(\mathbf{n}_j \cdot \mathbf{u}) = \frac{1}{2}\,(\boldsymbol{\tau}_j \cdot \mathbf{u})(\mathbf{n}_j \cdot \mathbf{u})|_{\mathbf{x}_j}^{\mathbf{x}_{j+1}} + \frac{1}{2}\,u_1 u_2|_{\mathbf{x}_j}^{\mathbf{x}_{j+1}} - \int_{\Gamma_j} u_1 du_2.$$

Summing over the edges yields the result. The result for a general connected polygonal domain is established by cutting $\Omega_i$ into simply connected polygonal subdomains and adding the result. ∎

LEMMA 2.2. *For every* $\mathbf{u} \in \mathbf{W}_S^1$, *we have*

$$\int\int_{\Omega} \partial_1 u_1 \partial_2 u_2 dz = \int\int_{\Omega} \partial_2 u_1 \partial_1 u_2 dz. \tag{2.15}$$

*Proof.* First, let $\mathbf{u} \in H_S^2(\Omega) \cap \mathbf{W}$. The space $\mathbf{W}$ is characterized by the property that, for $\mathbf{u} \in \mathbf{W}$, both $\sqrt{a}\mathbf{n} \cdot \mathbf{u}$ and $\frac{1}{\sqrt{a}}\boldsymbol{\tau} \cdot \mathbf{u}$ are continuous (a.e.) across any curve in $\Omega$. Thus, $(\mathbf{n} \cdot \mathbf{u})(\boldsymbol{\tau} \cdot \mathbf{u})$ is continuous (a.e.). In particular, this holds for the polygonal boundaries between the regions $\Omega_i$. Let $\Gamma_{ij}$ denote the edge joining $\Omega_i$ and $\Omega_j$. Summing the boundary integrals in (2.14) over each $\Omega_i$ shows that $\Gamma_{ij}$ is traversed once in each direction. Thus, only integrals on the boundary of $\Omega$ survive. This yields

$$\int\int_{\Omega} \partial_1 u_1 \partial_2 u_2 = \int\int_{\Omega} \partial_2 u_1 \partial_1 u_2 \tag{2.16}$$

$$+ \frac{1}{2}\sum_{j=1}^{\tilde{K}} \left((\tilde{\boldsymbol{\tau}}_j \cdot \mathbf{u})(\tilde{\mathbf{n}}_j \cdot \mathbf{u}) - (\tilde{\boldsymbol{\tau}}_{j-1} \cdot \mathbf{u})(\tilde{\mathbf{n}}_{j-1} \cdot \mathbf{u})\right)|_{\tilde{\mathbf{x}}_j}, \tag{2.17}$$

where the $\tilde{\mathbf{x}}_j$ now denote the $\tilde{K}$ vertices $\tilde{\mathbf{x}}_j$ on the boundary of $\Omega$, and the $\tilde{\mathbf{n}}_j$ and $\tilde{\boldsymbol{\tau}}_j$ are the corresponding standard normal and tangent vectors. The boundary conditions imposed on $\mathbf{W}$ now imply (2.15) for $\mathbf{u} \in H_S^2(\Omega) \cap \mathbf{W}$. The proof is completed by noting that Lemma 4.3.1.3 in [19] implies that $H_S^2(\Omega) \cap \mathbf{W}$ is dense in $\mathbf{W}_S^1 = H_S^1(\Omega) \cap \mathbf{W}$. ∎

The next result has important implications for the decomposition of $\mathbf{W}$.

THEOREM 2.3. *Suppose* $a = \omega_i$ *(constant) on* $\Omega_i$. *Then*

$$|\mathbf{u}|_{1,S} = |\mathbf{u}|_{\mathbf{W}} \quad \textit{for every} \ \ \mathbf{u} \in \mathbf{W}_S^1. \tag{2.18}$$

*Proof.* By definition,

$$\begin{aligned} |\mathbf{u}|_{\mathbf{W}}^2 &= \left\|\frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u}\right\|_{0,\Omega}^2 + \left\|\sqrt{a}\nabla \times \frac{1}{\sqrt{a}}\mathbf{u}\right\|_{0,\Omega}^2 \\ &= \sum_{i=1}^{J}\left(\left\|\frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u}\right\|_{0,\Omega_i}^2 + \left\|\sqrt{a}\nabla \times \frac{1}{\sqrt{a}}\mathbf{u}\right\|_{0,\Omega_i}^2\right) \\ &= \sum_{i=1}^{J}(\|\nabla \cdot \mathbf{u}\|_{0,\Omega_i}^2 + \|\nabla \times \mathbf{u}\|_{0,\Omega_i}^2). \end{aligned}$$

The theorem now follows from Lemma 2.2 and the easily verified relation

$$\|\nabla\cdot\mathbf{u}\|_{0,\Omega_i}^2+\|\nabla\times\mathbf{u}\|_{0,\Omega_i}^2=|\mathbf{u}|_{1,\Omega_i}+2\langle\partial_1u_1,\,\partial_2u_2\rangle_{0,\Omega_i}-2\langle\partial_2u_1,\,\partial_1u_2\rangle_{0,\Omega_i}. \qquad \square$$

COROLLARY 2.4. *Suppose that $a(x, y)$ is now allowed to vary according to* (2.8) *and* (2.9). *Then,*

$$\frac{1}{\delta}\|\mathbf{u}\|_{\mathbf{W}}\le\|\mathbf{u}\|_{1,S}\le\delta\|\mathbf{u}\|_{\mathbf{W}}\quad \textit{for}\quad \mathbf{u}\in\mathbf{W}_S^1,$$

*where*

$$\delta=\sqrt{1+c_3\left(\frac{c_3+\sqrt{c_3^2+8}}{4}\right)}$$

*and $c_3$ is defined in* (2.9).

*Proof.* Observe that

$$\left\|\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}\right\|_{0,\Omega_i}\le\|\nabla\cdot\mathbf{u}\|_{0,\Omega_i}+\left\|\frac{1}{2}(\nabla a)\cdot\mathbf{u}\right\|_{0,\Omega_i},$$
$$\left\|\sqrt{a}\nabla\times\frac{1}{\sqrt{a}}\mathbf{u}\right\|_{0,\Omega_i}\le\|\nabla\times\mathbf{u}\|_{0,\Omega_i}+\left\|\frac{1}{2}(\nabla^\perp a)\cdot\mathbf{u}\right\|_{0,\Omega_i}.$$

(Here, we use the notation $\nabla^\perp a := (-\partial_2 a,\,\partial_1 a)^t$.) Using the $\epsilon$-inequality twice now yields

$$\begin{aligned}|\mathbf{u}|_{\mathbf{W},\Omega_i}^2&\le\left(\|\nabla\cdot\mathbf{u}\|_{0,\Omega_i}+\frac{c_3}{2}\|\mathbf{u}\|_{0,\Omega_i}\right)^2+\left(\|\nabla\times\mathbf{u}\|_{0,\Omega_i}+\frac{c_3}{2}\|\mathbf{u}\|_{0,\Omega_i}\right)^2\\&\le(1+\epsilon)(\|\nabla\cdot\mathbf{u}\|_{0,\Omega_i}^2+\|\nabla\times\mathbf{u}\|_{0,\Omega_i}^2)+\left(1+\frac{1}{\epsilon}\right)\frac{c_3^2}{2}\|\mathbf{u}\|_{0,\Omega_i}^2\end{aligned}$$

for any $\epsilon>0$. Choosing $\epsilon=c_3\big(\frac{c_3+\sqrt{c_3^2+8}}{4}\big)$, summing over $i$, and appealing to Theorem 2.3 yields the lower bound. The upper bound is proved in a similar fashion. $\square$

*Remark* 1. Following the development in section 4.3 in [19], the above results can be extended to problem (2.1) with boundary conditions that involve both the conormal and tangential derivatives, as long as the coefficients remain constant on each edge. We believe that Theorem 2.3 also holds for regions $\Omega$ for which $\partial\Omega_i$ are piecewise $C^{1,1}$, but this remains an open question.

