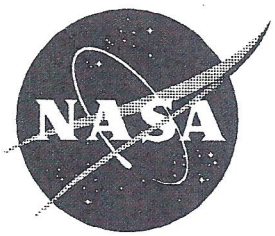

NASA Conference Publication 3339
Part 1

Seventh Copper Mountain Conference on Multigrid Methods

Edited by
N. Duane Melson
Langley Research Center • Hampton, Virginia

Tom A. Manteuffel and Steve F. McCormick
University of Colorado • Boulder, Colorado

Craig C. Douglas
IBM Thomas J. Watson Research Center • Yorktown Heights, New York
Yale University • New Haven, Connecticut

Proceedings of a workshop cosponsored by the National Aeronautics and Space Administration, Washington, D.C., and the Department of Energy, Washington, D.C., and held at Copper Mountain, Colorado April 2–7, 1995

National Aeronautics and Space Administration
Langley Research Center • Hampton, Virginia 23681-0001

September 1996

First-Order System Least-Squares for Second-Order Elliptic Problems with Discontinuous Coefficients

Thomas A. Manteuffel Stephen F. McCormick Gerhard Starke*

Abstract

The first-order system least-squares methodology represents an alternative to standard mixed finite element methods. Among its advantages is the fact that the finite element spaces approximating the pressure and flux variables are not restricted by the inf-sup condition and that the least-squares functional itself serves as an appropiate error measure. This paper studies the first-order system least-squares approach for scalar second-order elliptic boundary value problems with discontinuous coefficients. Ellipticity of an appropriately scaled least-squares bilinear form is shown independently of the size of the jumps in the coefficients leading to adequate finite element approximation results. The occurrence of singularities at interface corners and cross-points is discussed, and a weighted least-squares functional is introduced to handle such cases. Numerical experiments are presented for two test problems to illustrate the performance of this approach.

Introduction

The purpose of this paper is to apply the first-order system least-squares approach developed in [4] and [5] to scalar second-order elliptic boundary value problems in two dimensions with discontinuous coefficients. Such problems arise in various application areas, including flow in heterogeneous porous media (see, e.g., [12]), neutron transport [1], and biophysics [7]. In many physical applications, one is interested not only in an 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 modelled by the equation

$$-\nabla \cdot (a\nabla p) = f \tag{1}$$

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 fluid velocity

$$\mathbf{u} = a\nabla p\,, \tag{2}$$

a concern which led to the development of mixed finite element methods (see, e.g., [3, Chapter 10]). In mixed methods, both p and $\mathbf{u}$ are approximated by not necessarily identical finite elements and, roughly speaking, a Galerkin condition is imposed on the first-order system resulting from (1) and (2).

An alternative to mixed finite elements is the first-order system least-squares approach developed and analyzed, e.g., in [4], [5], [11], and [10]. This methodology replaces the Galerkin condition by the minimization of a least-squares functional associated with a first-order system derived from (1) and (2). Augmenting the basic system

*Program in Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526. E-mail {tmanteuf,stevem,starke}@boulder.colorado.edu

with the curl-condition $\nabla \times (\mathbf{u}/a) = 0$ (see [5], [10]) leads to ellipticity with respect to the $H^1(\Omega)$ norm in the individual variables. Important practical advantages of this least-squares approach over standard mixed methods are: (i) the finite element spaces approximating the pressure and flux variables are not restricted by the inf-sup condition of Ladyzhenskaya-Babuška-Brezzi (cf. [3, Section 10.5]) and (ii) the least-squares functional serves as an appropriate error measure. Moreover, if the problem is sufficiently regular (e.g., if $a \in C^{1,1}(\Omega)$ and Ω has certain properties (cf. [5])), then (iii) optimal accuracy is guaranteed in each variable, including the velocities, in the H^1 norm and (iv) optimal computational complexity for the solution of the resulting discrete systems is achieved with standard multigrid methods (see [5]).

For problems with discontinuous coefficients, which is our focus in this paper, the velocity components will, in general, not be in $H^1(\Omega)$. While the theory developed in [4] and [5] already allows for discontinuous coefficients, special care must be taken in order to prove ellipticity, in an appropriate norm, with constants independent of 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. This results in ellipticity, independently of the size of coefficient jumps, and consequently in finite element approximation results, with respect to a norm that is suitably scaled depending on the size of a. This scaling is presented in the following section.

At interface corners and cross-points (i.e., where two smooth interface components intersect), the components of $\mathbf{u}$ will, in general, be unbounded, and singularities naturally arise (see, for example, Strang and Fix [14, Ch. 8]). The shape of these singularities is determined by the angle at an interface corner (or between two intersecting interfaces) and the jumps in the coefficients. We will show how the parameters describing these singularities can be computed from the coefficient jumps and corner angles. We are particularly interested in the exponent associated with the singular function at a corner or cross-points since this indicates how much we have to unweight the least-squares functional in the neighborhood of such a point. The performance of this scaled least-squares approach will be studied using bilinear finite elements for the pressure and fluxes (based on the same grid) and a full multigrid algorithm for the solution of the resulting discrete system. Finally, computational experiments for two test problems are presented.

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

The Least-Squares Functional

Consider the following prototype problem on a bounded domain $\Omega \subset \mathbb{R}^2$:

$$\begin{aligned} -\nabla \cdot (a \nabla p) &= f, && \text{in } \Omega, \\ p &= 0, && \text{on } \Gamma_D, \\ \mathbf{n} \cdot \nabla p &= 0, && \text{on } \Gamma_N, \end{aligned} \tag{3}$$

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 Ω but may have large jumps across interfaces. We assume that $\Gamma_D \neq \emptyset$, so that the Poincaré-Friedrichs inequality

$$\|p\|_{0,\Omega} \leq \gamma \|\nabla p\|_{0,\Omega} \tag{4}$$

holds and (3) has a unique solution in $H^1(\Omega)$. Following [5], we rewrite (3) as a first-order system by introducing the flux variable $\mathbf{u} = a \nabla p$:

$$\left\{ \begin{aligned} \mathbf{u} - a \nabla p &= 0, && \text{in } \Omega, \\ -\nabla \cdot \mathbf{u} &= f, && \text{in } \Omega, \\ p &= 0, && \text{on } \Gamma_D, \\ \mathbf{n} \cdot \mathbf{u} &= 0, && \text{on } \Gamma_N. \end{aligned} \right. \tag{5}$$

Since $\mathbf{u}/a = \nabla p$ with $p \in H^1(\Omega)$, then we have (cf. [6, Theorem 2.9])

$$\nabla \times (\mathbf{u}/a) \equiv \partial_1(u_2/a) - \partial_2(u_1/a) = 0\,, \text{ in } \Omega\,.$$

Moreover, the homogeneous Dirichlet boundary condition on Γ_D implies the tangential flux condition

$$\mathbf{n} \times (\mathbf{u}/a) \equiv (n_1 u_2 - n_2 u_1)/a = 0\,, \text{ on } \Gamma_D\,.$$

Adding these equations to first-order system (5) yields the augmented system

$$\begin{array}{rcll} \mathbf{u} - a\nabla p &=& \mathbf{0}\,, & \text{in } \Omega\,, \\ -\nabla \cdot \mathbf{u} &=& f\,, & \text{in } \Omega\,, \\ \nabla \times (\mathbf{u}/a) &=& 0\,, & \text{in } \Omega\,, \\ p &=& 0\,, & \text{on } \Gamma_D\,, \\ \mathbf{n} \cdot \mathbf{u} &=& 0\,, & \text{on } \Gamma_N\,, \\ \mathbf{n} \times (\mathbf{u}/a) &=& 0\,, & \text{on } \Gamma_D\,. \end{array} \tag{6}$$

In addition to $L^2(\Omega)$ and $H^1(\Omega)$ with the respective norms $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{1,\Omega}$, we will need the spaces

$$\begin{aligned} H(\mathrm{div};\Omega) &= \{\mathbf{v} \in L^2(\Omega)^2 : \nabla \cdot \mathbf{v} \in L^2(\Omega)\}\,, \\ H(\mathrm{curl}\, a;\Omega) &= \{\mathbf{v} \in L^2(\Omega)^2 : \nabla \times (\mathbf{v}/a) \in L^2(\Omega)\} \end{aligned}$$

and

$$\begin{aligned} V &= \{q \in H^1(\Omega) : q = 0 \text{ on } \Gamma_D\}\,, \\ \mathbf{W} &= \{\mathbf{v} \in H(\mathrm{div};\Omega) \cap H(\mathrm{curl}\, a;\Omega) : \mathbf{n} \cdot \mathbf{v} = 0 \text{ on } \Gamma_N\,,\ \mathbf{n} \times (\mathbf{v}/a) = 0 \text{ on } \Gamma_D\}\,. \end{aligned} \tag{7}$$

Clearly, for the solution of (3), we have $p \in V$ and $\mathbf{u} \in \mathbf{W}$, so it is appropriate to pose (6) on these spaces.

