SCATTERING OF TIME-HARMONIC ELECTROMAGNETIC WAVES BY ANISOTROPIC INHOMOGENEOUS SCATTERERS OR IMPENETRABLE OBSTACLES

PETER MONK* AND JOE COYLE†

Abstract. We investigate an overlapping solution technique to compute the scattering of time-harmonic electromagnetic waves in two dimensions. The technique can be used to compute waves scattered by penetrable anisotropic inhomogeneous scatterers or impenetrable obstacles. The major focus is on implementing the method using finite elements. We prove existence of a unique solution to the disctretized problem and derive an optimal convergence rate for the scheme, which is verified numerical by examples.

Key words. Anisotropic, finite element, electromagnetic scattering

AMS subject classifications. 35Q60, 65N30

1. Introduction. We consider the scattering of time-harmonic electromagnetic waves by a penetrable anisotropic medium of compact support or by a bounded impenetrable obstacle. Scattering by anisotropic media appears in various medical applications since the body, as a medium, is anisotropic (Colton and Monk [9]). The aim is to obtain the scattered field given a known incident field and sufficient knowledge of the scatterer and background in which it is contained. The scattered field propagates in an unbounded region, which poses a problem when discretizing the equations and numerically computing the field.

The first step in overcoming this difficulty is usually to introduce an artificial boundary containing the scatterer. On this boundary, the Dirichlet-to-Neumann map (Keller and Givoli [12], Masmoudi [21]) then provides a non-local boundary condition accounting for the infinite domain. There are a variety of ways of implementing these non-local conditions. For example, boundary integral equations, leading to weakly singular integrals, can be used to approximate the Dirichlet-to-Neumann map (Chen and Zhou [4], Hsiao [15, 16], Kirsch and Monk [20]). Alternatively, on a simple auxiliary boundary, it is possible to use special function series (Keller and Givoli [12], Kirsch and Monk [19]). An alternate approach is to approximate the Dirichlet-to-Neumann map using local differential operators to obtain a local absorbing condition. This is a very popular approach (Engquist and Majda [10], Stupfel and Mittra [26], Jin [18]), but the accuracy of such boundary conditions is difficult to assess.

A new class of methods due to Bérenger [1] also deserves mention. This method perturbs the differential equation in a layer (the "perfectly matched layer") near the artificial boundary to absorb scattering solutions and prevent unphysical reflections from the artificial boundary. Although very effective, this technique is limited to rather simple convex artificial boundaries (Collino and Monk [7], Chew and Teixeria [5]).

In this paper, we use a technique first suggested by Jami and Lenoir [17] and used extensively by Hazard and Lenoir [13]. This method is based on an overlapping solution technique, which we shall discuss shortly. It has the advantage of guaranteed

*Department of Mathematical Sciences, University of Delaware, Newark, DE, 19716, USA (monk@math.udel.edu).

†Department of Mathematical Sciences, University of Delaware, Newark, DE, 19716, USA (cassidy@math.udel.edu)

accuracy (shared with the "exact" methods coupling integral equations or series solution methods discussed previously), but without the need to evaluate weakly singular integrals. In addition, the artificial boundary can be of arbitrary shape. The major drawbacks are that the resulting matrix is general (not even complex symmetric), and the coupling procedure reduces the sparsity of the matrix as we shall discuss in §5.2. We are not aware of any existing discretization error study for methods of this type. Deriving such a result as well as demonstrating it numerically are the major results of this paper.

2. Setting up the problem. Consider the scattering of electromagnetic waves from an infinitely long cylinder containing an anisotropic inhomogeneous medium. Denote by ε and μ the electric permittivity and magnetic permeability. The electric and magnetic fields, denoted $\hat{\mathbf{E}}$ and $\hat{\mathbf{H}}$, satisfy the following Maxwell equations:

$$\varepsilon\frac{\partial\hat{\mathbf{E}}}{\partial t}+\sigma\hat{\mathbf{E}}-\nabla\times\hat{\mathbf{H}}=0,\qquad \mu\frac{\partial\hat{\mathbf{H}}}{\partial t}+\nabla\times\hat{\mathbf{E}}=0.$$

We are interested in finding the solution at a fixed frequency ω. Let ε_0 and μ_0 denote the electric permittivity and magnetic permeability of free space. Define the wave number $k=\omega\sqrt{\varepsilon_0\mu_0}$ and the index of refraction

$$N(\boldsymbol{x})=\frac{1}{\varepsilon_0}\left(\varepsilon(\boldsymbol{x})+i\frac{\sigma(\boldsymbol{x})}{\omega}\right),$$

and consider the special case of an anisotropic medium: an orthotropic medium. We then have

$$\varepsilon(\boldsymbol{x})=\begin{pmatrix}\varepsilon_{11}(\boldsymbol{x}) & \varepsilon_{12}(\boldsymbol{x}) & 0\\ \varepsilon_{21}(\boldsymbol{x}) & \varepsilon_{22}(\boldsymbol{x}) & 0\\ 0 & 0 & \varepsilon_{33}(\boldsymbol{x})\end{pmatrix}.$$

We also assume that $\sigma(\boldsymbol{x})$ and $\mu(\boldsymbol{x})$ have the same form as ε and that $\boldsymbol{x}=(x,y)$, so these quantities are independent of z. Also define $n(\boldsymbol{x})=\frac{\mu_{33}}{\mu_0}$. For a fixed frequency, the time-harmonic electric and magnetic fields can be written $\hat{\mathbf{E}}(\boldsymbol{x},t)=\varepsilon_o{}^{-\frac{1}{2}}\mathbf{E}(\boldsymbol{x})e^{-i\omega t}$ and $\hat{\mathbf{H}}(\boldsymbol{x},t)=\mu_o{}^{-\frac{1}{2}}\mathbf{H}(\boldsymbol{x})e^{-i\omega t}$ so that

$$\nabla\times\mathbf{E}-ikn\mathbf{H}=0,\qquad \vec{\nabla}\times\mathbf{H}+ikN\mathbf{E}=0, \tag{2.1}$$

where $\vec{\nabla}$ is the vector curl of a scalar. Here, $\mathbf{E}$ and $\mathbf{H}$ are assumed to be independent of z (with $\mathbf{H}$ perpendicular to the xy-plane):

$$\mathbf{E}=\begin{pmatrix}E_1(x,y)\\ E_2(x,y)\\ 0\end{pmatrix},\qquad \mathbf{H}=\begin{pmatrix}0\\ 0\\ H_3(x,y)\end{pmatrix}.$$

Under these assumptions, the Maxwell system (2.1) reduces to solving the following general Helmholtz equation for $u=H_3(x,y)$:

$$\nabla\cdot\mathcal{A}\nabla u+k^2n(\boldsymbol{x})u=0,$$

where

$$\mathcal{A}=\frac{1}{N_{11}N_{22}-N_{12}N_{21}}\begin{pmatrix}N_{11} & N_{21}\\ N_{12} & N_{22}\end{pmatrix}. \tag{2.2}$$

To provide a broad setting for the theory and consequently for implementing the finite element scheme, we consider a bounded impenetrable scatterer, D, with smooth boundary Γ contained in a bounded region outside of which $\mathcal{A} = I$ and $n = 1$. This corresponds to the cross section of the cylinder. We denote the unbounded complement of $\bar{D}$ in $\mathbb{R}^2$ by Ω.

Thus, the problem we wish to approximate is the problem **P** of finding u such that

$$\nabla \cdot \mathcal{A}\nabla u + k^2 n u = f \qquad \text{in} \quad \Omega, \tag{2.3}$$

$$u = 0 \qquad \text{on} \quad \Gamma, \tag{2.4}$$

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{\partial u^s}{\partial r} - iku^s\right) = 0, \tag{2.5}$$

$$u = u^i + u^s \qquad \text{in} \quad \Omega. \tag{2.6}$$

Here, the incident field $u^i(\boldsymbol{x})$ is taken to be either a plane wave,

$$u^i(\boldsymbol{x}) = u^i(\boldsymbol{x}, \boldsymbol{d}) = e^{ik\boldsymbol{x}\cdot\boldsymbol{d}}$$

in the direction $\boldsymbol{d}$, $|\boldsymbol{d}| = 1$, where we take $f = 0$, or a point source (the fundamental solution),

$$u^i(\boldsymbol{x}) = u^i(\boldsymbol{x}, \boldsymbol{y}) = \frac{i}{4} H_0^{(1)}\left(k|\boldsymbol{x} - \boldsymbol{y}|\right),$$

where $\boldsymbol{x}$ is the observation point, $\boldsymbol{y}$ is the source point located outside of the scatterer, and $f = \delta(\boldsymbol{x} - \boldsymbol{y})$. Condition (2.5) is known as the Sommerfeld radiation condition and holds uniformly in azimuth angle θ. Equation (2.4) is the metallic or perfectly conducting boundary condition. Motivated by Potthast [23], we assume $\mathcal{A}$ (given by (2.2)) is a complex and uniformly bounded matrix that can be pointwise diagonalized by a unitary complex matrix, $\mathcal{U}$:

$$\mathcal{A}(\boldsymbol{x}) = \mathcal{U}^*(\boldsymbol{x})\mathcal{A}_\Lambda(\boldsymbol{x})\mathcal{U}(\boldsymbol{x}),$$

where $\mathcal{A}_\Lambda$ is a diagonal matrix. We assume further that the real part of $\mathcal{A}_\Lambda$ has uniformly positive diagonal entries, so that

$$a_{min}|\boldsymbol{s}|^2 \le \boldsymbol{s}^T Re\left(\mathcal{A}_\Lambda\right)\overline{\boldsymbol{s}} \le a_{max}|\boldsymbol{s}|^2 \qquad \text{for all } \boldsymbol{s} \text{ in } \mathbb{C}^2, \tag{2.7}$$

where $0 < a_{min} \le a_{max} < \infty$.