**3. The least-squares functional.** We now turn to the construction of the least-squares functional. An appropriate scaling of the equations in (2.4) leads to

(3.1)
$$G_\alpha(\mathbf{u},p;f):=\alpha\|\mathbf{u}-\sqrt{a}\nabla p\|_{0,\Omega}^2+\left\|\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}+\frac{1}{\sqrt{a}}f\right\|_{0,\Omega}^2+\left\|\sqrt{a}\,\nabla\times\frac{1}{\sqrt{a}}\mathbf{u}\right\|_{0,\Omega}^2$$

and associated bilinear form

(3.2)
$$\begin{aligned}\mathcal{F}_\alpha((\mathbf{u},p);(\mathbf{v},q))&=\alpha\langle\mathbf{u}-\sqrt{a}\nabla p,\mathbf{v}-\sqrt{a}\nabla q\rangle_{0,\Omega}\\&+\left\langle\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u},\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{v}\right\rangle_{0,\Omega}+\left\langle\sqrt{a}\,\nabla\times\frac{1}{\sqrt{a}}\mathbf{u},\sqrt{a}\,\nabla\times\frac{1}{\sqrt{a}}\mathbf{v}\right\rangle_{0,\Omega},\end{aligned}$$

where $\alpha \geq 0$ will be determined later. Here, for the sake of notational simplicity, we agree that $\langle \cdot, \cdot \rangle_{0,\Omega}$ is meant componentwise for vector functions, e.g., if $\mathbf{w} = (w_1, w_2)$ and $\mathbf{z} = (z_1, z_2)$, then

$$\langle \mathbf{w}, \mathbf{z} \rangle_{0,\Omega} = \langle w_1, z_1 \rangle_{0,\Omega} + \langle w_2, z_2 \rangle_{0,\Omega} .$$

The solution of (2.4) also solves the minimization problem

$$G_\alpha(\mathbf{u}, p; f) = \min_{(\mathbf{v},q)\in \mathbf{W}\times V} G_\alpha(\mathbf{v}, q; f) \tag{3.3}$$

and, therefore, the variational problem

$$\mathcal{F}_\alpha((\mathbf{u}, p); (\mathbf{v}, q)) = -\left\langle \frac{1}{\sqrt{a}} f, \frac{1}{\sqrt{a}} \nabla \cdot \sqrt{a}\mathbf{v} \right\rangle_{0,\Omega} \quad \text{for all } (\mathbf{v}, q) \in \mathbf{W} \times V . \tag{3.4}$$

In Theorem 3.2, we will show that $(\mathcal{F}_\alpha((\mathbf{v}, q); (\mathbf{v}, q)))^{1/2}$ is uniformly equivalent to the scaled norm defined for $(\mathbf{v}, q) \in \mathbf{W} \times V$ by

$$\begin{aligned} &|||(\mathbf{v}, q)|||_\alpha \\ &:= \left( \left\| \frac{1}{\sqrt{a}} \nabla \cdot \sqrt{a}\mathbf{v} \right\|_{0,\Omega}^2 + \left\| \sqrt{a} \nabla \times \frac{1}{\sqrt{a}} \mathbf{v} \right\|_{0,\Omega}^2 + \alpha \|\mathbf{v}\|_{0,\Omega}^2 + \alpha \|\sqrt{a}\nabla q\|_{0,\Omega}^2 \right)^{1/2} . \end{aligned} \tag{3.5}$$

Note that, for sufficiently smooth $a$, we get

$$|||(\mathbf{v}, q)|||_\alpha \sim \left( \|\nabla \cdot \mathbf{v}\|_{0,\Omega}^2 + \|\nabla \times \mathbf{v}\|_{0,\Omega}^2 + \alpha \|\mathbf{v}\|_{0,\Omega}^2 + \alpha \|\sqrt{a}\nabla q\|_{0,\Omega}^2 \right)^{1/2} , \tag{3.6}$$

although our assumptions on $a$ do not admit this equivalence in general.

Before we prove the main result, we must establish a scaled Poincaré–Friedrichs inequality. By assumption, $\Gamma_D$ in (2.1) is a set of positive measures on $\partial\Omega$. Thus, a standard proof can be used to establish

$$\|p\|_{0,\Omega} \leq \gamma_0 \|\nabla p\|_{0,\Omega}, \tag{3.7}$$

for $p \in V$, where $\gamma_0$ depends only on $\Omega$. In fact, we may choose $\gamma_0$ so that (3.7) holds on any subdomain composed of a union of the $\Omega_i$ whose closure is connected and intersects $\Gamma_D$ in a set of positive measure. In this sense, $\gamma_0$ depends also on the partitioning (2.7).

Instead of (3.7), we seek scaled inequalities of the form

$$\|\sqrt{a}p\|_{0,\Omega} \leq c_4 \gamma_0 \|\sqrt{a}\nabla p\|_{0,\Omega} \quad \text{and} \quad \left\| \frac{1}{\sqrt{a}} p \right\|_{0,\Omega} \leq c_5 \gamma_0 \left\| \frac{1}{\sqrt{a}} \nabla^\perp p \right\|_{0,\Omega} ,$$

for $p \in V$. Of course, if each subdomain is such that $\Gamma_D \cap \overline{\Omega}_i$ is of positive measure, then we may choose, for example, $c_4 = \sqrt{c_2/c_1}$ (see (2.8)). In general, $c_4$ and $c_5$ depend on $a(x_1, x_2)$ in a more complicated way that we now characterize.

For each $\Omega_i$, there is a connected path $\lambda_i$ in $\Omega$ from $\Gamma_D$ to $\Omega_i$ that passes through, say, $\overline{\Omega}_{j_1}, \overline{\Omega}_{j_2}, \ldots, \overline{\Omega}_{j_k} = \overline{\Omega}_i$ $(k \leq J)$ in turn, where $\Gamma_D \cap \overline{\Omega}_{j_1}$ and $\overline{\Omega}_{j_\ell} \cap \overline{\Omega}_{j_{\ell-1}}, \ell = 2, \ldots, k$, all have positive measure. We call such a path admissible. Now, let $c_1$, $c_2$, and $\omega_i$ be as in (2.8) and define

$$C_i = \min_{\lambda_i} \max_{\ell=1,\ldots,k} \frac{\omega_i}{\omega_{j_\ell}}, \quad D_i = \min_{\lambda_i} \max_{\ell=1,\ldots,k} \frac{\omega_{j_\ell}}{\omega_i}, \tag{3.8}$$

and

$$c_4 = \sqrt{\frac{c_2}{c_1}} \max_{i=1,\dots,J} \sqrt{C_i}, \quad c_5 = \sqrt{\frac{c_2}{c_1}} \max_{i=1,\dots,J} \sqrt{D_i}. \tag{3.9}$$

Note that, for certain geometries, $c_4$ or $c_5$ might depend on the maximum global variation in $a(x_1, x_2)$. However, for other geometries, $c_4$ or $c_5$ may be small even for arbitrary large global $a$-variations. We refer to this property by saying that $c_4$ and $c_5$ are $P$-uniform, meaning that $c_4$ and $c_5$ depend on $a$-variations along the best path to $\Gamma_D$, but are otherwise independent of the jumps in $a$.

LEMMA 3.1. *There exists a $P$-uniform constant, $\gamma \in (0, \sqrt{J}\gamma_0]$, such that*

$$\|\sqrt{a}p\|_{0,\Omega} \le c_4\gamma \|\sqrt{a}\nabla p\|_{0,\Omega} \quad \text{for all} \ \ p \in V, \tag{3.10}$$

$$\left\| \frac{1}{\sqrt{a}} p \right\|_{0,\Omega} \le c_5\gamma \left\| \frac{1}{\sqrt{a}} \nabla^{\perp} p \right\|_{0,\Omega} \quad \text{for all} \ \ p \in V, \tag{3.11}$$

*where $c_4$ and $c_5$ are the $P$-uniform constants defined in* (3.9).