As mentioned above, our main interest is in the solution of (3) when $a(x_1, x_2)$ has large jumps. Following Bramble, Pasciak, Wang, and Xu [2], we assume that

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

with $\{\Omega_i\}$ being mutually disjoint open polygonal regions; that the restriction of $a(x_1, x_2)$ to Ω_i is in $C^1(\Omega_i)$; and that

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

with constants c_1, c_2 of order one and arbitrary positive constants ω_i. In other words, $a(x_1, x_2)$ is assumed to be of approximate size ω_i throughout Ω_i for each i while large variations in $\{\omega_i\}$ over i are allowed. The bounds derived below will be independent of this variation in $\{\omega_i\}$, but the constants in these bounds will depend on the variation within each Ω_i, that is, on c_1 and c_2.

An appropriate scaling of the equations in (6) leads to the least-squares functional

$$G(\mathbf{u}, p; f) = \|\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p\|_{0,\Omega}^2 + \|\nabla \cdot \mathbf{u} + f\|_{0,\Omega}^2 + \|a\,\nabla \times (\mathbf{u}/a)\|_{0,\Omega}^2 \tag{8}$$

and associated bilinear form

$$\begin{aligned} \mathcal{F}(\mathbf{u}, p; \mathbf{v}, q) &= (\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p, \mathbf{v}/\sqrt{a} - \sqrt{a}\nabla q)_{0,\Omega} \\ &+ (\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v})_{0,\Omega} + (a\,\nabla \times (\mathbf{u}/a), a\,\nabla \times (\mathbf{v}/a))_{0,\Omega}\,. \end{aligned} \tag{9}$$

Here, for the sake of notational simplicity, we agree that $(\cdot,\cdot)_{0,\Omega}$ is meant componentwise for vector functions. That is, if $\mathbf{w} = (w_1, w_2)$ and $\mathbf{z} = (z_1, z_2)$, then

$$(\mathbf{w}, \mathbf{z})_{0,\Omega} = (w_1, z_1)_{0,\Omega} + (w_2, z_2)_{0,\Omega}\,.$$

The solution of (5) will also solve the minimization problem

$$G(\mathbf{u},p;f) = \min_{(\mathbf{v},q)\in \mathbf{W}\times V} G(\mathbf{v},q;f) \tag{10}$$

and, therefore, the variational problem

$$\mathcal{F}(\mathbf{u},p;\mathbf{v},q) = -(f,\nabla\cdot\mathbf{v})_{0,\Omega} \text{ for all } (\mathbf{v},q)\in \mathbf{W}\times V\,. \tag{11}$$

Here we show that $\mathcal{F}(\mathbf{v},q;\mathbf{v},q)$ is uniformly equivalent to the scaled norm defined for $(\mathbf{v},q)\in \mathbf{W}\times V$ by

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

Theorem 1 *Under the above assumptions, there exist constants γ_1 and γ_2, independent of the size of the jumps in $\{\omega_i\}$, such that*

$$\mathcal{F}(\mathbf{u},p;\mathbf{u},p) \geq \gamma_1 |||(\mathbf{u},p)|||^2 \quad \textit{for all } (\mathbf{u},p)\in \mathbf{W}\times V \tag{12}$$

and

$$\mathcal{F}(\mathbf{u},p;\mathbf{v},q) \leq \gamma_2 |||(\mathbf{u},p)||| \, |||(\mathbf{v},q)||| \quad \textit{for all } (\mathbf{u},p)\,,\ (\mathbf{v},q)\in \mathbf{W}\times V\,. \tag{13}$$

Proof. The proof is similar to the proof of [4, Theorem 3.1] (see also [10, Theorems 2.1 and 2.2]). We include it here because we must confirm that the constants γ_1 and γ_2 are independent of the jumps in a. The main part of the proof consists in showing that the functionals

$$\hat{\mathcal{F}}(\mathbf{u},p;\mathbf{v},q) = (\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p, \mathbf{v}/\sqrt{a} - \sqrt{a}\nabla q)_{0,\Omega} + (\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})_{0,\Omega}$$

and

$$\mathcal{S}(\mathbf{u},p;\mathbf{v},q) = (\mathbf{u}/\sqrt{a},\mathbf{v}/\sqrt{a})_{0,\Omega} + (\sqrt{a}\nabla p,\sqrt{a}\nabla q)_{0,\Omega} + (\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})_{0,\Omega}\,,$$

satisfy

$$c_1\mathcal{S}(\mathbf{u},p;\mathbf{u},p) \leq \hat{\mathcal{F}}(\mathbf{u},p;\mathbf{u},p) \tag{14}$$

and

$$\hat{\mathcal{F}}(\mathbf{u},p;\mathbf{v},q) \leq c_2(\mathcal{S}(\mathbf{u},p;\mathbf{u},p))^{1/2}(\mathcal{S}(\mathbf{v},q;\mathbf{v},q))^{1/2} \tag{15}$$

with constants c_1 and c_2 that are independent of the jumps in a.

For the proof of (14), we rewrite Poincaré-Friedrichs inequality (4) as

$$\|p\|^2_{0,\Omega} \leq \tilde{\gamma}\|\sqrt{a}\nabla p\|^2_{0,\Omega}\,. \tag{16}$$

Note that $\tilde{\gamma}$, and consequently the quantity γ_1 in (12), depends on $\min_{\mathbf{x}\in\Omega} a(\mathbf{x}) > 0$. It does not introduce, however, any dependence of (12) and (13) on the size of the jumps in a. Since on $\partial\Omega$ we either have $p = 0$ or $\mathbf{n}\cdot\mathbf{u} = 0$, then integration by parts confirms that

$$(\mathbf{u},\nabla p)_{0,\Omega} + (\nabla\cdot\mathbf{u},p)_{0,\Omega} = 0\,.$$

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

$$\begin{aligned}
&\hat{\mathcal{F}}(\mathbf{u},p;\mathbf{u},p)\\
&= (\mathbf{u}/\sqrt{a},\mathbf{u}/\sqrt{a})_{0,\Omega} + (\sqrt{a}\nabla p,\sqrt{a}\nabla p)_{0,\Omega} - 2(\mathbf{u},\nabla p)_{0,\Omega} + (\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{u})_{0,\Omega}\\
&\quad +2\tau(\nabla\cdot\mathbf{u},p)_{0,\Omega} + 2\tau(\mathbf{u},\nabla p)_{0,\Omega} + \tau^2(p,p)_{0,\Omega} - \tau^2(p,p)_{0,\Omega}\\
&= (\mathbf{u}/\sqrt{a} + (\tau-1)\sqrt{a}\nabla p, \mathbf{u}/\sqrt{a} + (\tau-1)\sqrt{a}\nabla p)_{0,\Omega}\\
&\quad +(\nabla\cdot\mathbf{u} + \tau p, \nabla\cdot\mathbf{u} + \tau p)_{0,\Omega} - \tau^2(p,p)_{0,\Omega} + (2\tau-\tau^2)(\sqrt{a}\nabla p,\sqrt{a}\nabla p)_{0,\Omega}\\
&\geq (2\tau-\tau^2)(\sqrt{a}\nabla p,\sqrt{a}\nabla p)_{0,\Omega} - \tau^2(p,p)_{0,\Omega}\\
&\geq (2\tau-(1+\tilde{\gamma})\tau^2)\|\sqrt{a}\nabla p\|^2_{0,\Omega}\,.
\end{aligned}$$

Choosing $\tau = 1/(1+\tilde{\gamma})$ leads to

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

We then also have

$$\|\mathbf{u}/\sqrt{a}\|_{0,\Omega}^2 \leq 2(\|\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p\|_{0,\Omega}^2 + \|\sqrt{a}\nabla p\|_{0,\Omega}^2) \leq 2(1+1/\tau)\hat{\mathcal{F}}(\mathbf{u},p;\mathbf{u},p)$$

and, clearly,

$$\|\nabla\cdot\mathbf{u}\|_{0,\Omega}^2 \leq \hat{\mathcal{F}}(\mathbf{u},p;\mathbf{u},p) ,$$

which completes the proof of (14).