We also assume that the domain, Ω, can be decomposed into a finite number of disjoint open sets Ω_m, $m = 0...M$, where $\bigcup_{i=1}^M \overline{\Omega}_i$ ($\overline{\Omega}_i$ denotes the closure of Ω_i) completely contains the anisotropic inhomogeneous medium and each $\Omega_i, i \neq 0$, is bounded with a uniformly Lipschitz-continuous boundary. We choose Ω_0 to be the exterior of $\bigcup_{i=1}^M \overline{\Omega}_i$ and assume that in every subdomain $\mathcal{A}$ and n are continuously differentiable and satisfy one of the following conditions:

1. $\mathcal{A}$ is a positive definite matrix, n is a strictly positive scalar function, both are real-valued, and each component of $\mathcal{A}$, as well as n, are in $H^3(\Omega_m)$ (this implies that each component of $\mathcal{A}$, as well as n, are continuously differentiable in $\overline{\Omega}_m$, [24] Corollary 6.92); or
2. $\mathcal{A}$ and n are complex-valued with $\mathcal{A}$ being semi-coercive, by which we mean

$$-\left(\boldsymbol{s}^T Im(\mathcal{A}_\Lambda)\overline{\boldsymbol{s}}\right) \ge \alpha|\boldsymbol{s}|^2 \qquad \text{for all } \boldsymbol{s} \text{ in } \mathbb{C}^2, \tag{2.8}$$

where either

$$\textbf{i. } \alpha > 0 \text{ and } Im\,(n) \geq 0; \text{ or}$$
$$\textbf{ii. } \alpha \geq 0 \text{ and } Im\,(n) \geq \delta > 0.$$

Note that (2.8) implies, on domains for which condition **2** is satisfied, that

$$-Im\left(\int_{\Omega_m} \overline{\nabla u} \cdot \mathcal{A}\nabla u dA\right) = -\left(\int_{\Omega_m} \overline{\mathcal{U}\nabla u} \cdot Im(\mathcal{A}_\Lambda)\mathcal{U}\nabla u dA\right)$$

$$\geq \alpha \left\|\nabla u\right\|^2_{L^2(\Omega_m)}. \tag{2.9}$$

The purpose of the next two sections is to establish uniqueness of the weak solution to (2.3)-(2.6) and then show that an equivalent problem to **P** can be written as a Fredholm equation from which existence follows. In these sections, we follow closely the techniques of Hazard and Lenoir [13].

For the integral equation approach to this problem in [23], it is assumed that N (and, consequently, $\mathcal{A}$) is continuously differentiable. The approach in this paper is by a variational formulation and allows us to establish existence and uniqueness for piecewise smooth coefficients $\mathcal{A}$ and n. Another motivating factor for presenting this analysis is that it leads directly to a finite element scheme.

3. Uniqueness. We show here that, under the conditions outlined in the previous section, the problem **P** has at most one weak solution. By linearity, this amounts to proving that if the incident wave vanishes, the only solution of **P** in $H^1_{loc}(\Omega)$ is $u \equiv 0$, where

$$H^1_{loc}(\Omega) = \left\{w : \phi w \in H^1_0(\Omega) \quad \forall \phi \in C^\infty_0(\Omega)\right\}.$$

THEOREM 3.1. *If $u \in H^1_{loc}(\Omega)$ is a solution of* **P** *where the coefficients satisfy the conditions outlined in §2 and there is no incident wave, then $u \equiv 0$.*

Proof. Since there is no incident wave, $u^i = 0$ and $f = 0$ in **P**. We begin by proving $u = 0$ on Ω_o and then work through the remaining subdomains.

Let B_R be a ball of radius R and define $\Omega_R = B_R \bigcap \Omega$. If u is a solution of **P**, then u satisfies

$$-\int_{\Omega_R} \overline{\nabla u} \cdot \mathcal{A}\nabla u dA + \int_{\Omega_R} \nabla \cdot (\overline{u}\mathcal{A}\nabla u)\, dA + k^2 \int_{\Omega_R} n|u|^2 dA = 0. \tag{3.1}$$

For R sufficiently large, using the Divergence Theorem and the boundary condition (2.5) on Γ, we have

$$\int_{\partial B_R} \overline{u}\frac{\partial u}{\partial \nu} ds = \int_{\Omega_R} \overline{\nabla u} \cdot \mathcal{A}\nabla u dA - k^2 \int_{\Omega_R} n|u|^2 dA. \tag{3.2}$$

Hence,

$$Im\left(\int_{\partial B_R} \overline{u}\frac{\partial u}{\partial \nu} ds\right) = Im\left(\int_{\Omega_R} \overline{\nabla u} \cdot \mathcal{A}\nabla u dA - k^2 \int_{\Omega_R} n|u|^2 dA\right)$$

and since subdomains that satisfy condition **1**, including Ω_0, do not contribute to the right-hand side,

$$Im\left(\int_{\partial B_R} \overline{u}\frac{\partial u}{\partial \nu} ds\right) = Im\left(\sum_{\Omega_m \text{ of cond. } \mathbf{2}} \int_{\Omega_m} \overline{\nabla u} \cdot \mathcal{A}\nabla u dA - k^2 \int_{\Omega_m} n|u|^2 dA\right).$$

The right-hand side is non-positive due to (2.8). Thus, for all sufficiently large R,

$$Im\left(\int_{\partial B_R} \overline{u}\frac{\partial u}{\partial \nu} ds\right) \leq 0,$$

from which it follows via Rellich's Lemma (Colton and Kress [8]) that $u = 0$ in Ω_o. Note that, in Ω_o, $u \in H^1_{loc}(\Omega_0)$ is a weak solution of $\triangle u + k^2 u = 0$ and hence a classical solution in Ω_0 [24] so the result of [8] applies.

Using (3.2) and the fact that, on Ω_0 and $\partial\Omega$, $u = 0$, it follows by continuity that u vanishes in $\partial\Omega_0$, and we have

$$\sum_{m=1}^{M}\left(\int_{\Omega_m} -\overline{\nabla u}\cdot\mathcal{A}\nabla u dA + k^2\int_{\Omega_m} n|u|^2 dA\right) = 0.$$

If $\mathcal{A}$ satisfies condition **2** on a particular subdomain, then taking the imaginary part leaves

$$\sum_{\Omega_m \, of \, cond. \, \mathbf{2}}\left(Im\left(\int_{\Omega_m} -\overline{\nabla u}\cdot\mathcal{A}\nabla u dA\right) + k^2 Im\left(\int_{\Omega_m} n|u|^2 dA\right)\right) = 0.$$

By (2.9), for each Ω_m where $\mathcal{A}$ and n satisfy condition **2**, we have

$$Im\left(\int_{\Omega_m} -\overline{\nabla u}\cdot\mathcal{A}\nabla u dA\right) = Im\left(k^2\int_{\Omega_m} n|u|^2 dA\right) = 0.$$

Under the assumptions of condition **2i**, we first conclude that $\nabla u = 0$ (u is a constant) and it then follows, since u satisfies (2.3) with $f = 0$, that $u = 0$. If condition **2ii** holds, $u = 0$ directly. In either situation, we conclude that $u = 0$ in each subdomain where $\mathcal{A}$ and n satisfy condition **2**.

Let Ω_m ($m \geq 0$) be a given subdomain of Ω where $\mathcal{A}$ and n satisfy condition **1**. It will be shown that if a solution u vanishes in some subdomain $\Omega_{m'}$ adjacent to Ω_m (i.e., they share an edge), then it also vanishes in Ω_m. On Ω_m, u satisfies the equation

$$-\int_{\Omega_m}\overline{\nabla u}\cdot\mathcal{A}\nabla u dA + \int_{\partial\Omega_m}\overline{u}\,(\mathcal{A}\nabla u)\cdot\nu ds + k^2\int_{\Omega_m} n|u|^2 dA = 0.$$

The following unique continuation result is due to Hörmander ([14], Theorem 17.2.1). We give a special case applicable to this problem.

THEOREM 3.2. *Let $\mathcal{O}$ be a bounded domain in $\mathbb{R}^2$ and suppose that $u \in H^1_{loc}(\mathcal{O})$ satisfies (2.3) with $\nabla\cdot\mathcal{A}\nabla u \in L^2_{loc}(\mathcal{O})$. Suppose in addition that $\mathcal{A}$ and n are real-valued and uniformly Lipschitz continuous in $\mathcal{O}$. Finally, suppose that u vanishes on a ball, $B \subset \mathcal{O}$. Then $u \equiv 0$ in the whole domain $\mathcal{O}$.*

It is shown in Gilbarg and Trudinger [11] that, under the assumptions on $\mathcal{A}$ and n, $u \in H^2_{loc}(\Omega_i)$ (and, hence, $\nabla\cdot\mathcal{A}\nabla u \in L^2_{loc}(\Omega_i)$), where Ω_i is a subdomain in which $\mathcal{A}$ and n satisfy condition **1**. We can thus apply Hörmander's theorem.