*Proof.* Choose $\Omega_i$ and any of its admissible paths. By (3.7), we have

$$\sum_{\ell=1}^{k} \|p\|^2_{0,\Omega_{j_\ell}} \le \gamma_0^2 \sum_{\ell=1}^{k} \|\nabla p\|^2_{0,\Omega_{j_\ell}}.$$

In particular,

$$\|p\|^2_{0,\Omega_i} \le \gamma_0^2 \sum_{\ell=1}^{k} \|\nabla p\|^2_{0,\Omega_{j_\ell}}.$$

From (2.8), we have

$$\begin{aligned} \|\sqrt{a}p\|^2_{0,\Omega_i} &\le c_2\omega_i \|p\|^2_{0,\Omega_i} \le c_2\omega_i\gamma_0^2 \sum_{\ell=1}^{k} \|\nabla p\|^2_{0,\Omega_{j_\ell}} \\ &= c_2\gamma_0^2 \sum_{\ell=1}^{k} \frac{\omega_i}{\omega_{j_\ell}} \omega_{j_\ell} \|\nabla p\|^2_{0,\Omega_{j_\ell}} \le \frac{c_2}{c_1}\gamma_0^2 C_i \sum_{\ell=1}^{k} \|\sqrt{a}\nabla p\|^2_{0,\Omega_{j_\ell}}. \end{aligned}$$

Summation over $i$ now yields (3.10) with $\gamma \le \sqrt{J}\gamma_0$. The proof of (3.11) is analogous. ∎

THEOREM 3.2. *If we choose $\alpha \le 1/c_4^2$, where $c_4$ is defined in* (3.9)*, then there exist $P$-uniform constants $\gamma_1$ and $\gamma_2$ such that*

$$\mathcal{F}_\alpha((\mathbf{u}, p); (\mathbf{u}, p)) \ge \gamma_1 |||(\mathbf{u}, p)|||^2_\alpha \quad \text{for all } (\mathbf{u}, p) \in \mathbf{W} \times V, \tag{3.12}$$

*and*

$$\mathcal{F}_\alpha((\mathbf{u}, p); (\mathbf{v}, q)) \le \gamma_2 |||(\mathbf{u}, p)|||_\alpha \, |||(\mathbf{v}, q)|||_\alpha \quad \text{for all } (\mathbf{u}, p)\,,\ (\mathbf{v}, q) \in \mathbf{W} \times V. \tag{3.13}$$

*Proof.* The proof is similar to the proof of [11, Theorem 3.1] (see also [27, Theorems 2.1 and 2.2]). We include it here because we must confirm that the constants $\gamma_1$ and $\gamma_2$ are $P$-uniform. The main part of the proof consists of showing that the functionals

$$\hat{\mathcal{F}}_\alpha((\mathbf{u}, p); (\mathbf{v}, q)) := \alpha\langle \mathbf{u} - \sqrt{a}\nabla p, \mathbf{v} - \sqrt{a}\nabla q\rangle_{0,\Omega} + \left\langle \frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}, \frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{v} \right\rangle_{0,\Omega}$$

and

$$\hat{\mathcal{S}}_\alpha(\mathbf{u},p;\mathbf{v},q) := \alpha\langle \mathbf{u},\mathbf{v}\rangle_{0,\Omega} + \alpha\langle\sqrt{a}\nabla p,\sqrt{a}\nabla q\rangle_{0,\Omega} + \left\langle \frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}, \frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{v}\right\rangle_{0,\Omega}$$

satisfy

$$C_1\hat{\mathcal{S}}_\alpha(\mathbf{u},p;\mathbf{u},p) \le \hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p)) \tag{3.14}$$

and

$$\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{v},q)) \le C_2(\hat{\mathcal{S}}_\alpha(\mathbf{u},p;\mathbf{u},p))^{1/2}(\hat{\mathcal{S}}_\alpha(\mathbf{v},q;\mathbf{v},q))^{1/2}, \tag{3.15}$$

with constants $C_1$ and $C_2$ that are $P$-uniform. Since on $\partial\Omega$ we either have $p=0$ or $\mathbf{n}\cdot\sqrt{a}\mathbf{u}=0$, then integration by parts confirms that

$$\langle\mathbf{u},\sqrt{a}\nabla p\rangle_{0,\Omega} + \langle\nabla\cdot\sqrt{a}\mathbf{u},p\rangle_{0,\Omega} = 0.$$

For any $\beta>0$, which we specify later, we have

$$\begin{aligned}
\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p)) &= \alpha\langle\mathbf{u},\mathbf{u}\rangle_{0,\Omega} + \alpha\langle\sqrt{a}\nabla p,\sqrt{a}\nabla p\rangle_{0,\Omega} - 2\alpha\langle\mathbf{u},\sqrt{a}\nabla p\rangle_{0,\Omega}\\
&\quad + \left\langle\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u},\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}\right\rangle_{0,\Omega} + 2\alpha\beta\langle\nabla\cdot\sqrt{a}\mathbf{u},p\rangle_{0,\Omega}\\
&\quad + 2\alpha\beta\langle\mathbf{u},\sqrt{a}\nabla p\rangle_{0,\Omega} + \alpha^2\beta^2\langle\sqrt{a}p,\sqrt{a}p\rangle_{0,\Omega} - \alpha^2\beta^2\langle\sqrt{a}p,\sqrt{a}p\rangle_{0,\Omega}\\
&= \alpha\langle\mathbf{u}+(\beta-1)\sqrt{a}\nabla p,\mathbf{u}+(\beta-1)\sqrt{a}\nabla p\rangle_{0,\Omega}\\
&\quad + \left\langle\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}+\alpha\beta\sqrt{a}p,\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}+\alpha\beta\sqrt{a}p\right\rangle_{0,\Omega}\\
&\quad + \alpha(2\beta-\beta^2)\langle\sqrt{a}\nabla p,\sqrt{a}\nabla p\rangle_{0,\Omega} - \alpha^2\beta^2\langle\sqrt{a}p,\sqrt{a}p\rangle_{0,\Omega}\\
&\ge \alpha(2\beta-\beta^2)\langle\sqrt{a}\nabla p,\sqrt{a}\nabla p\rangle_{0,\Omega} - \alpha^2\beta^2\langle\sqrt{a}p,\sqrt{a}p\rangle_{0,\Omega}\\
&\ge \alpha(2\beta-(1+\gamma^2)\beta^2)\|\sqrt{a}\nabla p\|_{0,\Omega}^2\,,
\end{aligned}$$

where we used the assumption that $\alpha\le 1/c_4^2$ and where $\gamma$ is from Lemma 3.1. Choosing $\beta = 1/(1+\gamma^2)$ leads to

$$\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p)) \ge \beta\alpha\|\sqrt{a}\nabla p\|_{0,\Omega}^2\,.$$

We then also have

$$\alpha\|\mathbf{u}\|_{0,\Omega}^2 \le 2\alpha(\|\mathbf{u}-\sqrt{a}\nabla p\|_{0,\Omega}^2 + \|\sqrt{a}\nabla p\|_{0,\Omega}^2) \le 2\left(1+\frac{1}{\beta}\right)\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p))$$

and, clearly,

$$\left\|\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}\right\|_{0,\Omega}^2 \le \hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p)),$$

which completes the proof of (3.14). ∎

Upper bound (3.15) follows from

$$\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{v},q)) \le 2(\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p)))^{1/2}(\hat{\mathcal{F}}_\alpha((\mathbf{v},q);(\mathbf{v},q)))^{1/2}$$

and

$$\begin{aligned}\hat{\mathcal{F}}_\alpha((\mathbf{u},p);(\mathbf{u},p)) &= \alpha\|\mathbf{u}-\sqrt{a}\nabla p\|^2_{0,\Omega} + \left\|\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}\right\|^2_{0,\Omega} \\ &\le 2\left(\alpha\|\mathbf{u}\|^2_{0,\Omega} + \alpha\left\|\sqrt{a}\nabla p\right\|^2_{0,\Omega} + \left\|\frac{1}{\sqrt{a}}\nabla\cdot\sqrt{a}\mathbf{u}\right\|^2_{0,\Omega}\right) \\ &= 2\hat{\mathcal{S}}_\alpha(\mathbf{u},p;\mathbf{u},p).\end{aligned} \tag{3.16}$$

The proof of Theorem 3.2 is completed by adding the term $\|\sqrt{a}\nabla\times(\mathbf{u}/\sqrt{a})\|^2_{0,\Omega}$ to both sides of inequalities (3.14) and (3.16). ∎

Theorem 3.2 establishes coercivity and continuity of the least-squares bilinear form $\mathcal{F}_\alpha((\cdot,\cdot);(\cdot,\cdot))$ in terms of the norm $|||(\cdot,\cdot)|||_\alpha$. This norm equivalence depends on the jumps in $a$ along the best path to the Dirichlet boundary, but is otherwise independent of the jumps in $a$.