Upper bound (15) follows from

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

and

$$\begin{aligned}\hat{\mathcal{F}}(\mathbf{u},p;\mathbf{u},p) &= \|\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p\|_{0,\Omega}^2 + \|\nabla\cdot\mathbf{u}\|_{0,\Omega}^2 \\ &\leq 2(\|\mathbf{u}/\sqrt{a}\|_{0,\Omega}^2 + \|\sqrt{a}\nabla p\|_{0,\Omega}^2 + \|\nabla\cdot\mathbf{u}\|_{0,\Omega}^2) = \mathcal{S}(\mathbf{u},p;\mathbf{u},p) .\end{aligned} \tag{17}$$

The proof of Theorem 1 is completed by adding the term $\|a\nabla\times(\mathbf{u}/a)\|_{0,\Omega}$ to both sides of the inequalities (14) and (17). ∎

Theorem 1 states that ellipticity and continuity of the least-squares bilinear form $\mathcal{F}(\cdot,\cdot;\cdot,\cdot)$ in terms of the norm $|||(\cdot,\cdot)|||$ is independent of the jumps in a. Note, however, that the ellipticity constant γ_1 in (12) depends on the size of a, in particular, on the positive constant $\min_{\mathbf{x}\in\Omega} a(\mathbf{x})$ through the Poincaré-Friedrichs inequality (16).

The scaling of the norm $|||(\cdot,\cdot)|||$ has the following physical interpretation. In areas where a is relatively small, ∇p is allowed to be relatively large, and one has to expect a less accurate approximation there compared to areas where a is large and ∇p is therefore small. In contrast, the velocity $\mathbf{u} = a\nabla p$ can be expected to be more accurate in areas where a is small and less accurate, in general, where a is large. Ellipticity with constants that are independent of the jumps in a asserts that the scaling in $\mathcal{F}(\cdot,\cdot;\cdot,\cdot)$ correctly reflects these attributes.

Singularities at Interface Corners and Cross-Points

This section is concerned with the behavior of p and $\mathbf{u}$ at or near the interface curve. Most of what we present in this section is well-known; we refer to Strang and Fix [14, Chapter 8] for further details.

Recall from the previous section that the solution of (6) satisfies $\mathbf{u} \in H(\mathrm{div};\Omega) \cap H(\mathrm{curl}\ a;\Omega)$. This implies that, at a point on a smooth segment of the interface curve, the normal component $\mathbf{n}\cdot\mathbf{u}$ and the tangential component $\mathbf{n}\times(\mathbf{u}/a)$ must be continuous. Assume that $\overline{\Omega} = \overline{\Omega}^+ \cup \overline{\Omega}^-$ with constant diffusion coefficients a^+ and a^-, respectively, and let $\mathbf{u}^+ = (u_1^+, u_2^+)$ and $\mathbf{u}^- = (u_1^-, u_2^-)$ denote the solution restricted to the respective subdomains (see Figure 1). Then u_1 and u_2 must satisfy the jump conditions

$$n_1u_1^+ + n_2u_2^+ = n_1u_1^- + n_2u_2^- \text{ and } n_2\frac{u_1^+}{a^+} - n_1\frac{u_2^+}{a^+} = n_2\frac{u_1^-}{a^-} - n_1\frac{u_2^-}{a^-} . \tag{18}$$

For example, consider the situation shown in Figure 1 (which we will encounter again as Example 2 in the final section of this paper). Across the vertical part of the interface, $u_1 = \mathbf{n}\cdot\mathbf{u}$ will be continuous while $u_2 = \mathbf{n}\times\mathbf{u}$ has a jump factor of a^+/a^-. Similarly, across the horizontal part of the interface, $u_1 = -\mathbf{n}\times\mathbf{u}$ has a jump factor of a^+/a^- while $u_2 = \mathbf{n}\cdot\mathbf{u}$ is continuous. At the interface corner, both of these conditions must be satisfied, i.e., u_1 and u_2 must jump by a factor a^+/a^- and be continuous at the same

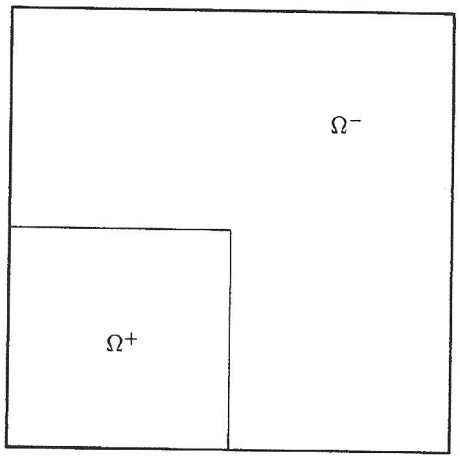

Figure 1: Interface with corner

time. Obviously, there are only two ways for this to happen: either $\mathbf{u} = \mathbf{0}$ or $\mathbf{u} = \infty$ at the interface corner. In general, the latter case is encountered at interface corners—the behavior of $\mathbf{u}$ is singular there.

Without loss of generality, assume that the singularity occurs at the origin, and consider the polar coordinate representation

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\,\cos\theta \\ r\,\sin\theta \end{pmatrix}$$

The solution of (3) then admits the representation

$$p(r,\theta) = \begin{cases} r^\alpha(\lambda_c^+ \cos\alpha\theta + \lambda_s^+ \sin\alpha\theta) + \tilde{p}^+(r,\theta)\,, & \text{in } \Omega^+\,, \\ r^\alpha(\lambda_c^- \cos\alpha\theta + \lambda_s^- \sin\alpha\theta) + \tilde{p}^-(r,\theta)\,, & \text{in } \Omega^-\,, \end{cases}$$

where $\tilde{p}^+ \in H^2(\Omega^+), \tilde{p}^- \in H^2(\Omega^-)$ (cf. [14, Section 8.1]), $\alpha \in (1/2, 1)$, and $\lambda_c^\pm, \lambda_s^\pm$ are constants. Using

$$\nabla = \begin{pmatrix} \partial_1 \\ \partial_2 \end{pmatrix} = \begin{pmatrix} \cos\theta\,\frac{\partial}{\partial r} - \sin\theta\,\frac{1}{r}\,\frac{\partial}{\partial\theta} \\ \sin\theta\,\frac{\partial}{\partial r} + \cos\theta\,\frac{1}{r}\,\frac{\partial}{\partial\theta} \end{pmatrix} \tag{19}$$

leads to

$$u_1(r,\theta) = \begin{cases} \alpha a^+ r^{\alpha-1}(\lambda_c^+ \cos(\alpha-1)\theta + \lambda_s^+ \sin(\alpha-1)\theta) + \tilde{u}_1^+(r,\theta)\,, & \text{in } \Omega^+\,, \\ \alpha a^- r^{\alpha-1}(\lambda_c^- \cos(\alpha-1)\theta + \lambda_s^- \sin(\alpha-1)\theta) + \tilde{u}_1^-(r,\theta)\,, & \text{in } \Omega^-\,, \end{cases} \tag{20}$$