Suppose that $\mathcal{A}$ and n satisfy condition **1** in Ω_m and that the subdomain $\Omega_{m'}$ has part of its boundary in common with $\partial\Omega_m$. Consider the domain $\tilde{\Omega} = \Omega_m \cup B$, where B is a small ball centered at a point of $\partial\Omega_m \cap \partial\Omega_{m'}$, with B contained in $\Omega_m \cup \Omega_{m'}$. Let $\tilde{\mathcal{A}}$ and $\tilde{n}$, respectively, denote uniformly Lipschitz continuous and real-valued extensions of $\mathcal{A}$ and n from Ω_m to $\tilde{\Omega}$. These functions can be built by using the Calderon-Zygmund Extension Theorem (Wloka [27], Theorem 5.4) applied

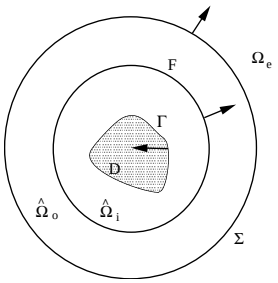


FIG. 4.1. *Orientation of normals and geometry used in the existence proof. The shaded region is the impenetrable scatterer.*

to functions in $H^3(\Omega_m)$. This extension preserves continuity and differentiability at the boundary of Ω_m (Wloka [27], Addendum 5.2). If B is small enough, $\tilde{\mathcal{A}}$ is positive definite and $\tilde{n}$ is positive. Using the fact that u vanishes on $\Omega_{m'}$ and satisfies (2.3) with $f = 0$ on Ω_m, we see that u satisfies

$$\nabla \cdot \tilde{\mathcal{A}} \nabla u + k^2 \tilde{n} u = 0 \qquad \text{in} \quad \tilde{\Omega},$$

and there is an open ball $\tilde{\mathrm{B}}$ contained in B (and, hence, contained in $\tilde{\Omega}$) on which u vanishes. Hence, by the result of Hörmander, u vanishes in $\tilde{\Omega}$, which shows that u vanishes in Ω_m. If $\mathcal{A}$ and n satisfy condition **1** in Ω_m, then, by (2.6), u is constant. Since u vanishes on $\partial\Omega_m \cap \partial\Omega_{m'}$, u is zero in Ω_m. This proves that the only solution to **P**, with $u^i = 0$ and $f = 0$, is $u = 0$, so the problem has at most one solution. Hence, the uniqueness theorem has been proved. □

4. Existence. The first step in proving existence is to derive a reduced problem suitable for later finite element discretization. Let F be a closed uniformly Lipschitz curve surrounding D and Σ a closed uniformly Lipschitz curve surrounding F which has no point in common with F. We assume $\mathcal{A} = I$ and $n = 1$ in a neighborhood and outside of F. Denote by $\hat{\Omega}$ the bounded part of Ω delimited by Σ, and by $\hat{\Omega}_i$ and $\hat{\Omega}_0$ the parts of $\hat{\Omega}$ that are located, respectively, inside and outside of F. Let Ω_e denote the region exterior to Σ. These regions and the orientation of the unit normals are indicated in Figure 4.1.

Let $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \frac{i}{4} H_0^{(1)}(k|\boldsymbol{x} - \boldsymbol{y}|)$, where $H_0^{(1)}(k|\boldsymbol{x} - \boldsymbol{y}|)$ is the Hankel function of the first kind and order zero. Note that Φ satisfies (2.3) with $\mathcal{A} = I$ and $n = 1$, as well as (2.4) with respect to both variables $\boldsymbol{x}$ and $\boldsymbol{y}$, except where $\boldsymbol{x} = \boldsymbol{y}$.

Outside of F, the solution u of **P** satisfies

$$\Delta u + k^2 u = f, \tag{4.1}$$

$$\lim_{r \to \infty} \sqrt{r} \left(\frac{\partial u^s}{\partial r} - iku^s \right) = 0, \tag{4.2}$$

$$u = u^i + u^s. \tag{4.3}$$

In order to obtain a representation of u outside F, the following theorem (found in [3]) is a useful starting point.

THEOREM 4.1. *Let Ω_F denote the unbounded region outside by F. Let $w \in C^2(\Omega_F) \cap C^1(\overline{\Omega}_F)$ be a solution of (4.1) satisfying (4.2). Then, for $\boldsymbol{x} \in \Omega_F$,*

$$w(\boldsymbol{x}) = \int_F \left(w(\boldsymbol{y}) \frac{\partial \Phi}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{x}, \boldsymbol{y}) - \frac{\partial w}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{y}) \Phi(\boldsymbol{x}, \boldsymbol{y}) \right) ds_{\boldsymbol{y}}.$$

Hence, we define

$$I[F;u] := \int_F \left(u(\boldsymbol{y}) \frac{\partial \Phi}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}) - \frac{\partial u}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{y}) \Phi(\boldsymbol{x},\boldsymbol{y}) \right) ds_{\boldsymbol{y}}. \tag{4.4}$$

Note that, for $\boldsymbol{x}$ outside of F, we have

$$u^s(\boldsymbol{x}) = I[F;u^s]. \tag{4.5}$$

Since $I[F;\cdot]$ is linear and $I[F;u^i] = 0$ for $\boldsymbol{x}$ outside F, it follows that

$$u(\boldsymbol{x}) = u^i + I[F;u].$$

Now define the boundary operator on Σ as follows:

$$\mathcal{L}(u) := \left(\frac{\partial u}{\partial \nu_x} - i\lambda u \right)\Big|_{\Sigma},$$

where λ is a nonzero real parameter.

The restriction of u to $\hat{\Omega}$, denoted by $\hat{u}$, is then a solution to the following problem, $\mathbf{P}'$, set in the bounded domain $\hat{\Omega}$:

$$\begin{aligned} \nabla \cdot \mathcal{A} \nabla \hat{u} + k^2 n \hat{u} &= 0 && \text{in } \hat{\Omega}, \\ \hat{u} &= 0 && \text{on } \Gamma, \\ \mathcal{L}(\hat{u} - I[F;\hat{u}]) &= \mathcal{L}\left(\hat{u}^i\right) && \text{on } \Sigma. \end{aligned}$$

To obtain a variational formulation of the above problem, it is necessary to modify the formula for $I[F;\cdot]$ to allow for less smooth functions, i.e., functions in $H^1(\hat{\Omega})$. Let $\Psi(\boldsymbol{y})$ be a function in $C_0^\infty(\hat{\Omega})$ such that $\Psi = 0$ in a neighborhood of Σ and $\Psi = 1$ in a neighborhood of F. Define $(R\Phi)(\boldsymbol{x},\boldsymbol{y}) = \Psi(\boldsymbol{y})\Phi(\boldsymbol{x},\boldsymbol{y})$ and note that

$$R\Phi|_F = \Phi \tag{4.6}$$

and

$$R\Phi|_\Sigma = 0. \tag{4.7}$$

Using (4.4) and (4.6), we have

$$u^s(\boldsymbol{x}) = \int_F u^s(\boldsymbol{y}) \frac{\partial \Phi}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y})\, ds_{\boldsymbol{y}} - \int_F \frac{\partial u^s}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{y}) R\Phi(\boldsymbol{x},\boldsymbol{y})\, ds_{\boldsymbol{y}}. \tag{4.8}$$

By Green's first identity and taking into account the direction of the normals we obtain

$$\begin{aligned} u^s(\boldsymbol{x}) = -k^2 \int_{\hat{\Omega}_o} R\Phi(\boldsymbol{x},\boldsymbol{y})\, u^s(\boldsymbol{y})\, dA_{\boldsymbol{y}} + \int_F u^s(\boldsymbol{y}) \frac{\partial \Phi}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y})\, ds_{\boldsymbol{y}} \\ + \int_{\hat{\Omega}_o} \nabla_{\boldsymbol{y}} u^s(\boldsymbol{y}) \cdot \nabla_{\boldsymbol{y}} R\Phi(\boldsymbol{x},\boldsymbol{y})\, dA_{\boldsymbol{y}} := I^R[F;u^s]. \end{aligned} \tag{4.9}$$

More generally, for every field $u \in C^2(\Omega_o) \bigcap C^1\left(\overline{\Omega}_o\right)$ which satisfies (2.3) in $\hat{\Omega}_o$ it is true that

$$I[F;u] = I^R[F;u].$$

However, $I^R[F;\cdot]$ extends $I[F;\cdot]$ to a map from $H^1(\hat{\Omega}_o) \to C^\infty(\Omega_e)$.

Define the space

$$W := \left\{ f \in L^2(\hat{\Omega}) : \ f_x, f_y \in L^2(\hat{\Omega}) \text{ and } f|_\Gamma = 0 \right\},$$

equipped with the usual $H^1(\hat{\Omega})$ norm, denoted $\|\cdot\|_W$, and inner product

$$(u,v)_W = \int_{\hat{\Omega}} \left(u(\boldsymbol{x})\overline{v}(\boldsymbol{x}) + \nabla u(\boldsymbol{x}) \cdot \nabla \overline{v}(\boldsymbol{x}) \right) d\boldsymbol{x}.$$

The reduced problem $\mathbf{P}''$ can now be written:

Find $\hat{u}$ in W such that

$$\nabla \cdot \mathcal{A} \nabla \hat{u} + k^2 n \hat{u} = 0 \qquad \text{in} \quad \hat{\Omega}, \tag{4.10}$$

$$\hat{u} = 0 \qquad \text{on} \quad \Gamma, \tag{4.11}$$

$$\mathcal{L}\left(\hat{u} - I^R[F;\hat{u}]\right) = \mathcal{L}\left(\hat{u}^i\right) \qquad \text{on} \quad \Sigma. \tag{4.12}$$

It can be seen, by construction, that any solution $u \in H^1_{loc}(\Omega)$ of (2.3)-(2.6) satisfies (4.10)-(4.12). Furthermore, if $\hat{u} \in W$ satisfies (4.10)-(4.12), then

$$u = \begin{cases} \hat{u} & \text{in} \quad \hat{\Omega}, \\ u^i + I[F;\hat{u}] & \text{in} \quad \Omega_e, \end{cases}$$

satisfies (2.3)-(2.6). So the problems are equivalent and uniqueness is established for (4.10)-(4.12) by our argument from the last section.