The scaling of the norm $|||(\cdot,\cdot)|||_\alpha$ has the following physical interpretation. Focusing first on $p$, imagine that the error $q$ as measured by the term $\|\sqrt{a}\nabla q\|^2_{0,\Omega}$ is balanced over the domain; that is, $\sqrt{a}\nabla q$ is roughly constant. Then, in areas where $\sqrt{a}$ is relatively small, $\nabla q$ is correspondingly relatively large, and one has to expect a less accurate approximation (in the $L^2$ sense) there compared to areas where $\sqrt{a}$ is large and $\nabla q$ is therefore small. In contrast, approximation of the velocity $\mathbf{u}=\sqrt{a}\nabla p$ (assuming the error $\mathbf{v}$ is balanced in the sense of the term $|\mathbf{v}|^2_{1,\Omega}+\alpha\|\mathbf{v}\|^2_{0,\Omega}$; see (3.6)) can be expected to have balanced accuracy (in the $L^2$ sense) over $\Omega$. Ellipticity with constants that are independent of the global jumps in $a$ asserts that the scaling in $\mathcal{F}_\alpha((\cdot,\cdot);(\cdot,\cdot))$ correctly reflects these attributes.

Uniform coercivity and continuity of $\mathcal{F}$ in the norm $|||(\cdot,\cdot)|||_\alpha$ allows for effective computation of $\mathbf{u}$ and $p$ together by finite element and multigrid techniques. Notice that the result is valid for all $\alpha\in[0,1/c_4^2]$. Proof of Theorem 3.2 for the case $\alpha=0$ is trivial, with $\gamma_1=\gamma_2=1$. Moreover, this choice reveals a perhaps simpler alternative: we can use a two-stage approach (cf. [13]) that first minimizes the flux-only functional,

$$G_0(\mathbf{u};f)=\left\|\frac{1}{\sqrt{a}}(\nabla\cdot\sqrt{a}\mathbf{u}+f)\right\|^2_{0,\Omega}+\left\|\sqrt{a}\,\nabla\times\left(\frac{\mathbf{u}}{\sqrt{a}}\right)\right\|^2_{0,\Omega}, \tag{3.17}$$

over $\mathbf{u}\in\mathbf{W}$, then fixes $\mathbf{u}/\sqrt{\mathbf{a}}$ and minimizes the Poisson functional,

$$G_P\left(p;\frac{\mathbf{u}}{\sqrt{a}}\right)=\left\|\nabla p-\frac{\mathbf{u}}{\sqrt{a}}\right\|^2_{0,\Omega},$$

over $p\in V$. The efficacy of this two-stage approach is confirmed by the uniform coercivity and continuity of $G_P(p;0)$ in the $H^1(\Omega)$ seminorm $\|\nabla p\|^2_{0,\Omega}$, which by (3.7) is itself a norm on $V$, and of $G_1(\mathbf{u};0)$ in the $\mathbf{W}$ seminorm as defined in (2.6), which we now demonstrate is a norm on $\mathbf{W}$ by establishing a Poincaré–Friedrichs inequality.

LEMMA 3.3. *We have*

$$\|\mathbf{u}\|_{0,\Omega}\le c_6\gamma|\mathbf{u}|_{\mathbf{W}} \quad \textit{for all} \quad \mathbf{u}\in\mathbf{W}, \tag{3.18}$$

*where* $c_6=\max\{c_4,c_5\}$ *(see* 3.9*) and* $\gamma$ *is from Lemma* 3.1.

*Proof.* Consider a Helmholtz decomposition on $\mathbf{W}$: for $\mathbf{u}\in\mathbf{W}$, there exist $p$, $\psi\in H^1(\Omega)$ such that

$$\mathbf{u}=\sqrt{a}\nabla p+\frac{1}{\sqrt{a}}\nabla^\perp\psi, \tag{3.19}$$

where $p$ is unique the solution of (2.1) with $f = -\nabla \cdot \sqrt{a}\mathbf{u}$ and $\psi$ is the unique (up to a constant) solution of

$$\begin{aligned} -\nabla \cdot \left(\frac{1}{a}\nabla\psi\right) &= -\nabla \times \frac{1}{\sqrt{a}}\mathbf{u} && \text{in } \Omega, \\ \psi &= C_i && \text{on } \Gamma_{N_i}, \\ \mathbf{n} \cdot \frac{1}{a}\nabla\psi &= 0 && \text{on } \Gamma_D, \end{aligned} \tag{3.20}$$

where $C_i$ are arbitrary constants, one of which may be set to zero. Since $\mathbf{u} \in \mathbf{W}$, it satisfies the integral constraints

$$\int_{\Gamma_{N_i}} \boldsymbol{\tau} \cdot \frac{1}{\sqrt{a}}\mathbf{u} = 0$$

for each disjoint piece of $\Gamma_N$. Thus, we may set the constants $C_i = 0$, and (3.20) will have a unique solution.

Note that the decomposition is orthogonal in the $L^2$ sense:

$$\left\langle \sqrt{a}\nabla p,\, \frac{1}{\sqrt{a}}\nabla^\perp\psi \right\rangle_{0,\Omega} = 0. \tag{3.21}$$

We thus have

$$\|\mathbf{u}\|_{0,\Omega}^2 = \|\sqrt{a}\nabla p\|_{0,\Omega}^2 + \left\| \frac{1}{\sqrt{a}}\nabla^\perp\psi \right\|_{0,\Omega}^2. \tag{3.22}$$

Now,

$$-\nabla \cdot a\nabla p = -\nabla \cdot \sqrt{a}\mathbf{u},$$

so that, using (3.10),

$$\begin{aligned} \|\sqrt{a}\nabla p\|_{0,\Omega}^2 &= \langle -\nabla \cdot a\nabla p,\, p\rangle_{0,\Omega} \\ &= \langle -\nabla \cdot \sqrt{a}\mathbf{u},\, p\rangle_{0,\Omega} \\ &= \left\langle -\frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u},\, \sqrt{a}p \right\rangle_{0,\Omega} \\ &\le \left\| \frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u} \right\|_{0,\Omega} \|\sqrt{a}p\|_{0,\Omega} \\ &\le c_4\gamma \left\| \frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u} \right\|_{0,\Omega} \|\sqrt{a}\nabla p\|_{0,\Omega}, \end{aligned}$$

which yields

$$\|\sqrt{a}\nabla p\|_{0,\Omega} \le c_4\gamma \left\| \frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u} \right\|_{0,\Omega}. \tag{3.23}$$

Similarly, using (3.11),

$$\begin{aligned}\left\| \frac{1}{\sqrt{a}} \nabla^{\perp} \psi \right\|_{0,\Omega}^2 &= \left\langle -\nabla \times \frac{1}{a} \nabla^{\perp} \psi,\, \psi \right\rangle_{0,\Omega} \\ &= \left\langle -\nabla \times \frac{1}{\sqrt{a}} \mathbf{u},\, \psi \right\rangle_{0,\Omega} \\ &= \left\langle -\sqrt{a} \nabla \times \frac{1}{\sqrt{a}} \mathbf{u},\, \frac{1}{\sqrt{a}} \psi \right\rangle_{0,\Omega} \\ &\le \left\| \sqrt{a} \nabla \times \frac{1}{\sqrt{a}} \mathbf{u} \right\|_{0,\Omega} \left\| \frac{1}{\sqrt{a}} \psi \right\|_{0,\Omega} \\ &\le c_5 \gamma \left\| \sqrt{a} \nabla \times \frac{1}{\sqrt{a}} \mathbf{u} \right\|_{0,\Omega} \left\| \frac{1}{\sqrt{a}} \nabla^{\perp} \psi \right\|_{0,\Omega},\end{aligned}$$

which yields

$$\left\| \frac{1}{\sqrt{a}} \nabla^{\perp} \psi \right\|_{0,\Omega} \le c_5 \gamma \left\| \sqrt{a} \nabla \times \frac{1}{\sqrt{a}} \mathbf{u} \right\|_{0,\Omega}. \tag{3.24}$$

The result now follows from (3.22)–(3.24), where $c_6 = \max\{c_4, c_5\}$. ◻

For simplicity of discussion, the following sections focus on the two-stage approach described above.