and

$$u_2(r,\theta) = \begin{cases} \alpha a^+ r^{\alpha-1}(-\lambda_c^+ \sin(\alpha-1)\theta + \lambda_s^+ \cos(\alpha-1)\theta) + \tilde{u}_2^+(r,\theta)\,, & \text{in } \Omega^+\,, \\ \alpha a^- r^{\alpha-1}(-\lambda_c^- \sin(\alpha-1)\theta + \lambda_s^- \cos(\alpha-1)\theta) + \tilde{u}_2^-(r,\theta)\,, & \text{in } \Omega^-\,, \end{cases} \tag{21}$$

with $\tilde{u}_1^+, \tilde{u}_2^+ \in H^1(\Omega^+)$ and $\tilde{u}_1^-, \tilde{u}_2^- \in H^1(\Omega^-)$. The parameters $\alpha, \lambda_c^+, \lambda_s^+, \lambda_c^-$, and λ_s^- are computed such that conditions (18) are fulfilled. Setting $\mu = a^+/a^-$ leads to the matrix equation

$$\begin{bmatrix} -\mu\sin\alpha\frac{3}{2}\pi & \mu\cos\alpha\frac{3}{2}\pi & -\sin\alpha\frac{\pi}{2} & -\cos\alpha\frac{\pi}{2} \\ -\cos\alpha\frac{3}{2}\pi & -\sin\alpha\frac{3}{2}\pi & \cos\alpha\frac{\pi}{2} & -\sin\alpha\frac{\pi}{2} \\ -\cos\alpha\pi & -\sin\alpha\pi & \cos\alpha\pi & \sin\alpha\pi \\ \mu\sin\alpha\pi & -\mu\cos\alpha\pi & -\sin\alpha\pi & \cos\alpha\pi \end{bmatrix} \begin{bmatrix} \lambda_c^+ \\ \lambda_s^+ \\ \lambda_c^- \\ \lambda_s^- \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

For this homogeneous system of linear equations to have a nontrivial solution, its determinant must vanish, which leads to

$$\frac{1}{2}(\mu + \frac{1}{\mu})(\cos \pi\alpha - \cos 2\pi\alpha) + 2 - \cos \pi\alpha - \cos 2\pi\alpha = 0 \,. \tag{22}$$

The exponent α that determines the degree of the singularity apparently depends on the size of the jump μ. It can be shown that (22) always has a unique solution $\alpha \in (1/2, 1)$. For $\mu \to 1$, i.e., as the jump disappears, we have $\alpha \to 1$, i.e., the singularity disappears as well. For $\mu \to 0$ or $\mu \to \infty$, α tends to $2/3$, which is exactly the value obtained for a reentrant corner with exterior angle $\pi/2$. It is straightforward to extend the procedure outlined above to any number of adjoining subdomains and any size of angles (cf. [8]). We therefore have a computational technique to compute the shape of the singularity at interface corners and cross-points where two interfaces intersect. This technique will be fundamental for the finite element approach described in the next section.

Finite Element Approximation

The minimum of $G(\mathbf{u}, p; f)$ is approximated using a Rayleigh-Ritz finite element method. Let $\mathcal{T}^h$ be a triangulation of Ω, which we assume to be quasi-uniform (cf. [3, Chapter 4]), and let $\mathbf{W}^h$ and V^h be appropriate finite-dimensional spaces. The interface is required to be the union of edges of the triangulation. If the interface is cutting through elements of the triangulation, then special techniques have to be considered in order to average the parameters properly, which complicates the whole approach. We do not address this task or the problems associated with it here, but instead assume that the interfaces are restricted to edges of the triangulation. For the sake of exposition, we also assume that each segment of the interface curves is parallel to one of the coordinate axes. It is easy to see that the following development of the finite element approach can be generalized to isoparametric elements, where the interface curves are logically aligned with coordinate axes.

It is desirable, in general, to use conforming finite elements, where the finite-dimensional spaces satisfy $\mathbf{W}^h \subset \mathbf{W}$ and $V^h \subset V$. Along straight segments of the interface curve, this can be accomplished by enforcing condition (18) on the finite element basis functions. Using bilinear finite elements on rectangles, for example, a basis function for u_1 at a node on a horizontal interface segment is continuous in the x_1-direction and has a jump of size a^+/a^- in the x_2-direction. Such a basis function for u_1 at a node on a vertical interface segment is continuous (in both coordinate directions). Under the assumption that all the interface curves are straight lines which do not intersect each other (we will address the case of interface corners or cross-points later), we can therefore construct piecewise bilinear finite element spaces:

$$V^h = \{q \in V : q|_T \text{ bilinear on } T \text{ for all } T \in \mathcal{T}^h\}$$
$$\mathbf{W}^h = \{\mathbf{v} \in H(\mathrm{div}, \Omega) \cap H(\mathrm{curl}\, a, \Omega) : v_i|_T \text{ bilinear on } T \text{ for all } T \in \mathcal{T}^h\} \,.$$

The finite element approximation $(\mathbf{u}^h, p^h) \in \mathbf{W}^h \times V^h$ is then defined as the solution of the minimization problem

$$G(\mathbf{u}^h, p^h; f) = \min_{(\mathbf{v}^h, q^h) \in \mathbf{W}^h \times V^h} G(\mathbf{v}^h, q^h; f) \,. \tag{23}$$

One of the main practical advantages of the least-squares finite element approach over other variational formulations consists in the fact that the minimum of the functional constitutes an a posteriori error measure. This follows from the general relation between the least-squares functional and corresponding bilinear form. The main point here is

the fact that the least-squares functional is zero at the solution $(\mathbf{u}, p)$, which leads to

$$\begin{aligned} & G(\mathbf{u}^h, p^h; f) \\ = \; & G(\mathbf{u}^h, p^h; f) - G(\mathbf{u}, p; f) \\ = \; & \mathcal{F}(\mathbf{u}^h, p^h; \mathbf{u}^h, p^h) + 2(f, \nabla \cdot \mathbf{u}^h)_{0,\Omega} - \mathcal{F}^h(\mathbf{u}, p; \mathbf{u}, p) - 2(f, \nabla \cdot \mathbf{u})_{0,\Omega} \\ = \; & \mathcal{F}^h(\mathbf{u} - \mathbf{u}^h, p - p^h; \mathbf{u} - \mathbf{u}^h, p - p^h) \; . \end{aligned}$$

Under the above assumptions, we get the following convergence result for the finite element approximation.

Theorem 2 *Assume that for* $(\mathbf{u}, p)$, *the solution of* (10), *we have* $(\mathbf{u}, p)|_{\Omega_i} \in (H^{1+\delta}(\Omega_i))^3$ *for some* $\delta \in (0, 1]$ *and for* $i = 1, \ldots, J$. *Let* $(\mathbf{u}^h, p^h) \in \mathbf{W}^h \times V^h$ *be the solution of* (23). *Then*

$$|||(\mathbf{u}, p) - (\mathbf{u}^h, p^h)||| \leq C h^\delta \sum_{i=1}^{J} (\|\mathbf{u}\|_{1+\delta,\Omega_i} + \|\sqrt{\omega_i} p\|_{1+\delta,\Omega_i}) \tag{24}$$

where the constant C *is independent of* h *and of the size of the jumps in* $\{\omega_i\}$.