4.1. The Fredholm result. To show that (4.10)-(4.12) can be rewritten as a Fredholm equation, we use a variational formulation suitable for finite element discretization . Multiplying (4.10) by an arbitrary v in $W$and integrating over $\hat{\Omega}$ yields

$$\int_{\hat{\Omega}} \overline{v} \left(\nabla \cdot \mathcal{A} \nabla \hat{u} \right) dA + k^2 \int_{\hat{\Omega}} \overline{v} n \hat{u} dA = 0.$$

Using the vector identity

$$\nabla \cdot (\overline{v} A \nabla \hat{u}) = \nabla \overline{v} \cdot A \nabla \hat{u} + \overline{v} \left(\nabla \cdot A \nabla \hat{u} \right),$$

and the fact that $\partial\hat{\Omega} = \Gamma \cup \Sigma$, we obtain

$$0 = \int_{\hat{\Omega}} \nabla \overline{v} \cdot \left(\mathcal{A} \nabla \hat{u} \right) dA - k^2 \int_{\hat{\Omega}} n \hat{u} \overline{v} dA - \int_\Sigma \frac{\partial \hat{u}}{\partial \nu} \overline{v} ds.$$

On Σ,

$$\frac{\partial \hat{u}}{\partial \nu} = \mathcal{L}(\hat{u}) + i\lambda \hat{u} = \mathcal{L}\left(I^R[F;\hat{u}]\right) + \mathcal{L}\left(u^i\right) + i\lambda \hat{u},$$

and this results in

$$\int_\Sigma \overline{v} \mathcal{L}\left(u^i\right) ds = -k^2 \int_{\hat{\Omega}} \overline{v} n \hat{u} dA + \int_{\hat{\Omega}} \overline{\nabla v} \cdot \mathcal{A} \nabla \hat{u} dA$$

$$- \int_\Sigma \overline{v} \mathcal{L}\left(I^R[F,\hat{u}]\right) ds - i\lambda \int_\Sigma \overline{v} \hat{u} ds.$$

Now it is easily seen that $\mathbf{P}''$ is equivalent to:

Find u in W such that

$$a_r(u,v) = l_r(v) \quad \text{for all } v \text{ in } W \tag{4.13}$$

where $a_r(\cdot,\cdot)$ is the sesquilinear form defined on W by

$$a_r(u,v) = -k^2 \int_{\hat{\Omega}} \overline{v} n u dA + \int_{\hat{\Omega}} \overline{\nabla v} \cdot \mathcal{A} \nabla u dA$$

$$- \int_{\Sigma} \overline{v} \mathcal{L}\left(I^R[F,u]\right) ds - i\lambda \int_{\Sigma} \overline{v} u ds \tag{4.14}$$

and $l_r(\cdot)$ is the semilinear form given by

$$l_r(v) = \int_{\Sigma} \overline{v} \mathcal{L}\left(u^i\right) ds.$$

Consider then the operators J_r and K_r defined on W by

$$(J_r u, v)_W = \int_{\hat{\Omega}} \overline{\nabla v} \cdot \mathcal{A} \nabla u dA - i\lambda \int_{\Sigma} \overline{v} u ds + k^2 \int_{\hat{\Omega}} \overline{v} u dA \qquad \text{for all } v \text{ in } W,$$

and

$$(K_r u, v)_W = -k^2 \int_{\hat{\Omega}} \overline{v} n u dA - \int_{\Sigma} \overline{v} \mathcal{L}\left(I^R[F,u]\right) ds - k^2 \int_{\hat{\Omega}} \overline{v} u dA \qquad \text{for all } v \text{ in } W.$$

Let L_r be the vector of W associated with the semilinear form $l_r(\cdot)$ by the relation

$$(L_r, v)_W = l_r(v) \qquad \text{for all } v \text{ in } W.$$

The variational formulation (4.13) amounts to the following operator equation:

$$(J_r + K_r) u = L_r \qquad \text{for } u \text{ in } W. \tag{4.15}$$

THEOREM 4.2. *The reduced problem $\mathbf{P}''$ has a unique solution in W.*

The proof is based on showing that J_r and K_r, respectively, are an isomorphism and a compact operator in W. The Fredholm alternative shows that if the only solution to (4.15) with $L_r = 0$ is the trivial solution $u = 0$, then (4.15) has exactly one solution for every $L_r \in W$. The required uniqueness property follows from the uniqueness of the solution of the problem $\mathbf{P}$ (Theorem 3.1), and the previously established equivalence between the latter problem and $\mathbf{P}''$.

LEMMA 4.3. *J_r is a bounded invertible operator in W with bounded inverse.*

Proof. There exists positive constants C_1 and C_2 such that

$$|(J_r v, u)_W| \leq C_1 \int_{\hat{\Omega}} \left|\overline{\nabla v} \nabla u\right| dA + C_2 \int_{\Sigma} |\overline{v} u| \, ds + |k|^2 \int_{\hat{\Omega}} |\overline{v} u| \, dA.$$

Then, by the Schwarz inequality and the trace theorem, there exists a positive constant C_3 such that

$$|(J_r v, u)_W| \leq C_3 \|u\|_W \|v\|_W$$

and J_r is continuous.

By virtue of the Lax-Milgram Lemma, it is enough to show that the sesquilinear form associated with J_r is coercive in W, i.e.,

$$|(J_r v, v)_W| \geq \alpha \|v\|_W^2 .$$

But

$$|(J_r v, v)_W| \geq \left| \left(Re\,(A)\,\nabla v, \overline{\nabla v} \right)_{L^2(\hat{\Omega})} + k^2 \|v\|_{L^2(\hat{\Omega})}^2 \right|,$$

and, by (2.7),

$$|(J_r v, v)_W| \geq C \left(\|\nabla v\|_{L^2(\hat{\Omega})}^2 + k^2 \|v\|_{L^2(\hat{\Omega})}^2 \right),$$

and the lemma follows. □

LEMMA 4.4. *K_r is a compact operator in W.*

Proof. Write K_r as follows:

$$K_r = -K_r^{\Sigma} - k^2 K_r^{\hat{\Omega}},$$

where

$$\left(K_r^{\Sigma} u, v \right)_W = \int_{\Sigma} \overline{v} \mathcal{L} \left(I^R[F, u] \right) ds \qquad \text{for all} \quad v \in W$$

and

$$\left(K_r^{\hat{\Omega}} u, v \right)_W = \int_{\hat{\Omega}} \overline{v} u \,(n+1)\, dx \qquad \text{for all} \quad v \in W.$$

Existence of these operators follows from the Lax-Milgram Lemma.

(**i**) To see that $K_r^{\hat{\Omega}}$ is compact, note that

$$\left| \left(K_r^{\hat{\Omega}} u, v \right)_W \right| \leq \int_{\hat{\Omega}} |\overline{v} u \,(n+1)|\, dx \leq C \|u\|_{L^2(\hat{\Omega})} \|v\|_{L^2(\hat{\Omega})} ,$$

where $C = \underset{\boldsymbol{x} \in \hat{\Omega}}{max} [n\,(\boldsymbol{x})] + 1$. Thus, if W^* is the dual space of $W$and $(K_r^{\hat{\Omega}})^*$ the adjoint, then

$$\left\| \left(K_r^{\hat{\Omega}} \right)^* v \right\|_W = \underset{u \in W}{sup} \frac{\left| \left(K_r^{\hat{\Omega}} u, v \right)_W \right|}{\|u\|_W} \leq \frac{C \|u\|_{L^2(\hat{\Omega})}^2 \|v\|_{L^2(\hat{\Omega})}^2}{\|u\|_W} \leq C \|v\|_{L^2(\hat{\Omega})} .$$

Thus, $(K_r^{\hat{\Omega}})^*$ is a continuous operator from $L^2(\hat{\Omega})$ into W. Compactness of $(K_r^{\hat{\Omega}})^*$ and, hence, of $K_r^{\hat{\Omega}}$ follows from the compactness of the canonical injection from W into $L^2(\hat{\Omega})$.

(**ii**) To see that K_r^{Σ} is compact it is first easily seen that $I^R[F; u]$ is infinitely differentiable in a vicinity of Σ, and if $\mathcal{P}$ denotes any derivative operator on Σ (of any order), it follows that

$$\left| \mathcal{P} I^R[F; u]\,(\boldsymbol{x}) \right| \leq C_{\mathcal{P}} \left(\|u\|_{L^2(F)} + \|u\|_{L^2(\hat{\Omega}_0)} + \|\nabla u\|_{L^2(\hat{\Omega}_0)} \right)$$

at every point $\boldsymbol{x} \in \Sigma$. Note here that $C_{\mathcal{P}}$ is positive and depends on P. Thus

$$|\mathcal{P} I^R[F; u]\,(\boldsymbol{x})\,| \leq C_{\mathcal{P}} \|u\|_W .$$

In particular, there exist a constant, $C \geq 0$, such that

$$\left\| \mathcal{L}\left(I^R[F;u]\right)\right\|_{L^2(\Sigma)} \leq C \left\|u\right\|_W .$$

It then follows that

$$\left| \left(K_r^\Sigma u, v\right)_W \right| \leq C \left\|v\right\|_{L^2(\Sigma)} \left\|u\right\|_W$$

and

$$\left\| \left(K_r^\Sigma\right)^* v \right\|_W = \sup_{u \in W} \frac{\left| \left(K_r^\Sigma u, v\right)_W \right|}{\left\|u\right\|_W} \leq C \left\|v\right\|_{L^2(\Sigma)} ,$$

where $\left(K_r^\Sigma\right)^*$ denotes the adjoint.