**4. Scaled div-curl operator.** We are now in a position to define the scaled div-curl operator and develop some tools that will aid in the proof of the decomposition of $\mathbf{W}$ in the next section. Define $\mathcal{L} : \mathbf{W} \to (L^2(\Omega))^2$ as follows:

$$\mathcal{L} := \begin{bmatrix} \frac{1}{\sqrt{a}} \nabla \cdot \sqrt{a} \\ \sqrt{a} \nabla \times \frac{1}{\sqrt{a}} \end{bmatrix}, \tag{4.1}$$

with domain $\mathcal{D}(\mathcal{L}) = \mathbf{W}$. It is straightforward to verify that the adjoint of $\mathcal{L}$ is given by

$$\mathcal{L}^* := -\left[ \sqrt{a} \nabla \frac{1}{\sqrt{a}},\, \frac{1}{\sqrt{a}} \nabla^{\perp} \sqrt{a} \right], \tag{4.2}$$

with domain

(4.3)

$$\mathcal{D}(\mathcal{L}^*) := \left\{ \mathbf{q} \;:\; \left( \frac{1}{\sqrt{a}} q_1,\, \sqrt{a} q_2 \right)^t \in (H^1(\Omega))^2,\, q_1 = 0 \text{ on } \Gamma_D,\, q_2 = C_i \text{ on } \Gamma_{N_i} \right\},$$

where $C_i$ are arbitrary constants, one of which may be set to zero. We summarize properties of $\mathcal{L}$ and $\mathcal{L}^*$ in the following lemma.

LEMMA 4.1. *The operator $\mathcal{L}$ is continuous and coercive on $\mathbf{W}$, the range $\mathcal{R}(\mathcal{L})$ is closed in $(L^2(\Omega))^2$, and*

$$\mathcal{R}(\mathcal{L})^{\perp} = \mathcal{N}(\mathcal{L}^*) = \left\{ \begin{pmatrix} 0 \\ \frac{1}{\sqrt{a}} \end{pmatrix} \right\}.$$

*Proof.* The first result follows directly from Lemma 3.3. For the second result, note for $\mathbf{u} \in \mathbf{W}$ we have

$$\|\mathbf{u}\|_{\mathbf{W}} \le (c_6\gamma + 1)|\mathbf{u}|_{\mathbf{W}} = \|\mathcal{L}\mathbf{u}\| \le \|\mathbf{u}\|_{\mathbf{W}}, \tag{4.4}$$

which implies that $\mathcal{R}(\mathcal{L})$ is closed in $(L^2(\Omega))^2$. For the last result, note for $\mathbf{u} \in \mathbf{W}$ that

$$\left\langle \frac{1}{\sqrt{a}}\nabla \cdot \sqrt{a}\mathbf{u},\, 0 \right\rangle + \left\langle \sqrt{a}\nabla \times \frac{1}{\sqrt{a}}\mathbf{u},\, \frac{1}{\sqrt{a}} \right\rangle = \int\int_\Omega \nabla \times \frac{1}{\sqrt{a}}\mathbf{u} = \oint \boldsymbol{\tau} \cdot \frac{1}{\sqrt{a}}\mathbf{u} = 0.$$

The last equality follows from the boundary conditions imposed on $\mathbf{u}$. Thus, $(0, \frac{1}{\sqrt{a}})^t \in \mathcal{R}(\mathcal{L})^\perp = \mathcal{N}(\mathcal{L}^*)$.

To show that this function spans $\mathcal{N}(\mathcal{L}^*)$, suppose that $\mathbf{q} \in \mathcal{D}(\mathcal{L}^*)$ satisfies

$$-\mathcal{L}^*\mathbf{q} = \sqrt{a}\nabla \frac{1}{\sqrt{a}}q_1 + \frac{1}{\sqrt{a}}\nabla^\perp \sqrt{a}q_2 = \mathbf{0}. \tag{4.5}$$

Let $p_1 = q_1/\sqrt{a}$, $p_2 = \sqrt{a}q_2$. From the boundary conditions on $\mathbf{q}$ and (4.5), we see that

$$\mathbf{n} \cdot \sqrt{a}\nabla p_1 = \mathbf{n} \cdot \left( \sqrt{a}\nabla p_1 + \frac{1}{\sqrt{a}}\nabla^\perp p_2 \right) = 0 \qquad \text{on} \;\; \Gamma_N. \tag{4.6}$$

Since $\frac{1}{\sqrt{a}}\nabla^\perp p_2 \in H(\operatorname{div} a; \Omega)$, then $\sqrt{a}\nabla p_1 \in H(\operatorname{div} a; \Omega)$. Thus, $p_1$ satisfies (2.1) with homogeneous data, which implies that $p_1 = 0$. This leaves $\nabla^\perp p_2 = 0$, which implies $p_2 = C$ and finally $q_2 = \frac{C}{\sqrt{a}}$ for some arbitrary constant $C$. Since this is the only solution of (4.5), the result is proved. ◻

Next, we define the restriction of $\mathcal{L}$ to $\mathbf{W}_S^1$:

$$\widehat{\mathcal{L}} := \mathcal{L}|_{\mathbf{W}_S^1}. \tag{4.7}$$

Since $\widehat{\mathcal{L}} \subseteq \mathcal{L}$, we know that $\mathcal{L}^* \subseteq \widehat{\mathcal{L}}^*$; that is,

$$\mathcal{D}(\widehat{\mathcal{L}}^*) = \left\{ \mathbf{q} \in (L^2(\Omega))^2 \;:\; \mathcal{L}^*\mathbf{q} \in (L^2(\Omega))^2,\, q_1 = 0 \text{ on } \Gamma_D,\, q_2 = C_i \text{ on } \Gamma_{N_i} \right\}. \tag{4.8}$$

This larger definition of $\mathcal{D}(\widehat{\mathcal{L}}^*)$ will be important in proving the decomposition in the next section. Finally, we have the following result.

LEMMA 4.2. *Subspace $\mathbf{W}_S^1$ is closed in $\mathbf{W}$ and $\mathcal{R}(\widehat{\mathcal{L}}) \subseteq \mathcal{R}(\mathcal{L})$ are both closed in $(L^2(\Omega))^2$.*

*Proof.* The result is an immediate consequence of Theorem 2.3, Corollary 2.4, and Lemma 4.1. ◻

**5. Solution decomposition.** Here, we introduce a splitting of the flux space $\mathbf{W}$ into a finite-dimensional space spanned by singular functions and locally smooth functions, that is, functions that are $H_S^1(\Omega)$. As a result, the flux $\mathbf{u}$ can be discretized as the sum of singular basis functions and standard basis functions that satisfy the interface conditions. This splitting provides the foundation for the finite element method that we present in [3]. For a detailed description of the finite element spaces, see also [2].

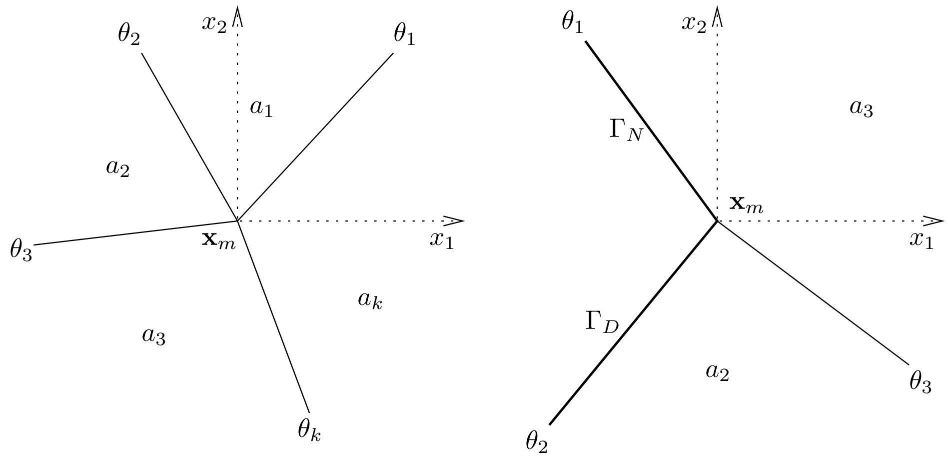


FIG. 5.1. *Cross point (on the left, $K = 4$), and boundary cross point (on the right, $K = 3$).*

In this context, a singular function is any function $\mathbf{u} \in \mathbf{W}$ such that $\mathbf{u} \notin \mathbf{W}^1_S(\Omega)$. This leads to a decomposition of any $\mathbf{u} \in \mathbf{W}$ as

$$\mathbf{u} = \mathbf{u}_0 + \sum_{m=1}^{M} \sum_{n=1}^{N_m} b_{m,n} \mathbf{s}_{m,n}, \tag{5.1}$$

where $\mathbf{u}_0 \in \mathbf{W}^1_S$, and $\mathbf{s}_{m,n}$, $n = 1, \ldots, N_m$, are singular functions associated with singular points $\mathbf{x}_m$, $m = 1, \ldots, M$.