Proof. From Theorem 1 and Cea's Lemma (see, for example, [3, Theorem 2.8.1]), we obtain

$$|||(\mathbf{u}, p) - (\mathbf{u}^h, p^h)||| \leq \frac{\gamma_2}{\gamma_1} \min_{(\mathbf{v}^h, q^h) \in \mathbf{W}^h \times V^h} |||(\mathbf{u}, p) - (\mathbf{v}^h, q^h)||| \; .$$

Moreover, for $(\mathbf{v}, q) \in \mathbf{W} \times V$, we have

$$\begin{aligned} |||(\mathbf{v}, q)|||^2 &= \|\nabla \cdot \mathbf{v}\|^2_{0,\Omega} + \|a \nabla \times (\mathbf{v}/a)\|^2_{0,\Omega} + \|\mathbf{v}/\sqrt{a}\|^2_{0,\Omega} + \|\sqrt{a} \nabla q\|^2_{0,\Omega} \\ &= \sum_{i=1}^{J} (\|\nabla \cdot \mathbf{v}\|^2_{0,\Omega_i} + \|a \nabla \times (\mathbf{v}/a)\|^2_{0,\Omega_i} + \|\mathbf{v}/\sqrt{a}\|^2_{0,\Omega_i} + \|\sqrt{a} \nabla q\|^2_{0,\Omega_i}) \\ &\leq c_1 \sum_{i=1}^{J} (\|\nabla \cdot \mathbf{v}\|^2_{0,\Omega_i} + \|\nabla \times \mathbf{v}\|^2_{0,\Omega_i} + \|\mathbf{v}/\sqrt{a}\|^2_{0,\Omega_i} + \|\sqrt{a} \nabla q\|^2_{0,\Omega_i}) \; . \end{aligned}$$

Since by assumption $\mathbf{u}|_{\Omega_i} \in H^1(\Omega_i)$ and, similarly, $\mathbf{v}^h|_{\Omega_i} \in H^1(\Omega_i)$ for each $\mathbf{v}^h \in \mathbf{W}^h$, then for $i = 1, \ldots, J$ we have

$$\|\nabla \cdot (\mathbf{u} - \mathbf{v}^h)\|^2_{0,\Omega_i} + \|\nabla \times (\mathbf{u} - \mathbf{v}^h)\|^2_{0,\Omega_i} \leq c_2 |\mathbf{u} - \mathbf{v}^h|^2_{1,\Omega_i} \; .$$

This leads to

$$|||(\mathbf{u}, p) - (\mathbf{v}^h, q^h)||| \leq c_3 \sum_{i=1}^{J} (|\mathbf{u} - \mathbf{v}^h|_{1,\Omega_i} + \|(\mathbf{u} - \mathbf{v}^h)/\sqrt{\omega_i}\|_{0,\Omega_i} + \|\sqrt{\omega_i}(p - q^h)\|_{1,\Omega_i}) \; .$$

Standard interpolation properties of piecewise bilinear functions (see, for example, [3, Theorems 12.3.3 and 12.3.12]) lead to

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}^h\|_{1,\Omega_i} &\leq c_4 h^\delta \|\mathbf{u}\|_{1+\delta,\Omega_i} \\ \|p - q^h\|_{1,\Omega_i} &\leq c_5 h^\delta \|p\|_{1+\delta,\Omega_i} \end{aligned}$$

which completes the proof. ∎

If the interface curve is not a straight line, or, more generally, not sufficiently smooth, then the finite element approximation becomes excessively more complicated. In the preceding section we saw that, for the solution $(\mathbf{u}, p)$ of (10), $\mathbf{u}$ has the singular behavior shown in (20) and (21). It is easy to see that this implies $\mathbf{u}|_{\Omega_i} \notin (H^1(\Omega_i))^2$ for all

subregions Ω_i adjacent to the interface corner, and therefore the standard finite element approximation results do not apply.

Moreover, in order to have $\mathbf{u}^h \in H(\mathrm{div},\Omega) \cap H(\mathrm{curl}\, a,\Omega)$ in the neighborhood of an interface corner, it is necessary and sufficient to require $\mathbf{u}^h$ to have the form of (20) and (21). In other words, in order to have conforming finite elements, we must include a singular basis function at each interface corner (or cross-point): The tools developed in the previous section allow us, in principle, to compute the exact shape of such a singularity. Multiplied by a standard piecewise bilinear function, such a singular function could then serve as a basis function at that point. A procedure of this type is described in [14, Section 8.2] along with special techniques to solve the resulting discrete system. However, this approach requires special stencils for these singular points, which complicates the overall finite element approach. Instead, we consider an alternative nonconforming finite element method, based on simple basis functions like bilinears on rectangles.

We construct $\mathbf{W}^h$ observing the fact that, for the right-hand side in (11) to be defined, we must have $\mathbf{W}^h \subset H(\mathrm{div},\Omega)$. This implies that, for $\mathbf{u}^h \in \mathbf{W}^h$, $\mathbf{n}\cdot\mathbf{u}^h$ must be continuous across all interfaces. Now consider the bilinear finite element basis function associated with the interface corner in Figure 1. For $\mathbf{u}^h \in \mathbf{W}^h \subset H(\mathrm{div},\Omega)$, we must require that u_1 is continuous in the x_1-direction across the horizontal portion of the interface; that u_2 is continuous in the x_2-direction across the vertical portion of the interface; and that both u_1, u_2 are continuous elsewhere. From (18) we see that $\mathbf{u} \in H(\mathrm{curl}\, a,\Omega)$ requires u_2 to have a jump across the vertical portion of the interface, while u_1 must have a jump across the horizontal portion. This causes a conflict at the corner. The finite-dimensional space $\mathbf{W}^h$ will, therefore, not be contained in $H(\mathrm{curl}\, a,\Omega)$, in general, and $\mathbf{W}^h \times V^h \not\subset \mathbf{W}\times V$. In particular, the bilinear form $\mathcal{F}(\cdot,\cdot;\cdot,\cdot)$ is not defined on $\mathbf{W}^h \times V^h$. For $\mathbf{u},\mathbf{v} \in \mathbf{W}+\mathbf{W}^h$ and $p,q \in V+V^h$, we define a modified least-squares bilinear form by

$$\begin{aligned}\mathcal{F}^h(\mathbf{u},p;\mathbf{v},q) &= (\mathbf{u}/\sqrt{a}-\sqrt{a}\nabla p, \mathbf{v}/\sqrt{a}-\sqrt{a}\nabla q)_{0,\Omega}\\ &+(\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})_{0,\Omega}+\textstyle\sum_{i=1}^{J}(\nabla\times\mathbf{u},\nabla\times\mathbf{v})_{0,\Omega_i}\,.\end{aligned} \tag{25}$$

On $\mathbf{W}\times V$, this bilinear form coincides with $\mathcal{F}(\cdot,\cdot;\cdot,\cdot)$. The least-squares functional corresponding to $\mathcal{F}^h(\cdot,\cdot;\cdot,\cdot)$ is

$$G^h(\mathbf{u},p;f) = \|\mathbf{u}/\sqrt{a}-\sqrt{a}\nabla p\|^2_{0,\Omega}+\|\nabla\cdot\mathbf{u}+f\|^2_{0,\Omega}+\sum_{i=1}^{J}\|\nabla\times\mathbf{u}\|^2_{0,\Omega_i}\,. \tag{26}$$

Let $(\mathbf{u},p)\in\mathbf{W}\times V$ be the solution of (10), and let $(\mathbf{u}^h,p^h)\in\mathbf{W}^h\times V^h$ be defined by

$$G^h(\mathbf{u}^h,p^h;f) = \min_{(\mathbf{v}^h,q^h)\in\mathbf{W}^h\times V^h} G^h(\mathbf{v}^h,q^h;f)\,. \tag{27}$$

Recall that, at an interface corner, $\mathbf{u}$ has a singularity of the form given in (20) and (21). This implies that we cannot expect to approximate $\mathbf{u}$ to the same accuracy by standard finite elements near a singularity as elsewhere in Ω. Moreover, since our finite element subspace $\mathbf{W}^h\times V^h$ is not contained in the space $\mathbf{W}\times V$ in which we have shown ellipticity, the relatively large error near a singularity will deterioriate the finite element approximation in the entire region. This phenomenon is reflected by the fact that, in the presence of singularities, $G^h(\mathbf{u}^h,p^h;f)$ does not decrease as h is made smaller. We will observe this behavior later in our computational experiments. It is therefore necessary to introduce a weight function which decreases near the singular point. The proper choice of weighting is motivated by the form of the singularity.