Thus, $\left(K_r^\Sigma\right)^*$ is as a continuous operator from $L^2(\Sigma)$ into W. The trace operator $v \to v|_{H^{\frac{1}{2}}(\Sigma)}$ is continuous. Compactness of $\left(K_r^\Sigma\right)^*$ (and, consequently, K_r^Σ) follows from the compactness of the canonical embedding of $H^{\frac{1}{2}}(\Sigma)$ into $L^2(\Sigma)$. ∎

Thus, we have shown that J_r and K_r are an isomorphism and a compact operator, respectively. Existence follows from the Fredholm alternative.

5. Finite element analysis. Here we discretize the variational formulation derived in the previous section and use the finite element method to compute the scattered field. Below we will show that there exists a unique solution to the discretized problem and, at the same time, derive an estimate of the rate of convergence.

Let W_h be a finite dimensional subspace of W made up of piecewise linear functions defined on a regular mesh, where h denotes the minimum diameter of a circle that could contain each element of the mesh. A key assumption is that F and Σ coincide with edges of the mesh.

Standard finite element theory (e.g., Brenner and Scott [2]) shows that the following inequality holds, where u_I denotes the interpolant of u:

$$\left\|u - u_I\right\|_{L^2(\hat{\Omega})} + h \left\|\nabla(u - u_I)\right\|_{L^2(\hat{\Omega})} \leq Ch^2 \left\|u\right\|_{H^2(\hat{\Omega})} . \tag{5.1}$$

The finite-dimensional problem is then to find $u_h \in W_h$ such that

$$a_{r,h}\left(u_h, v_h\right) = l_r\left(v\right) \qquad \text{for all } v_h \text{ in } W_h, \tag{5.2}$$

where $a_{r,h}\left(\cdot,\cdot\right)$ is the approximate bilinear form defined as

$$a_{r,h}\left(u_h, v_h\right) = -k^2 \int_{\hat{\Omega}} \overline{v_h} n u_h dA + \int_{\hat{\Omega}} \overline{\nabla v_h} \cdot \mathcal{A} \nabla u dA$$

$$- \int_\Sigma \overline{v_h} \mathcal{L}\left(I_h^R[F, u_h]\right) ds - i\lambda \int_\Sigma \overline{v_h} u_h ds$$

and we take I_h^R, the discrete version of I^R, to be defined as follows:

$$I_h^R[F;u_h] := \int_F u_h\left(\boldsymbol{y}\right) \frac{\partial \Phi}{\partial \nu_{\boldsymbol{y}}}\left(\boldsymbol{x},\boldsymbol{y}\right) ds_{\boldsymbol{y}} - k^2 \int_{\hat{\Omega}_o} v_h\left(\boldsymbol{x},\boldsymbol{y}\right) u_h\left(\boldsymbol{y}\right) dA_{\boldsymbol{y}}$$

$$+ \int_{\hat{\Omega}_o} \nabla_{\boldsymbol{y}} u_h\left(\boldsymbol{y}\right) \cdot \nabla_{\boldsymbol{y}} v_h\left(\boldsymbol{x},\boldsymbol{y}\right) dA_{\boldsymbol{y}} \tag{5.3}$$

where, for each $\boldsymbol{x}$ in a neighborhood of Σ, $v_h\left(\boldsymbol{x},\cdot\right) \in W_h|_{\hat{\Omega}_o}$ interpolates $R\Phi\left(\boldsymbol{x},\cdot\right)$ as a function on F and vanishes on Σ, but is otherwise arbitrary.

5.1. Existence and uniqueness. In this section, existence and uniqueness of the approximate solution $u_h \in W_h$ will be discussed. The sesquilinear form (4.14) is non-Hermitian and, more importantly, not coercive. As a result, the Lax-Milgram Lemma cannot be directly applied to prove existence and uniqueness. We need to use a Gårding-type inequality.

First, we prove a preliminary result that justifies our choice of v_h in (5.3).

LEMMA 5.1. *$I_h^R[F;u_h]$ is independent of the choice of v_h provided that u_h satisfies (5.2).*

Proof. Let $v_h^{(1)}$ and $v_h^{(2)}$ be two finite element functions as defined after (5.3) and define

$$I_h^{R,j}[F,u_h] := \int_F u_h(\boldsymbol{y}) \frac{\partial \Phi}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y})\, ds_{\boldsymbol{y}} - k^2 \int_{\hat{\Omega}_o} v_h^{(j)}(\boldsymbol{x},\boldsymbol{y})\, u_h(\boldsymbol{y})\, dA_{\boldsymbol{y}}$$

$$+ \int_{\hat{\Omega}_o} \nabla_{\boldsymbol{y}} u_h(\boldsymbol{y}) \cdot \nabla_{\boldsymbol{y}} v_h^{(j)}(\boldsymbol{x},\boldsymbol{y})\, dA_{\boldsymbol{y}}$$

for $j = 1, 2$. Taking the difference,

$$I_h^{R,1}[F,u_h] - I_h^{R,2}[F,u_h] = -k^2 \int_{\hat{\Omega}_o} \left(v_h^{(1)}(\boldsymbol{x},\boldsymbol{y}) - v_h^{(2)}(\boldsymbol{x},\boldsymbol{y}) \right) u_h(\boldsymbol{y})\, dA_{\boldsymbol{y}}$$

$$+ \int_{\hat{\Omega}_o} \nabla_{\boldsymbol{y}} u_h(\boldsymbol{y}) \cdot \nabla_{\boldsymbol{y}} \left(v_h^{(1)}(\boldsymbol{x},\boldsymbol{y}) - v_h^{(2)}(\boldsymbol{x},\boldsymbol{y}) \right) dA_{\boldsymbol{y}}.$$

But, $v_h^{(1)} - v_h^{(2)} = 0$ at the interpolation points on F and Σ and, since $v_h^{(1)} - v_h^{(2)}$ is piecewise linear and F and Σ coincide with the edges of the mesh, we conclude $v_h^{(1)} - v_h^{(2)} = 0$ on F and Σ. As a result, $v_h^{(1)} - v_h^{(2)}$ can be extended by zero to a function $w_h \in W_h$, where $w_h = 0$ on F and Σ. Hence,

$$I_h^{R,1}[F,u_h] - I_h^{R,2}[F,u_h] = a_{r,h}(u_h, w_h) = l_r(w_h).$$

Since $l_r(w_h)$ depends only on $w_h|_\Sigma$, we have that

$$l_r(w_h) = 0.$$

This implies that

$$I_h^{R,1}[F,u_h] = I_h^{R,2}[F,u_h],$$

which is the desired result. ∎

We now proceed with a series of lemmas that ultimately lead to the proof of existence of a unique u_h in W_h that satisfies (5.2).

LEMMA 5.2. *For every v in W, there exists a positive constant C such that*

$$\left\| \mathcal{L} I^R[F;v] - \mathcal{L} I_h^R[F;v] \right\|_{L^2(\Sigma)} \leq Ch \|v\|_W.$$

Furthermore, if v_I is the interpolant of a function v in $W \cap H^2(\hat{\Omega})$, then

$$\left\| \mathcal{L} I^R[F;v_I] - \mathcal{L} I_h^R[F;v_I] \right\|_{L^2(\Sigma)} \leq Ch \|v\|_{H^2(\hat{\Omega})}.$$