This decomposition will be established, following the development in Kellogg [24] and Grisvard [19], by demonstrating a linearly independent set of functions $\mathbf{s}_{m,n} \in \mathbf{W} \backslash \mathbf{W}^1_S$ and then using a counting argument to show that they span all of $\mathbf{W} \backslash \mathbf{W}^1_S$. In fact, we will demonstrate two sets of functions, one associated with singular solutions of (2.1) and the other associated with singular solutions of (3.20), and show that they span the same space. The fact that they span the same space will be essential to the counting argument.

We first examine singular functions of the original equation (2.1). A singular function of (2.1) is a function $p \in H^1(\Omega) \setminus H^2_S(\Omega)$ for which $\nabla \cdot a \nabla p \in L^2(\Omega)$. As described in the introduction, singular points are associated with cross points, boundary cross points, reentrant corners, and irregular boundary points.

We begin with interior singular points. Boundary singular points are handled in a similar manner. First, we restrict our attention to the ball of radius $R$, call it $B_m(R)$, centered at the singular point $\mathbf{x}_m$ that contains no other singular points, and we establish a polar coordinate system $(r, \theta)$ centered at $\mathbf{x}_m$. For example, consider Figure 5.1. Denote the angle of the boundaries between segments to the positive $x_1$-axis by $\theta_i$ for $i = 1, \ldots, K$. In the following, we use the convention that $\theta_{-1} = \theta_K$ and $\theta_{K+1} = \theta_1$.

We seek solutions of the homogeneous equation

$$\nabla \cdot a \nabla p = \partial_r a \partial_r p + \frac{1}{r} a \partial_r p + \frac{1}{r^2} \partial_\theta a \partial_\theta p = 0 \tag{5.2}$$

in $B_m(R)$. Substituting $p = r^\alpha T(\theta)$ and dividing by $r^{\alpha-2}$ yields the problem

$$-(a T_\theta(\theta))_\theta = (a\alpha^2 + r a_r \alpha) T(\theta). \tag{5.3}$$

Here, we make the additional assumption on $a$ that, within each segment, $\lim_{r\to 0} a_\theta = 0$. Since it was assumed above that $a \in C^{1,1}(\Omega_i)$ for each subdomain $\Omega_i$, we also know that $\lim_{r\to 0} r a_r = 0$. Thus, we may substitute the value

$$\tilde{a}_i = \lim_{r\to 0} a(r,\theta) \qquad \text{in} \quad \Omega_i. \tag{5.4}$$

With this replacement, (5.3) now becomes the the Sturm–Liouville eigenvalue problem

$$-(\tilde{a}T')' = \tilde{a}\alpha^2 T \quad \text{on } [0,2\pi). \tag{5.5}$$

Solutions of this equation are of the form

$$T_n(\theta) = A_{n,i}\cos(\alpha_n(\theta-\theta_i)) + B_{n,i}\sin(\alpha_n(\theta-\theta_i)), \tag{5.6}$$

for $\theta \in (\theta_i, \theta_{i+1})$, with corresponding eigenvalue

$$\lambda_n = \alpha_n^2. \tag{5.7}$$

The singular functions we seek are constructed by choosing only those $\alpha_n \in (0,1)$ for, say, $n = 1, \ldots, N_m$. Note that for any solution with $\alpha = \alpha_n \in (0,1)$, there is a solution with $\alpha = -\alpha_n \in (-1,0)$. These solutions will be important in the counting argument.

Now, let $\tilde{\delta}_m(r) \in H^2(0,R)$ be a smooth cut-off function that is equal to 1 for $r \in (0, R/2)$ and drops to 0 for $r \in (R/2, R)$. It is easy to see that

$$s_{m,n} := \tilde{\delta}_m(r) r^{\alpha_n} T_n(\theta) \tag{5.8}$$

is in the domain of boundary value problem (2.1). Moreover, for any cut-off function $\delta_m \in H^1(0,R)$, we see that

$$\mathbf{s}_{m,n} := \delta_m(r)\sqrt{a}\nabla r^{\alpha_n} T_n(\theta) \in \mathbf{W} \setminus \mathbf{W}_S^1. \tag{5.9}$$

The exponent $\alpha$ and the coefficients $(A_i, B_i)$ can be determined by enforcing continuity of both $T(\theta)$ and $aT'(\theta)$ across interfaces. (We have dropped the first subscript for convenience.) This may be expressed as

$$\begin{bmatrix} 1 & 0 \\ 0 & -\tilde{a}_i \end{bmatrix} \begin{pmatrix} A_i \\ B_i \end{pmatrix} = \begin{bmatrix} \cos(\alpha(\theta_i-\theta_{i-1})) & \sin(\alpha(\theta_i-\theta_{i-1})) \\ \tilde{a}_{i-1}\sin(\alpha(\theta_i-\theta_{i-1})) & -\tilde{a}_{i-1}\cos(\alpha(\theta_i-\theta_{i-1})) \end{bmatrix} \begin{pmatrix} A_{i-1} \\ B_{i-1} \end{pmatrix}, \tag{5.10}$$

for $i = 1, \ldots, K$. Divide the second equation by $\tilde{a}_{i-1}$, define $\delta_i := \tilde{a}_i/\tilde{a}_{i-1}$ and

$$D_i := \begin{bmatrix} 1 & 0 \\ 0 & -\delta_i \end{bmatrix}, \qquad C_i := \begin{bmatrix} \cos(\alpha(\theta_i-\theta_{i-1})) & \sin(\alpha(\theta_i-\theta_{i-1})) \\ \sin(\alpha(\theta_i-\theta_{i-1})) & -\cos(\alpha(\theta_i-\theta_{i-1})) \end{bmatrix}, \tag{5.11}$$

and finally define $\underline{\beta}_i := (A_i, B_i)^t$. Then, the above constraints may be expressed by the homogeneous system

$$M\underline{b} = \begin{bmatrix} D_1 & 0 & \cdots & C_K \\ -C_1 & D_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & -C_{K-1} & D_K \end{bmatrix} \begin{pmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \\ \vdots \\ \underline{\beta}_K \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{0} \\ \vdots \\ \underline{0} \end{pmatrix}. \tag{5.12}$$

A nontrivial solution exists only when the determinant of $M$ is zero. The corresponding null vector yields the coefficients.

We now turn our attention to singular solutions of the boundary value problem (3.20). In $B_m(R)$ we seek solutions to the homogeneous problem

$$\nabla \cdot \frac{1}{a} \nabla p = 0. \tag{5.13}$$

Following the same arguments, we are led to the Sturm–Liouville eigenvalue problem

$$-\left(\frac{1}{\tilde{a}} \hat{T}'\right)' = \frac{1}{\tilde{a}} \alpha^2 \hat{T} \quad \text{on } [0, 2\pi) \tag{5.14}$$

and solutions of the form

$$\hat{T}_n(\theta) = \hat{A}_{n,i} \cos(\alpha_n(\theta - \theta_i)) + \hat{B}_{n,i} \sin(\alpha_n(\theta - \theta_i)), \tag{5.15}$$

for $\theta \in (\theta_i, \theta i + 1)$.