In particular, (19), (20), and (21) imply $\nabla\mathbf{u}\sim r^{\alpha-2}$ in the neighborhood of the singularity. If T_s^h denotes an element of the triangulation $\mathcal{T}^h$ such that the interface corner appears as one of its vertices, then

$$\|r^{2-\alpha}\nabla(u_j-u_j^h)\|_{0,T_s^h} = O(h^2)\,.$$

If the right-hand side f and the restriction of a to Ω_i are sufficiently smooth, then we know that $\mathbf{u} \in (H^2_{\mathrm{loc}}(\Omega_i))^2$, i.e., $\mathbf{u} \in (H^2(\tilde{\Omega}))^2$ for any compact $\tilde{\Omega} \subset \Omega_i$. This implies that $\tilde{\mathbf{v}}^h \in \mathbf{W}^h$ exists such that

$$\|\nabla \cdot (\mathbf{u} - \tilde{\mathbf{v}}^h)\|_{0,\tilde{\Omega}} = O(h^2) .$$

The other terms in (26) can be treated in a similar way, which motivates the definition of the weighted least-squares functional

$$\begin{aligned} G^h_w(\mathbf{u}, p; f) &= \|\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p\|^2_{0,h,1-\alpha,\Omega} \\ &+ \|\nabla \cdot \mathbf{u} + f\|^2_{0,h,2-\alpha,\Omega} + \sum_{i=1}^{J} \|\nabla \times \mathbf{u}\|^2_{0,h,2-\alpha,\Omega_i} \end{aligned} \tag{28}$$

and corresponding bilinear form

$$\begin{aligned} \mathcal{F}^h_w(\mathbf{u}, p; \mathbf{v}, q) &= (\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p), \mathbf{v}/\sqrt{a} - \sqrt{a}\nabla q))_{0,h,1-\alpha,\Omega} \\ &+ (\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v})_{0,h,2-\alpha,\Omega} + \sum_{i=1}^{J} (\nabla \times \mathbf{u}), \nabla \times \mathbf{v})_{0,h,2-\alpha,\Omega_i} . \end{aligned} \tag{29}$$

The inner product $(\cdot,\cdot)_{0,h,\beta,\Omega}$ is defined as

$$(\mathbf{v}, \mathbf{w})_{0,h,\beta,\Omega} = (w^{h,\beta}\mathbf{v}, w^{h,\beta}\mathbf{w})_{0,\Omega}$$

with the weight function $w^{h,\beta}$ constructed in the following way: Consider a sequence of triangulations $\{\mathcal{T}^{h_l}, l = 0, \ldots, L\}$, with $H = h_0 \geq h_1 \geq \cdots \geq h_L = h$. Let $\Omega^{h_l}_s$ denote the union of of all elements $T^{h_l} \in \mathcal{T}^{h_l}$ with the singular point as one of their vertices. The weight function $w^{h,\beta}$ is defined as

$$w^{h,\beta}(\mathbf{x}) = \begin{cases} h^\beta \text{ for } \mathbf{x} \in \Omega^h_s , \\ h_l^\beta \text{ for } \mathbf{x} \in \Omega^{h_{l-1}}_s \backslash \Omega^{h_l}_s , \; l = 1, \ldots, L , \\ 1 \text{ for } \mathbf{x} \in \Omega \backslash \Omega^{h_0}_s . \end{cases} \tag{30}$$

Let $(\mathbf{u}^h_w, p^h_w) \in \mathbf{W}^h \times V^h$ be the solution of

$$G^h_w(\mathbf{u}^h_w, p^h_w; f) = \min_{(\mathbf{v}^h, q^h) \in \mathbf{W}^h \times V^h} G^h_w(\mathbf{v}^h, q^h; f) . \tag{31}$$

In the final section of this paper we will demonstrate, by means of numerical results, that the weighted functional $G^h_w(\mathbf{u}^h_w, p^h_w; f)$ actually decreases regularly as the triangulation is refined. Note, however, that this does not mean that the error $\mathbf{u} - \mathbf{u}^h_w$ is small throughout the region Ω. In particular, the pointwise accuracy usually deteriorates near singularities. This suggests that the weighted functional should be combined with local refinement techniques to guarantee satisfactory resolution in the entire region. Multilevel refinement techniques are especially effective in this context.

Multilevel Algorithms

Consider the sequence of triangulations $\{\mathcal{T}^{h_l}, l = 0, \ldots, L\}$ introduced earlier. Associated with each triangulation $\{\mathcal{T}^{h_l}\}$ is the finite element space $\mathbf{W}^{h_l} \times V^{h_l}$, which we may also denote by $\mathbf{W}_l \times V_l$. This leads to a nested sequence of spaces

$$\mathbf{W}_0 \times V_0 \subset \mathbf{W}_1 \times V_1 \subset \cdots \subset \mathbf{W}_L \times V_L = \mathbf{W}^h \times V^h .$$

On each level l, $0 \leq l \leq L$, an operator $\mathcal{F}_l : \mathbf{W}_l \times V_l \to \mathbf{W}_l \times V_l$ is defined by

$$\langle\langle \mathcal{F}_l(\mathbf{u}, p); (\mathbf{v}, q) \rangle\rangle = \mathcal{F}(\mathbf{u}, p; \mathbf{v}, q) \text{ for all } (\mathbf{v}, q) \in \mathbf{W}_l \times V_l ,$$

where the inner product $\langle\langle \cdot; \cdot \rangle\rangle$ is given by

$$\langle\langle (\mathbf{u}, p); (\mathbf{v}, q) \rangle\rangle = (\mathbf{u}, \mathbf{v})_{0,\Omega} + (\sqrt{a}\, p, \sqrt{a}\, q)_{0,\Omega} .$$

In terms of the operator $\mathcal{F}_l$, the discrete problem (23) can be written as

$$\mathcal{F}_l(\mathbf{u}_l, p_l) = F_l \tag{32}$$

where the right-hand side is defined by $\langle\langle F_l, (\mathbf{v}, q)\rangle\rangle = -(f, \nabla\cdot\mathbf{v})_{0,\Omega}$ for all $(\mathbf{v}, q) \in \mathbf{W}_l \times V_l$. For the solution of (32), it is natural to use an iterative method since this requires only a computational procedure for the action of the operator $\mathcal{F}_l$ for $l = 0, \ldots, L$. The cost for one call of such a procedure is proportional to the number of unknowns $N = O(h^{-2})$.

The conjugate gradient method (cf. [13, Section 8.7]) computes its iterates $(\mathbf{u}_l^{(n)}, p_l^{(n)}) \in \mathbf{W}_l \times V_l$ in the Krylov subspace

$$\mathcal{K}_n(F_l, \mathcal{F}_l) = \text{span}\{F_l, \mathcal{F}_l F_l, \ldots, \mathcal{F}_l^{n-1} F_l\}$$

according to the minimization property

$$G(\mathbf{u}_l^{(n)}, p_l^{(n)}; f) = \min_{(\mathbf{v}_l, q_l)\in\mathcal{K}_n(F_l, \mathcal{F}_l)} G(\mathbf{v}_l, q_l; f)\,.$$

Since the condition number of $\mathcal{F}_l$ is proportional to $O(h_l^{-2})$ (cf. [5, Theorem 3.2]), the number of conjugate gradient iterations required to achieve a certain accuracy grows like $O(h_l^{-1})$ (cf. [13, Section 8.7]). The overall computational complexity to solve a discrete problem on $\mathcal{T}^{h_l}$ using the conjugate gradient method therefore grows like $O(h_l^{-3})$.