Proof. Let v_I denote the interpolant of $R\Phi$ on $\hat{\Omega}$. From the definitions of $I_h^R[F,\cdot]$ and the operator $\mathcal{L}$, which is linear, we have

$$\begin{aligned}\mathcal{L}I^R[F;v]-\mathcal{L}I_h^R[F;v] = &\frac{\partial}{\partial\nu_{\boldsymbol{x}}}\int_{\hat{\Omega}_o}\nabla_{\boldsymbol{y}}v\left(\boldsymbol{y}\right)\cdot\nabla_{\boldsymbol{y}}E\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\\ &-k^2\frac{\partial}{\partial\nu_{\boldsymbol{x}}}\int_{\hat{\Omega}_o}E\left(\boldsymbol{x},\boldsymbol{y}\right)v\left(\boldsymbol{y}\right)dA_{\boldsymbol{y}}\\ &-i\lambda\int_{\hat{\Omega}_o}\nabla_{\boldsymbol{y}}v\left(\boldsymbol{y}\right)\cdot\nabla_{\boldsymbol{y}}E\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\\ &+i\lambda k^2\int_{\hat{\Omega}_o}E\left(\boldsymbol{x},\boldsymbol{y}\right)v\left(\boldsymbol{y}\right)dA_{\boldsymbol{y}},\end{aligned}$$

where $E\left(\boldsymbol{x},\boldsymbol{y}\right)=R\Phi\left(\boldsymbol{x},\boldsymbol{y}\right)-v_I\left(\boldsymbol{x},\boldsymbol{y}\right)$. Hence,

$$\begin{aligned}\left\|\mathcal{L}I^R[F;v]-\mathcal{L}I_h^R[F;v]\right\|_{L^2(\Sigma)} \leq C\Bigg(&\underset{\boldsymbol{x}\in\Sigma}{sup}\left|\int_{\hat{\Omega}_o}\left(\nu_{\boldsymbol{x}}\cdot\nabla_{\boldsymbol{x}}E\left(\boldsymbol{x},\boldsymbol{y}\right)\right)v\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\right|\\ &+\underset{\boldsymbol{x}\in\Sigma}{sup}\left|\int_{\hat{\Omega}_o}\left(\nabla_{\boldsymbol{y}}\nabla_{\boldsymbol{x}}E\left(\boldsymbol{x},\boldsymbol{y}\right)\right)\cdot\nabla_{\boldsymbol{y}}v\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\right|\\ &+\underset{\boldsymbol{x}\in\Sigma}{sup}\left|\int_{\hat{\Omega}_o}E\left(\boldsymbol{x},\boldsymbol{y}\right)v\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\right|\\ &+\underset{\boldsymbol{x}\in\Sigma}{sup}\left|\int_{\hat{\Omega}_o}\nabla_{\boldsymbol{y}}E\left(\boldsymbol{x},\boldsymbol{y}\right)\cdot\nabla_{\boldsymbol{y}}v\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\right|\Bigg).\end{aligned}$$

We can now estimate each term on the right-hand side of the above equation. For example,

$$\underset{\boldsymbol{x}\in\Sigma}{sup}\left|\int_{\hat{\Omega}_o}\left(\nu_{\boldsymbol{x}}\cdot\nabla_{\boldsymbol{x}}E\left(\boldsymbol{x},\boldsymbol{y}\right)\right)v\left(\boldsymbol{x},\boldsymbol{y}\right)dA_{\boldsymbol{y}}\right| \leq \underset{\boldsymbol{x}\in\Sigma}{sup}\left(\left\|\nabla_{\boldsymbol{x}}E\left(\boldsymbol{x},\boldsymbol{y}\right)\right\|_{L^2(\hat{\Omega}_o)}\right)\left\|v\right\|_{L^2(\hat{\Omega}_o)},$$

where

$$\nabla_{\boldsymbol{x}}E=\nabla_{\boldsymbol{x}}\left(\Psi(\boldsymbol{y})\Phi(\boldsymbol{x},\boldsymbol{y})\right)-\nabla_{\boldsymbol{x}}v_I(\boldsymbol{x},\boldsymbol{y})=\Psi(\boldsymbol{y})\nabla_{\boldsymbol{x}}\Phi(\boldsymbol{x},\boldsymbol{y})-\nabla_{\boldsymbol{x}}v_I(\boldsymbol{x},\boldsymbol{y}).$$

Clearly, $\nabla_{\boldsymbol{x}}E\in L^2(\hat{\Omega})$. Similar results hold for the other terms. The first result follows from the above inequalities and the second result follows by adding and subtracting v and using the triangle inequality:

$$\begin{aligned}\left\|\mathcal{L}I^R[F;v_I]-\mathcal{L}I_h^R[F;v_I]\right\|_{L^2(\Sigma)} &\leq Ch\|v_I\|_W\\ &\leq Ch\Big(\|v_I-v\|_W+\|v\|_W\Big)\\ &\leq Ch\Big(h\|v\|_{H^2(\hat{\Omega})}+\|v\|_W\Big)\\ &\leq Ch\|v\|_{H^2(\hat{\Omega})}.\end{aligned}$$

The lemma is proved. ∎

LEMMA 5.3. *For every v, w in W,*

$$|a_{r,h}(v,w) - a_r(v,w)| \leq Ch\|v\|_W\|w\|_W.$$

Also, if v_I is the interpolant of v in W and ξ_h in W_h is arbitrary, then

$$|a_{r,h}(v_I,\xi_h) - a_r(v_I,\xi_h)| \leq Ch\|v\|_{H^2(\hat{\Omega})}\|\xi_h\|_W.$$

Proof. Using the previous lemma,

$$\begin{aligned}|a_{r,h}(v_I,\xi_h) - a_r(v_I,\xi_h)| &= \left|\int_\Sigma \left(\mathcal{L}I^R[F;v_I] - \mathcal{L}I_h^R[F;v_I]\right)\,\xi_h ds\right| \\ &\leq C\left\|\mathcal{L}I^R[F;v_I] - \mathcal{L}I_h^R[F;v_I]\right\|_{L^2(\Sigma)}\|\xi_h\|_{L^2(\Sigma)} \\ &\leq Ch\|v\|_W\|\xi_h\|_W.\end{aligned}$$

The second result in the lemma follows by the triangle inequality:

$$\begin{aligned}|a_{r,h}(v_I,\xi_h) - a_r(v_I,\xi_h)| &\leq Ch\|v_I\|_W\|\xi_h\|_W \\ &\leq Ch\left(\|v_I - v\|_W + \|v\|_W\right)\|\xi_h\|_W \\ &\leq Ch\left(h\|v\|_{H^2(\hat{\Omega})} + \|v\|_W\right)\|\xi_h\|_W \\ &\leq Ch\|v\|_{H^2(\hat{\Omega})}\|\xi_h\|_W,\end{aligned}$$

and both inequalities are proved. ∎

We now have a bound for the discretized sesquilinear form.

LEMMA 5.4. *There exists a positive constant C such that, if $u \in H^2(\hat{\Omega})$,*

$$|a_{r,h}(e_h,e_h)| \leq Ch\|u\|_{H^2(\hat{\Omega})}\|e_h\|_W,$$

where $e_h = w_h - u_h$, $w_h \in W$.

Proof. The analysis is similar to the proof of the First Strang Lemma (see Ciarlet [6]). We first rewrite $a_{r,h}(e_h,e_h)$ as follows:

$$\begin{aligned}a_{r,h}(e_h,e_h) &= a_{r,h}(w_h,e_h) - a_{r,h}(u_h,e_h) \\ &= a_{r,h}(w_h,e_h) - l(e_h) \\ &= a_{r,h}(w_h,e_h) - l(e_h) + a_r(w_h,e_h) - a_r(w_h,e_h) \\ &= a_{r,h}(w_h,e_h) - l(e_h) + a_r(w_h - u,e_h) + a_r(u,e_h) - a_r(w_h,e_h) \\ &= a_r(w_h - u,e_h) + (a_{r,h}(w_h,e_h) - a_r(w_h,e_h)).\end{aligned}$$

Thus,

$$|a_{r,h}(e_h,e_h)| \leq |a_r(w_h - u,e_h)| + |a_{r,h}(w_h,e_h) - a_r(w_h,e_h)|.$$

By the continuity of $a_r(\cdot,\cdot)$ shown in the previous section and taking the supremum over ξ_h in W_h,

$$|a_{r,h}(e_h,e_h)| \leq C\left(\|w_h-u\|_W + \sup_{\xi_h\in W_h}\frac{|a_{r,h}(w_h,\xi_h)-a_r(w_h,\xi_h)|}{\|\xi_h\|_W}\right)\|e_h\|_W.$$

Choosing w_h to be the interpolant of u ($e_h = u_I - u$) and using Lemma 5.3 yields

$$|a_{r,h}(e_h,e_h)| \leq Ch\|u\|_{H^2(\hat{\Omega})}\|e_h\|_W,$$

which is the desired result. □

LEMMA 5.5. *The following Gårding-type inequality holds for all w in W:*

$$|a_{r,h}(w,w)| \geq C_1\|w\|_W^2 - C_2\left(\|w\|_{L^2(\hat{\Omega})}^2 + \|w\|_{L^2(F)}^2 + h^2\|w\|_W^2\right),$$

where C_1 and C_2 are positive constants.

Proof. Using the diagonalizability assumptions on $\mathcal{A}$ and standard estimates,

$$\begin{aligned}
|a_{r,h}(w,w)| \geq{}& \left|\int_{\hat{\Omega}}\overline{\nabla w}\cdot\mathcal{A}\nabla w dA - i\lambda\int_\Sigma |w|^2 ds\right| - |k|^2\int_{\hat{\Omega}}|n||w|^2 dA \\
&- \|w\|_{L^2(\Sigma)}\left\|\mathcal{L}\left(I_h^R[F,w]\right)\right\|_{L^2(\Sigma)} \\
\geq{}& \left[\left(\int_{\hat{\Omega}}\overline{\mathcal{U}\nabla w}\cdot Re(\mathcal{A}_\Lambda)\mathcal{U}\nabla w dA\right)^2\right. \\
&\left.+\left(\int_{\hat{\Omega}}\overline{\mathcal{U}\nabla w}\cdot Im(\mathcal{A}_\Lambda)\mathcal{U}\nabla w dA - \lambda\int_\Sigma |w|^2 ds\right)^2\right]^{\frac{1}{2}} \\
&- |k^2|\int_{\hat{\Omega}}|n||w|^2 dA - \|w\|_{L^2(\Sigma)}\left\|\mathcal{L}\left(I_h^R[F,w]\right)\right\|_{L^2(\Sigma)}.
\end{aligned}$$

By (2.9),

$$\left(\int_{\hat{\Omega}}\overline{\mathcal{U}\nabla w}\cdot Im(\mathcal{A}_\Lambda)\mathcal{U}\nabla w dA - \lambda\int_\Sigma |w| ds\right)^2 \geq \left(\lambda\int_\Sigma |w| ds\right)^2.$$