Again, we choose only those $\alpha_n \in (0, 1)$. With $\tilde{\delta}(r) \in H^2(0, R)$, solutions of this Sturm–Liouville problem yield

$$\hat{s}_{m,n} = \tilde{\delta}_m(r) r^{\alpha_n} \hat{T}_n(\theta) \tag{5.16}$$

in the domain of boundary value problem (3.20) and, with $\delta_m \in H^1(0, R)$,

$$\hat{\mathbf{s}}_{m,n} = \delta_m(r) \frac{1}{\sqrt{a}} \nabla^{\perp} r^{\alpha_n} \hat{T}_n(\theta) \in \mathbf{W} \setminus \mathbf{W}^1_S. \tag{5.17}$$

It would appear that there are at least two families of singular function in $\mathbf{W} \backslash \mathbf{W}^1_S$. We now show that they are in fact the same family. To see this, first notice that the only change to the continuity constraints (5.10) is that $\tilde{a}_i$, $\tilde{a}_{i-1}$ are replaced by $1/\tilde{a}_i$ and $1/\tilde{a}_{i-1}$ respectively, which results in replacing $D_i$ by $D_i^{-1}$. Thus, with the definition $\underline{\hat{\beta}}_i := (\hat{A}_i, \hat{B}_i)$ and similar notation for the other variables, the homogeneous system (5.12) becomes

$$\hat{M} \underline{\hat{b}} := \begin{bmatrix} D_1^{-1} & 0 & \cdots & C_K \\ -C_1 & D_2^{-1} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & -C_{K-1} & D_K^{-1} \end{bmatrix} \begin{pmatrix} \underline{\hat{\beta}}_1 \\ \underline{\hat{\beta}}_2 \\ \vdots \\ \underline{\hat{\beta}}_K \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{0} \\ \vdots \\ \underline{0} \end{pmatrix}. \tag{5.18}$$

We now show that $\det M = \det(\hat{M})$. Define the $2 \times 2$ rotation

$$Q_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{5.19}$$

and notice that $Q_2^t Q_2 = I_2, Q_2 C_i Q_2 = C_i$, and

$$Q_2 D_i Q_2 = \begin{bmatrix} \delta_i & 0 \\ 0 & -1 \end{bmatrix} = \delta_i D_i^{-1}. \tag{5.20}$$

Note that $\det(Q_2) = -1$ and define the $2K \times 2K$ block diagonal matrix $Q = \mathrm{diag}(Q_2, Q_2, \ldots, Q_2)$. This yields

$$QMQ = \begin{bmatrix} \delta_1 D_1^{-1} & 0 & \cdots & C_K \\ -C_1 & \delta_2 D_2^{-1} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & -C_{K-1} & \delta_K D_K^{-1} \end{bmatrix}. \tag{5.21}$$

Next, define the $2K \times 2K$ block matrices

$$\Delta_1 := \mathrm{diag}(\tilde{a}_1 I_2, \tilde{a}_2 I_2, \ldots, \tilde{a}_K I_2),$$
$$\Delta_2 := \mathrm{diag}(\tilde{a}_K I_2, \tilde{a}_1 I_2, \ldots, \tilde{a}_{K-1} I_2).$$

We can now establish

$$\Delta_2 QMQ\Delta_1^{-1} = \hat{M}, \tag{5.22}$$

which yields

$$\det(\hat{M}) = \det(\Delta_1)\det(\Delta_2^{-1})\det(Q)^2\det(M) = \det(M). \tag{5.23}$$

Let $\alpha_n \in (0,1)$ be a root of $\det(M) = 0$, and consider the associated null vector $M\underline{b}_n = 0$. Using the above relationships, we have

$$0 = (\Delta_2 QM)\underline{b}_n = (\Delta_2 QMQ\Delta_1^{-1})(\Delta_1 Q^t \underline{b}_n) = \hat{M}(\Delta_1 Q^t \underline{b}_n). \tag{5.24}$$

Thus, $\hat{b}_n = (\Delta_1 Q^t \underline{b}_n)$ is the corresponding null vector of $\hat{M}$, which yields

$$\begin{pmatrix} \hat{A}_{n,i} \\ \hat{B}_{n,i} \end{pmatrix} = \tilde{a}_i \begin{pmatrix} -B_{n,i} \\ A_{n,i} \end{pmatrix}. \tag{5.25}$$

For convenience, define

$$\phi_n(r,\theta) = r^{\alpha_n}\left(A_{n,i}\cos(\alpha_n(\theta-\theta_i)) + B_{n,i}\sin(\alpha_n(\theta-\theta_i))\right), \tag{5.26}$$
$$\psi_n(r,\theta) = r^{\alpha_n}(\hat{A}_{n,i}\cos(\alpha_n(\theta-\theta_i)) + \hat{B}_{n,i}\sin(\alpha_n(\theta-\theta_i))), \tag{5.27}$$

for $\theta \in (\theta_i, \theta_{i+1})$. Recall that

$$\nabla = \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix} = \begin{bmatrix} \cos(\theta) & -\frac{1}{r}\sin(\theta) \\ \sin(\theta) & \frac{1}{r}\cos(\theta) \end{bmatrix} \begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} \tag{5.28}$$

and that $\nabla^\perp = Q_2^t \nabla$. Using (5.25), (5.26), and (5.28), it is a simple matter to confirm that

$$\sqrt{a}\nabla\phi_n = \frac{1}{\sqrt{a}}\nabla^\perp\psi_n. \tag{5.29}$$

Boundary singular points are handled in a similar fashion. Now, instead of periodic boundary conditions, the Sturm–Liouville problem (5.5) would require $T(\theta) = 0$ for $\theta$ corresponding to a boundary segment in $\Gamma_D$, and $T'(\theta) = 0$ for $\theta$ corresponding to $\Gamma_N$, while problem (5.14) would reverse the roles. It is straightforward to verify that the relationship (5.29) holds for these singular functions as well.

We summarize the above discussion and complete the proof of the decomposition (5.1) in the following theorem.

THEOREM 5.1. *Every* $\mathbf{u} \in \mathbf{W}$ *has a unique decomposition*

$$\mathbf{u} = \mathbf{u}_0 + \sum_{m=1}^{M} \sum_{n=1}^{N_m} b_{m,n}\mathbf{s}_{m,n},$$

*where* $\mathbf{u}_0 \in \mathbf{W}_S^1$ *and* $\mathbf{s}_{m,n}$, $n = 1, \dots, N_m$, *are singular functions associated with singular points* $\mathbf{x}_m$, $m = 1, \dots, M$.

*Proof.* From Lemma 4.1, we know that $\mathbf{W}_S^1$ is closed in $\mathbf{W}$, that $\mathcal{R}(\widehat{\mathcal{L}}) \subseteq \mathcal{R}(\mathcal{L})$ are both closed in $(L^2(\Omega))^2$, and that both $\mathcal{L}$ and $\widehat{\mathcal{L}}$ are injective. Thus, the codimension of $\mathbf{W}_S^1$ in $\mathbf{W}$ is the same as the codimension of $\mathcal{R}(\widehat{\mathcal{L}})$ in $\mathcal{R}(\mathcal{L})$. By Lemma 4.1, we know that the dimension of $\mathcal{R}(\mathcal{L})^\perp$ is one. We now seek $\mathcal{R}(\widehat{\mathcal{L}})^\perp = \mathcal{N}(\widehat{\mathcal{L}}^*)$. At each singular point $\mathbf{x}_m$, let $\hat{\delta} \in H^2(0, R)$ be a smooth cut-off function and, for each $\alpha_{m,n} \in (0, 1)$, construct functions similar to (5.8) and (5.16) as follows:

$$s_{m,n}^- := \delta_m(r) r^{-\alpha_{m,n}} T_{m,n}(\theta),$$
$$\hat{s}_{m,n}^- := \delta_m(r) r^{-\alpha_{m,n}} \hat{T}_{m,n}(\theta),$$

and define

$$\mathbf{s}_{m,n}^- := (s_{m,n}^-, -\hat{s}_{m,n}^-)^t. \tag{5.30}$$

From (5.29) we see that $\mathbf{s}_{m,n}^- \in \mathcal{D}(\widehat{\mathcal{L}}^*) \setminus \mathcal{D}(\mathcal{L}^*)$ and $\widehat{\mathcal{L}}\mathbf{s}_{m,n}^- \in (L^2(\Omega))^2$. Since $\mathcal{L}^*$ is surjective, we can find $\mathbf{q}_{m,n} \in \mathcal{D}(\mathcal{L}^*)$ such that

$$\mathcal{L}^* \mathbf{q}_{m,n} = -\widehat{\mathcal{L}}^* \mathbf{s}_{m,n}^- \tag{5.31}$$

and set

$$\mathbf{f}_{m,n} = \mathbf{q}_{m,n} + \mathbf{s}_{m,n}^-. \tag{5.32}$$

Clearly, $\mathbf{f}_{m,n} \in \mathcal{N}(\widehat{\mathcal{L}}^*)$.