Optimal computational complexity, $O(h_l^{-2})$, can be achieved, under certain assumptions on $\mathcal{F}((\cdot,\cdot);(\cdot,\cdot))$, by a full multigrid algorithm. The basic ingredients for multilevel methods are the projection operators $\mathcal{P}_l, \mathcal{Q}_l : \mathbf{W}^h \times V^h \to \mathbf{W}_l \times V_l$ which are given by

$$\mathcal{F}(\mathcal{P}_l(\mathbf{u}, p); (\mathbf{v}, q)) = \mathcal{F}((\mathbf{u}, p); (\mathbf{v}, q)) \text{ for all } (\mathbf{v}, q) \in \mathbf{W}_l \times V_l$$

and

$$\langle\langle \mathcal{Q}_l(\mathbf{u}, p); (\mathbf{v}, q)\rangle\rangle = \langle\langle (\mathbf{u}, p); (\mathbf{v}, q)\rangle\rangle \text{ for all } (\mathbf{v}, q) \in \mathbf{W}_l \times V_l$$

and smoothing operators $\mathcal{R}_l : \mathbf{W}_l \times V_l \to \mathbf{W}_l \times V_l$ representing iterations on level l. With these tools, standard multilevel algorithms can be constructed (see [5, Section 4] for further details). A detailed study of the convergence properties of multilevel methods for first-order system least-squares applied to problems with discontinuous coefficients will be given in [9].

Computational Experiments

In our examples, we consider (3) on the unit square $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_1, x_2 < 1\}$, with $f \equiv 1$ and $\Gamma_D = \partial\Omega$. We show the results of two sets of experiments, one with a smooth interface curve and the other with an interface corner causing a singularity in $\mathbf{u}$.

Example 1. In this example, the interface curve is a straight line, so no singularity occurs. We consider

$$a(x_1, x_2) = \begin{cases} a^+\,, & 0 < x_2 < 0.5\,, \\ a^-\,, & 0.5 < x_2 < 1\,, \end{cases} \tag{33}$$

with different choices for the values for a^+ and a^-. The solution shown in Figure 3 was obtained for $a^+ = 10$ and $a^- = 0.1$.

The computational results shown in Table 1 indicate that the approximation of the solution improves nicely as the triangulation is refined, independently of the size of the jumps. The reduction factor displayed in parentheses is the ratio of the minimum values on the current and next coarser level. Note that they do not quite reach 0.25, which is due to the lack of regularity at the corners of the subdomains. In fact, due to the corners

Table 1: Example 1: Minimum value (reduction factor) of the functional G^h

a^+/a^-	1	10	10^2	10^4
h				
1/8	$2.42 \cdot 10^{-2}$	$3.50 \cdot 10^{-2}$	$4.13 \cdot 10^{-2}$	$7.81 \cdot 10^{-2}$
1/16	$7.18 \cdot 10^{-3}$ (0.30)	$1.07 \cdot 10^{-2}$ (0.31)	$1.26 \cdot 10^{-2}$ (0.31)	$2.30 \cdot 10^{-2}$ (0.29)
1/32	$2.08 \cdot 10^{-3}$ (0.29)	$3.14 \cdot 10^{-3}$ (0.29)	$3.71 \cdot 10^{-3}$ (0.29)	$6.41 \cdot 10^{-3}$ (0.28)
1/64	$5.92 \cdot 10^{-4}$ (0.28)	$9.05 \cdot 10^{-4}$ (0.29)	$1.07 \cdot 10^{-3}$ (0.29)	$1.75 \cdot 10^{-3}$ (0.27)

with interior angle $\pi/2$, we have neither $\mathbf{u} \in (H^2(\Omega^+))^2$ nor $\mathbf{u} \in (H^2(\Omega^-))^2$. Consequently, the finite element approximation deteriorates near these corners. In contrast to the situation at singularities, however, this behavior does not contaminate the solution elsewhere since the basis functions corresponding to these points are conforming.

Example 2. This example shows results for a problem with a singularity in $\mathbf{u}$. We choose

$$a(x_1, x_2) = \begin{cases} a^+, & 0 < x_1, x_2 < 0.5\,, \\ a^-, & \text{elsewhere} \end{cases} \tag{34}$$

(see Figure 1) with different choices for the values for a^+ and a^- (again with $a^+ = 10$ and $a^- = 0.1$ for the solution shown in Figure 4).

The exponents for this example with the three values for the coefficient jumps used in Table 2 are given by $\alpha = 0.7317, 0.6739$, and 0.6667, respectively. Note that the last number is very close to the value $\alpha = 2/3$ that one gets for a reentrant corner with interior angle $3/2\pi$. Using the weighting described earlier with $H = 1/8$ leads to the results listed in Table 2. The modified least-squares functional is again reduced nicely and regularly as the triangulation is refined. Note that using the weighted functional means that the pointwise approximation deteriorates close to the singular point, where local refinement can be used if a better pointwise resolution is needed.

Table 2: Example 2: Minimum value (reduction factor) of the weighted functional G_w^h

a^+/a^-	1	10	10^2	10^4
h				
1/8	$2.42 \cdot 10^{-2}$	$3.74 \cdot 10^{-2}$	$5.17 \cdot 10^{-2}$	$1.20 \cdot 10^{-1}$
1/16	$7.18 \cdot 10^{-3}$ (0.30)	$1.16 \cdot 10^{-2}$ (0.31)	$1.58 \cdot 10^{-2}$ (0.31)	$3.53 \cdot 10^{-2}$ (0.29)
1/32	$2.08 \cdot 10^{-3}$ (0.29)	$3.43 \cdot 10^{-3}$ (0.30)	$4.66 \cdot 10^{-3}$ (0.29)	$9.84 \cdot 10^{-3}$ (0.28)
1/64	$5.92 \cdot 10^{-4}$ (0.28)	$9.95 \cdot 10^{-4}$ (0.29)	$1.34 \cdot 10^{-3}$ (0.29)	$2.68 \cdot 10^{-3}$ (0.27)

Table 3: Example 2: Minimum value of the functional G^h

a^+/a^-	1	10	10^2	10^4
h				
1/8	$2.42 \cdot 10^{-2}$	$4.36 \cdot 10^{-2}$	$7.50 \cdot 10^{-2}$	$1.62 \cdot 10^{-1}$
1/16	$7.18 \cdot 10^{-3}$	$2.39 \cdot 10^{-2}$	$5.49 \cdot 10^{-2}$	$9.89 \cdot 10^{-2}$
1/32	$2.08 \cdot 10^{-3}$	$2.07 \cdot 10^{-2}$	$5.35 \cdot 10^{-2}$	$8.86 \cdot 10^{-2}$
1/64	$5.92 \cdot 10^{-4}$	$2.22 \cdot 10^{-2}$	$5.66 \cdot 10^{-2}$	$9.33 \cdot 10^{-2}$

In order to illustrate the necessity of modifying the functional in the neighborhood of a singular point, we also computed the results for the unmodified functional G^h instead of G_ω^h. The numbers in Table 3 show that this functional is not satisfactorily reduced in the course of refining the triangulation. Our numerical tests have shown that minimizing the unmodified functional leads to poor finite element approximations. Figure 2 shows the error with respect to the exact solution for p for the weighted

functional and for the unmodified functional. Obviously, for the unmodified functional, the resulting error between the discrete and exact solution is relatively large in the entire domain. This behavior seems to indicate that using the unmodified functional has the effect of trying too hard to satisfy the first-order system (6) close to the singularity, where it is impossible to get a good approximation with bilinear finite elements. For the weighted functional, however, the error is smaller and mainly occurs in a rather small neighborhood of the singular point.