Then, provided $\lambda > 0$, there exists positive constants C_1, C_2 and C_3 such that

$$\begin{aligned}
|a_{r,h}(w,w)| \geq{}& C_1\|\nabla w\|_{L^2(\hat{\Omega})}^2 + C_2\|w\|_{L^2(\Sigma)}^2 - C_3\|w\|_{L^2(\hat{\Omega})} \\
&- \|w\|_{L^2(\Sigma)}\left\|\mathcal{L}\left(I^R[F,w]\right)\right\|_{L^2(\Sigma)} \\
\geq{}& C_1\|\nabla w\|_{L^2(\hat{\Omega})}^2 + C_2\|w\|_{L^2(\Sigma)}^2 - C_3\|w\|_{L^2(\hat{\Omega})} \\
&- \|w\|_{L^2(\Sigma)}\left\|\mathcal{L}\left(I_h^R[F,w]\right) - \mathcal{L}\left(I^R[F,w]\right)\right\|_{L^2(\Sigma)} \\
&- \|w\|_{L^2(\Sigma)}^2\left\|\mathcal{L}\left(I^R[F,w]\right)\right\|_{L^2(\Sigma)}.
\end{aligned} \tag{5.4}$$

By applying the arithmetic-geometric mean to (5.4) for all $\epsilon > 0$ and for all $\delta > 0$,

$$\begin{aligned}|a_{r,h}(w,w)| \geq C_1 \|\nabla w\|^2_{L^2(\hat{\Omega})} + C_2 \|w\|^2_{L^2(\Sigma)} - C_3 \|w\|^2_{L^2(\hat{\Omega})} \\ - \left(\frac{\epsilon}{2}\|w\|^2_{L^2(\Sigma)} + \frac{1}{2\epsilon}\left\|\mathcal{L}\left(I^R[F,w]\right)\right\|^2_{L^2(\Sigma)}\right) \\ - \left(\frac{\delta}{2}\|w\|^2_{L^2(\Sigma)} + \frac{1}{2\delta}\left\|\mathcal{L}\left(I^R_h[F,w]\right) - \mathcal{L}\left(I^R[F,w]\right)\right\|^2_{L^2(\Sigma)}\right),\end{aligned}$$

and, taking $\epsilon = \delta = C_2$,

$$\begin{aligned}|a_{r,h}(w,w)| \geq C_1 \|\nabla w\|^2_{L^2(\hat{\Omega})} - C_3 \|w\|^2_{L^2(\hat{\Omega})} \\ -\frac{1}{2C_2}\left(\left\|\mathcal{L}\left(I^R[F,w]\right)\right\|^2_{L^2(\Sigma)}\right. \\ \left. + \left\|\mathcal{L}\left(I^R_h[F,w]\right) - \mathcal{L}\left(I^R[F,w]\right)\right\|^2_{L^2(\Sigma)}\right).\end{aligned} \tag{5.5}$$

The idea now is to bound (5.5). Using integration by parts,

$$\begin{aligned}I^R[F,w] = \int_F w(\boldsymbol{y})\frac{\partial\Phi}{\partial\nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y})\,ds_{\boldsymbol{y}} - k^2\int_{\hat{\Omega}_0} R\phi(\boldsymbol{x},\boldsymbol{y})\,w(\boldsymbol{y})\,dA_{\boldsymbol{y}} \\ + \int_{\hat{\Omega}_0} w(\boldsymbol{y})\,\triangle_{\boldsymbol{y}} R\Phi(\boldsymbol{x},\boldsymbol{y})\,dA_{\boldsymbol{y}} - \int_F w(\boldsymbol{y})\frac{\partial R\Phi}{\partial\nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y})\,ds_{\boldsymbol{y}}.\end{aligned}$$

The derivative, $\frac{\partial}{\partial\nu_{\boldsymbol{x}}}$, on the operator $\mathcal{L}$ is only applied to $R\Phi$ since w depends on the variable of integration $\boldsymbol{y}$; thus, there exists a positive C such that

$$\begin{aligned}\left\|\mathcal{L}\left(I^R[F,w]\right)\right\|^2_{L^2(\Sigma)} &\leq C\left(\|w\|^2_{L^2(F)} + \|w\|^2_{L^2(\hat{\Omega}_0)}\right) \\ &\leq C\left(\|w\|^2_{L^2(F)} + \|w\|^2_{L^2(\hat{\Omega})}\right).\end{aligned}$$

Using Lemma 5.2,

$$|a_{r,h}(w,w)| \geq C_1\|\nabla w\|^2_{L^2(\hat{\Omega})} - C_2\left(\|w\|^2_{L^2(\hat{\Omega})} + \|w\|^2_{L^2(F)} + h^2\|w\|^2_W\right),$$

and the result follows. □

Although existence and uniqueness are not guaranteed in general by the above Gårding-type inequality of the previous lemma, they can be shown to hold under certain circumstances. In particular, they will be shown here for h sufficiently small ($h \ll 1$). The proof uses the ideas of Shatz [25].

Let $\psi \in H^1(\hat{\Omega})$ be such that

$$a_r(v,\psi) = (v,e)_{L^2(\hat{\Omega})} \qquad \text{for all } v \in H^1(\hat{\Omega}), \tag{5.6}$$

where $e = u - u_h$.

We now show that ψ is well-defined. By interchanging the order of integration in $a_r(\cdot,\cdot)$, it can be seen that ψ is related to $w \in H^1(\Omega)$ which is a weak solution of

$$\nabla \cdot \overline{\mathcal{A}} \nabla w + k^2 \overline{n} w = e \qquad \text{in} \quad \Omega, \tag{5.7}$$

$$w = 0 \qquad \text{on} \quad \Gamma, \tag{5.8}$$

$$\lim_{r\to\infty} \sqrt{r} \left(\frac{\partial w}{\partial r} + ikw \right) = 0, \tag{5.9}$$

by the relation

$$w|_{\hat{\Omega}} := \begin{cases} \psi + (1-\Psi)\, T(\psi) & \text{in} \quad \hat{\Omega}_o, \\ \psi & \text{in} \quad \hat{\Omega}_i, \end{cases} \tag{5.10}$$

where T is defined as follows:

$$T(\psi) = \int_\Sigma \psi(\boldsymbol{y}) \left(\frac{\partial \overline{\Phi}}{\partial \nu_{\boldsymbol{y}}}(\boldsymbol{x},\boldsymbol{y}) + i\lambda \overline{\Phi}(\boldsymbol{x},\boldsymbol{y}) \right) ds_{\boldsymbol{y}}.$$

Recall that Ψ is defined in the discussion prior to equations (4.6) and (4.7). Existence and uniqueness of w follows the analysis of the original problem (using the appropriate radiation condition (5.9)). The integral equation

$$\psi + T(\psi) = w$$

is uniquely solvable for ψ on Σ (see Colton and Kress [8]) and, because T depends only on $\psi|_\Sigma$, existence and uniqueness of ψ follows.

It is assumed that there exists (for the index $0 < \gamma \leq 1$) a positive constant C such that the a priori estimate

$$\|\psi\|_{H^{1+\gamma}(\hat{\Omega}_i)} + \|\psi\|_{H^{1+\gamma}(\hat{\Omega}_o)} \leq C \|e\|_{L^2(\hat{\Omega})} \qquad \text{for all } e \text{ in } L^2(\hat{\Omega}) \tag{5.11}$$

holds, where $e = u - u_h$. Note also that, since $\gamma > 0$,

$$\|\psi - P\psi\|_W \leq Ch^\gamma \|\psi\|_{H^{1+\gamma}(\hat{\Omega})}, \tag{5.12}$$

where P is the orthogonal projection from W into W_h. For smooth coefficients and smooth boundaries, equation (5.11) holds for $\gamma = 1$.

LEMMA 5.6. *For every v in W, there exists a positive constant C such that*

$$\|v\|^2_{L^2(F)} \leq C \|v\|_{L^2(\hat{\Omega})} \|v\|_W.$$

Proof. Let $q \in H^1(\hat{\Omega})$ be such that

$$\Delta q - q = 0 \qquad \text{in } \hat{\Omega},$$

$$\frac{\partial q}{\partial \nu} = 0 \qquad \text{on } \Sigma,$$

$$\frac{\partial q}{\partial \nu} = 1 \qquad \text{on } F,$$

and choose $w = \nabla q$. Then

$$w \cdot \nu = \begin{cases} 0 & \text{on } \Sigma, \\ 1 & \text{on } F. \end{cases}$$

Using the Divergence Theorem,

$$\begin{aligned}\int_F v^2 w\cdot\nu ds &= \int_{\partial\hat{\Omega}} v^2 w\cdot\nu ds\\ &= \int_{\hat{\Omega}} \nabla\cdot\left(v^2 w\right) dA\\ &= \int_{\hat{\Omega}} \left(2v\nabla v\cdot w + v^2\nabla\cdot w\right) dA\\ &\leq C\left(\|v\|_{L^2(\hat{\Omega})}\|\nabla v\|_{L^2(\hat{\Omega})} + \|v\|^2_{L^2(\hat{\Omega})}\right),\end{aligned}$$

where C depends on w. It follows that

$$\|v\|^2_{L^2(F)} \leq C\|v\|_{L^2(\hat{\Omega})}\|v\|_W$$

and the inequality is established. □

Combining the above inequality, Lemma 5.4 and the Gårding-type inequality (with $w = e_h$),

$$\begin{aligned}C_1\|e_h\|^2_W - C_2\Big(\|e_h\|^2_{L^2(\hat{\Omega})} + \|e_h\|_{L^2(\hat{\Omega})} + \|e_h\|_W + h^2\|e_h\|^2_W\Big)\\ \leq Ch\|u\|_{H^2(\hat{\Omega})}\|e_h\|_W.\end{aligned}$$