It is straightforward to show that every element of $\mathcal{N}(\widehat{\mathcal{L}}^*)$ must be of this form, that is, must involve singular functions of both (2.1) and (3.20). Thus, the dimension of $\mathcal{N}(\widehat{\mathcal{L}}^*)$ is exactly equal to the number of such functions plus the one function in $\mathcal{N}(\mathcal{L}^*)$. We complete the proof by noting that the codimension of $\mathcal{N}(\mathcal{L}^*)$ in $\mathcal{N}(\widehat{\mathcal{L}}^*)$ is equal to the codimension of $\mathcal{R}(\widehat{\mathcal{L}})$ in $\mathcal{R}(\mathcal{L})$. ∎

This decomposition is the basis for the finite element discretization that is developed in the companion paper [3]. We only summarize the basic ideas here. Exponents and coefficients of singular basis functions $\mathbf{s}_{m,n}$ can be computed from the geometry of interfaces adjoining a singular point and the jumps in the coefficient $a$ across these interfaces. Although our theoretical development employed cut-off functions independent of $\theta$, any $H^1$ cut-off function may be used. We choose cut-off functions that equal one in a fixed region around the singular point and fall off to zero linearly in a small fringe region of width one grid cell.

The singular basis functions are included in the finite element space, together with standard elements, such as linear elements on triangles, that satisfy the interface conditions. Using functional $G_0$ to solve for the flux, inner products of standard elements with singular basis functions need only be calculated in the fringe region, thus saving a significant amount of work.

**6. Conclusions.** In this paper we have developed a FOSLS $L^2$ formulation for diffusion equations with discontinuous coefficients, irregular boundaries, and mixed boundary conditions. In Theorem 3.2, we showed the functional $G_\alpha$ in (3.1) to be coercive and continuous in $\mathbf{W} \times V$ with constants that are $P$-uniform. We then explored the flux-only functional, $G_0$ in (3.17), and in Lemma 3.3 and Lemma 4.1 showed that it is coercive and continuous in $\mathbf{W}$ with constants that are also $P$-uniform. Properties of the scaled div-curl operator (4.1) helped us to prove in Theorem 5.1 that $\mathbf{W}$ can be split into functions that are $H^1$ in each subdomain plus a finite number of singular basis functions with support in the neighborhood of the singular points.

These results form the theoretical basis for the finite element discretization of $\mathbf{W}$, a rigorous discretization error analysis, and a multilevel method, all of which are presented in the companion paper [3]. Our approach is different from others (see, for example, [9]) in that a rigorous discretization error analysis in the presence of approximate singular basis functions is possible, and a multilevel method can be devised that incorporates singular basis functions on all levels.

REFERENCES


[1] R. E. Alcouffe, A. Brandt, J. J. E. Dendy, Jr., and J. W. Painter, *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Stat. Comput., 2 (1981), pp. 430–454.

[2] M. Berndt, *Adaptive Refinement and the Treatment of Discontinuous Coefficients for Multilevel First-Order System Least Squares (FOSLS)*, Ph.D. thesis, Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO, 1999.

[3] M. Berndt, T. A. Manteuffel, and S. F. McCormick, *Analysis of first-order system least squares (FOSLS) for elliptic problems with discontinuous coefficients: Part* II, SIAM J. Numer. Anal., 43 (2005), pp. 409–436.

[4] P. Bochev, Z. Cai, T. A. Manteuffel, and S. F. McCormick, *Analysis of velocity-flux first-order system least-squares principles for the Navier–Stokes equations: Part* I, SIAM J. Numer. Anal, 35 (1998), pp. 990–1009.

[5] P. B. Bochev and M. D. Gunzburger, *Finite element methods of least-squares type*, SIAM Rev. 40 (1998), pp. 789–837.

[6] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *Least-squares methods for the Stokes equations based on a discrete minus one inner product*, J. Comp. Appl. Math., 74 (1996), pp. 155–173.

[7] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *A least-squares approach based on a discrete minus one inner product for first order systems*, Math. Comp., 66 (1997), pp. 935–955.

[8] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994.

[9] S. C. Brenner and L. Y. Sung, *Multigrid methods for the computation of singular solutions and stress intensity factors* II, BIT, 37 (1997), pp. 623–643.

[10] Z. Cai and S. Kim, *A finite element method using singular functions for the Poisson equation: Corner singularities*, SIAM J. Numer. Anal., 39 (2001), pp. 286–299.

[11] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part* I, SIAM J. Numer. Anal., 31 (1994), pp. 1785–1799.

[12] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part* II, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.

[13] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for the Stokes equations, with application to linear elasticity*, SIAM J. Numer. Anal., 34 (1997), pp. 1727–1741.

[14] Z. Cai, T. A. Manteuffel, S. F. McCormick, and J. Ruge, *First-order system $LL^*$ (FOSLL$^*$): Scalar elliptic partial differential equations*, SIAM J. Numer. Anal., 39 (2001), pp. 1418–1445.

[15] T. F. Chen and G. J. Fix, *Least squares finite element simulation of transonic flows*, Appl. Numer. Math., 2 (1986), pp. 399–408.

[16] C. L. Cox and G. J. Fix, *On the accuracy of least squares methods in the presence of corner singularities*, Comput. Math. Appl., 10 (1984), pp. 463–475.

[17] G. Fix and E. Stephan, *Finite Element Methods of the Least Squares Type for Regions with Corners*, Tech. Report 81-41, Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA, 1981.

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

[19] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.

[20] M. J. Holst, *Multilevel Methods for the Poisson–Boltzmann Equation*, Ph.D. thesis, Numerical Computing Group, University of Illinois at Urbana-Champaign, Urbana, IL, 1993.

[21] D. C. Jespersen, *A least-square decomposition method for solving elliptic systems*, Math. Comp., 31 (1977), pp. 873–880.

[22] B. N. Jiang and J. Z. Chai, *Least-squares finite element analysis of steady high subsonic plane potential flows*, Acta Mech. Sinica, 1 (1980), pp. 90–93.

[23] B.-N. Jiang and C. L. Chang, *Least-squares finite elements for the Stokes problem*, Comput. Methods Appl. Mech. Engrg., 78 (1990), pp. 297–311.

[24] R. B. Kellogg, *Singularities in interface problems*, in Proceedings of the 2nd Annual Symposium on the Numerical Solution of Partial Differential Equations, B. Hubbard, ed., Academic Press, New York, 1971, pp. 351–400.

[25] T. A. Manteuffel, S. F. McCormick, and G. Starke, *First-order system least-squares for second-order elliptic problems with discontinuous coefficients*, in Proceedings of the Seventh Annual Copper Mountain Conference on Multigrid Methods, N. D. Melson, T. A. Manteuffel, and S. F. McCormick, eds., NASA, Hampton, VA, 1995, pp. 535–550.

[26] P. Neittaanmäki and J. Saranen, *On finite element approximation of the gradient for the solution to Poisson equation*, Numer. Math., 37 (1981), pp. 131–148.

[27] A. I. Pehlivanov and G. F. Carey, *Error estimates for least-squares mixed finite elements*, RAIRO Modél. Math. Anal. Numer., 28 (1994), pp. 499–516.

[28] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1368–1377.

[29] T. F. Russell and M. F. Wheeler, *Finite element and finite difference methods for continuous flows in porous media*, in The Mathematics of Reservoir Simulation, R. E. Ewing, ed., Frontiers in Appl. Math. 1, SIAM, Philadelphia, 1983, pp. 35–106.

[30] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice–Hall, Englewood Cliffs, NJ, 1973.

[31] W. L. Wendland, *Elliptic Systems in the Plane*, Pitman, London, 1979.

[32] E. Zauderer, *Partial Differential Equations of Applied Mathematics*, 2nd ed., Ser. Pure Appl. Math., John Wiley & Sons, New York, 1988.