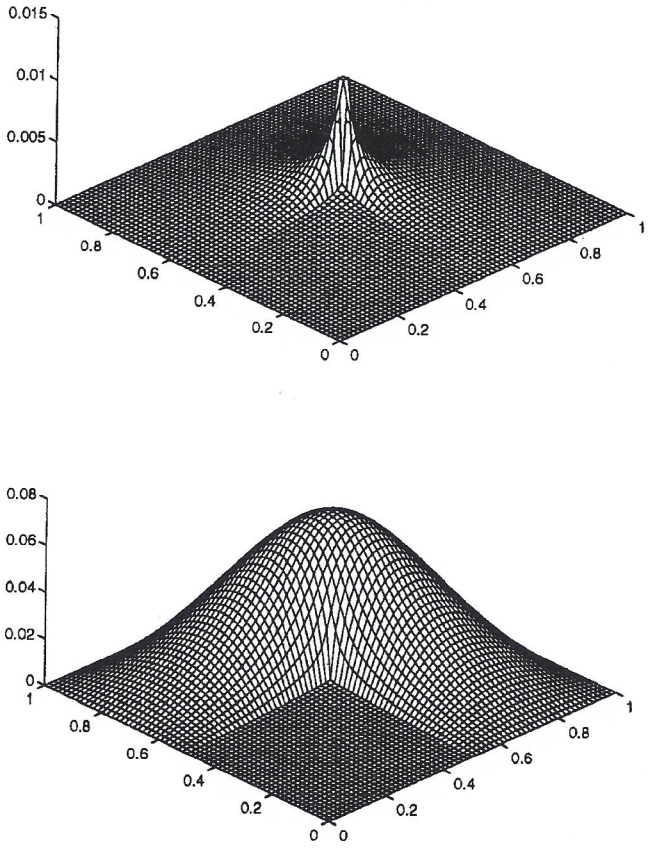

Figure 2: Example 2: Error in the pressure p for the weighted functional G_w^h (top) and the unmodified functional G^h (bottom)

REFERENCES

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

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

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

[4] 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.*, 31:1785–1799, 1994.

[5] 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.*, 1995. To Appear.

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

[7] M. J. Holst. *Multilevel Methods for the Poisson-Boltzmann Equation*. Ph.D. thesis, University of Illinois at Urbana-Champaign, 1993.

[8] O. E. Lafe, J. S. Montes, A. H. D. Cheng, J. A. Liggett, and P. L.-F. Liu. Singularities in Darcy flow through porous media. *J. ASCE Hydr. Div.*, 106:977–997, 1980.

[9] T. A. Manteuffel, S. F. McCormick, and G. Starke. Analysis of first-order system least-squares for elliptic problems with discontinuous coefficients. In Preparation.

[10] A. I. Pehlivanov and G. F. Carey. Error estimates for least-squares mixed finite elements. *RAIRO MMNA*, 28:499–516, 1994.

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

[12] T. F. Russell and M. F. Wheeler. Finite element and finite difference methods for continuous flows in porous media. In R. E. Ewing, editor, *The Mathematics of Reservoir Simulation*, chapter II. SIAM, 1983.

[13] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. 2nd edition. Springer, New York, 1993.

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

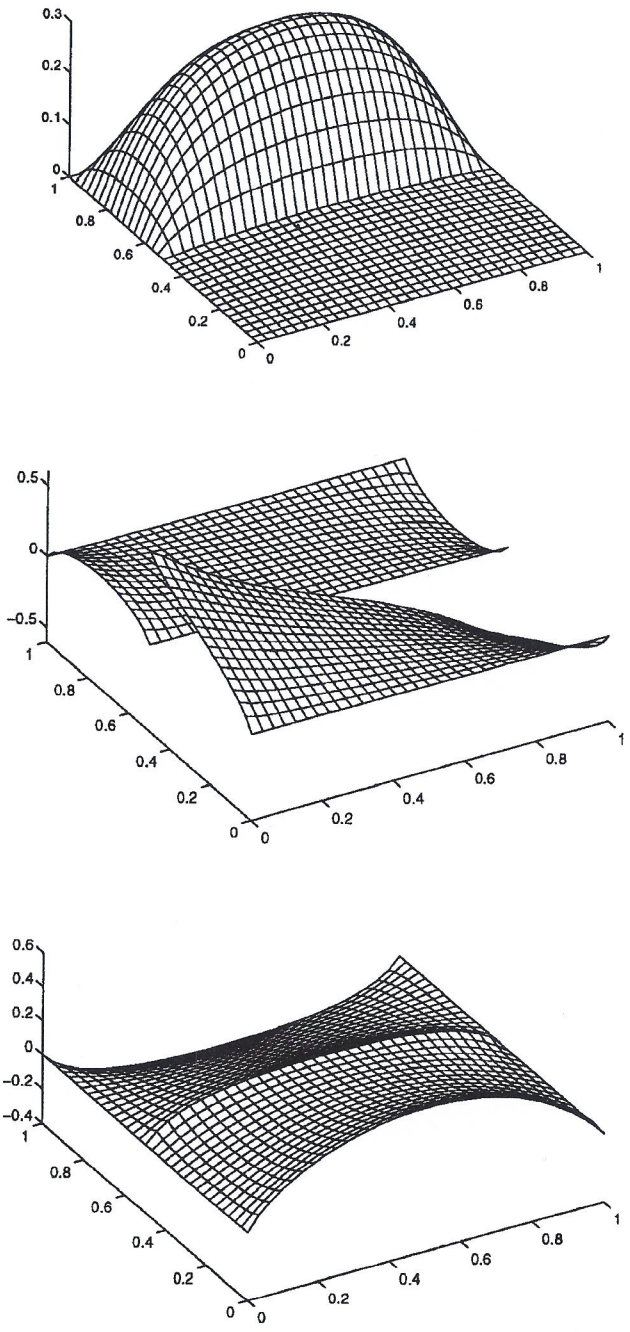

Figure 3: Example 1: Pressure p (top) and flux components u_1 and u_2

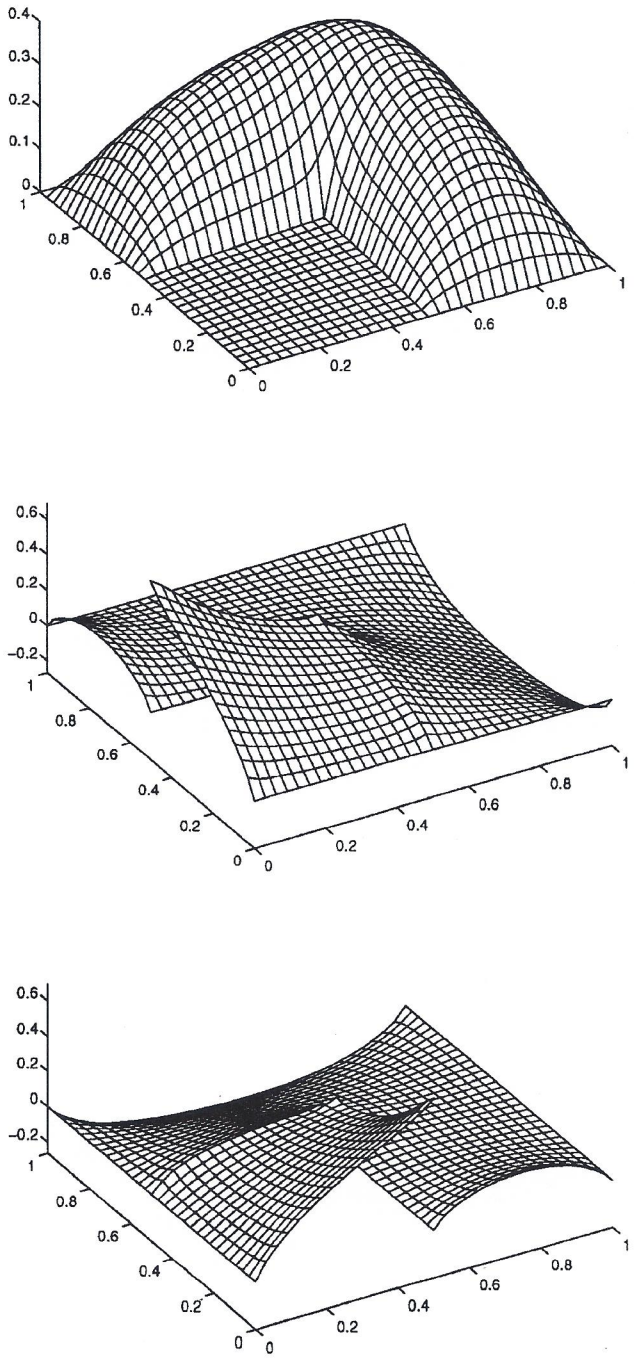

Figure 4: Example 2: Pressure p (top) and flux components u_1 and u_2