Using the arithmetic-geometric mean, for all $\epsilon > 0$ and for all $\delta > 0$,

$$\begin{aligned}C_1\|e_h\|^2_W - C_2\Big(\|e_h\|^2_{L^2(\hat{\Omega})} + \frac{\delta}{2}\|e_h\|^2_W + \frac{1}{2\delta}\|e_h\|^2_{L^2(\hat{\Omega})} + h^2\|e_h\|^2_W\Big)\\ \leq C\left(\frac{1}{2\epsilon}h^2\|u\|^2_{H^2(\hat{\Omega})} + \frac{\epsilon}{2}\|e_h\|^2_W\right).\end{aligned}$$

Choosing a δ and ϵ such that

$$\frac{\delta}{2}C_2 \leq \frac{C_1}{4}$$

and

$$\frac{\epsilon}{2}C \leq \frac{C_1}{4}$$

yields

$$\left(\frac{C_1}{2} - C_2h^2\right)\|e_h\|^2_W - C_2\left(1+\frac{\delta}{2}\right)\|e_h\|^2_{L^2(\hat{\Omega})} \leq \frac{C}{2\epsilon}h^2\|u\|^2_{H^2(\hat{\Omega})}.$$

If h is small enough, there are new constants C_1, C_2 and C_3 (all > 0) such that

$$C_1\|e_h\|^2_W - C_2\|e_h\|^2_{L^2(\hat{\Omega})} \leq C_3h^2\|u\|^2_{H^2(\hat{\Omega}},$$

independent of h. Hence,

$$C_1 \left(\|u-u_h\|_W - \|u-u_I\|_W\right)^2 - C_2 \left(\|u-u_h\|_{L^2(\hat{\Omega})} + \|u-u_I\|_{L^2(\hat{\Omega})}\right)^2$$

$$\leq h^2 C_3 \|u\|^2_{H^2(\hat{\Omega})},$$

or, with C_3 a new constant,

$$C_1 \|u-u_h\|^2_W - 2\left(C_1+C_2\right)\|u-u_h\|_W \|u-u_I\|_W$$

$$- C_2 \|u-u_h\|^2_{L^2(\hat{\Omega})} \;\leq\; h^2 C_3 \|u\|^2_{H^2(\hat{\Omega})}.$$

The arithmetic-geometric mean then yields, for all $\epsilon > 0$,

$$C_1 \|u-u_h\|^2_W - 2\left(C_1+C_2\right)\left(\frac{\epsilon}{2}\|u-u_h\|^2_W + \frac{1}{2\epsilon}\|u-w_h\|^2_W\right)$$

$$- C_2 \|u-u_h\|^2_{L^2(\hat{\Omega})} \;\leq\; h^2 C_3 \|u\|^2_{H^2(\hat{\Omega})},$$

and so, with new constants C_1, C_2 and C_3,

$$C_1 \|u-u_h\|^2_W - C_2 \|u-u_h\|^2_{L^2(\hat{\Omega})} \;\leq\; C_3 h^2 \|u\|^2_{H^2(\hat{\Omega})}. \tag{5.13}$$

LEMMA 5.7. *There exists a positive constant C such that*

$$\|e\|_{L^2(\hat{\Omega})} \leq C\left(h^\gamma \|e\|_W + h\|u\|_W\right). \tag{5.14}$$

Proof. By (5.6),

$$\begin{aligned}
(e,e)_{L^2(\hat{\Omega})} &= a_r(e,\psi) = a_r(u,\psi) - a_r(u_h,\psi)\\
&= a_r(u,\psi-P\psi) + a_r(u,P\psi) - a_r(u_h,\psi)\\
&\quad + a_{r,h}(u_h,\psi-P\psi) - a_{r,h}(u_h,\psi-P\psi)\\
&\quad + a_r(u_h,\psi-P\psi) - a_r(u_h,\psi-P\psi)\\
&= a_r(e,\psi-P\psi) + a_{r,h}(e,\psi) - a_r(e,\psi)\\
&\quad + a_{r,h}(u,\psi) - a_r(u,\psi)\\
&\quad + (a_r(e,\psi-P\psi) - a_{r,h}(e,\psi-P\psi)\\
&\quad + a_r(u,\psi-P\psi) - a_{r,h}(u,\psi-P\psi).
\end{aligned} \tag{5.15}$$

Using continuity of $a_r(\cdot,\cdot)$ in (5.15) and Lemma 5.4,

$$\begin{aligned}
\|e\|^2_{L^2(\hat{\Omega})} &\leq \|e\|_W \|\psi-P\psi\|_W + h\|e\|_W\|\psi\|_W\\
&\quad + h\|u\|_W\|\psi\|_W + h\|e\|_W\|\psi-P\psi\|_W\\
&\quad + h\|u\|_W\|\psi-P\psi\|_W.
\end{aligned}$$

Applying (5.11) and (5.12), as well as noting that $h(1+h^\gamma) \leq Ch$ if h is bounded above, yields the desired result. ∎

Using Lemma 5.7, equation (5.13) becomes

$$C_1 \|u-u_h\|_W^2 - C\left(h^\gamma \|u-u_h\|_W + h\|u\|_W\right)^2 \le Ch^2 \|u\|^2_{H^2(\hat{\Omega})},$$

or, with a new C,

$$C_1 \|u-u_h\|_W^2 - C_2 h^{2\gamma} \|u-u_h\|_W^2 - C_2 h^{2(\gamma+1)} \|u-u_h\|_W \|u\|_W \le Ch^2 \|u\|^2_{H^2(\hat{\Omega})}.$$

By applying the arithmetic-geometric mean, for all $\epsilon > 0$,

$$C_1 \|u-u_h\|_W^2 - C_2 h^{2\delta} \|u-u_h\|_W^2 - C_2 h^{2(\delta+1)} \left(\frac{\epsilon}{2}\|u-u_h\|_W^2 + \frac{1}{2\epsilon}\|u\|_W^2\right) \le Ch^2 \|u\|^2_{H^2(\hat{\Omega})}.$$

For h small enough,

$$C_1 - Ch^{2\delta} - h^{2(\delta-1)}\frac{\epsilon}{2} > 0,$$

and so, with C a new constant,

$$C_1 \|u-u_h\|_W^2 \le Ch^2 \|u\|^2_{H^2(\hat{\Omega})}.$$

Taking u to be zero in the above inequality proves the uniqueness of the discrete solution. We have proved:

THEOREM 5.8. *For h sufficiently small, there exists a unique solution to (5.2). Furthermore, if u is smooth enough, there exists a constant C, independent of h, (but, depending on k) such that*

$$\|u-u_h\|_W \le Ch\|u\|_{H^2(\hat{\Omega})}. \tag{5.16}$$

Remark. By a similar argument as the proceeding analysis, if piecewise $p-$degree polynomials are used to discretize the problem so that

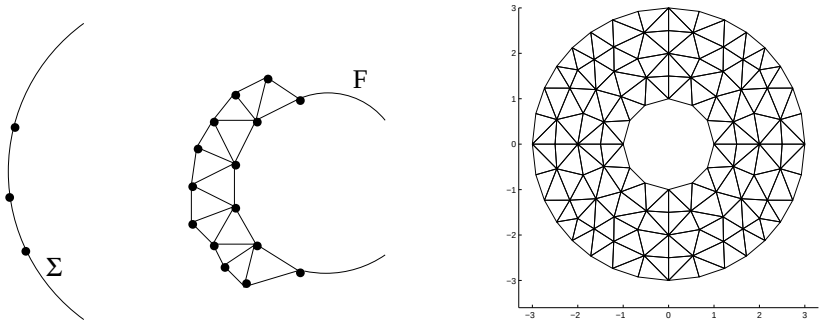
$$\|u-u_I\|_{L^2(\hat{\Omega})} + \|\nabla(u-u_I)\|_{L^2(\hat{\Omega})} \le Ch^p \|u\|_{H^{p+1}(\hat{\Omega})},$$

it then follows that

$$\|u-u_h\|_W \le h^p \|u\|_{H^{p+1}(\hat{\Omega})}.$$

5.2. Numerical integration. The mass and stiffness matrices for (5.2) can be approximated by quadrature in the normal way. The only difficult term of (5.2) involves $I_h^R[F,u_h]$. We use (5.3) and choose v_h that interpolates Φ on F and is zero on all elements that do not touch F (see Figure 5.1 *a*). Justification of this follows from the fact that the I_h^R was shown to be independent of the choice of the representation of $R\Phi$ from W_h. This limits the coupling of nodes and, as a result, the number of potentially nonzero entries in the system matrix. Thus, in an element with nodes $j=1,2,3$, at least one of which is on F, we have

$$v_h(\boldsymbol{x},\boldsymbol{y}) = \frac{i}{4}\sum_{j=1}^{3} p(j)\,\phi_j(\mathbf{y})\,H_0^{(1)}\left(k|\boldsymbol{x}-\boldsymbol{y}_j|\right), \tag{5.17}$$



(*a*) Coupling of the nodes between F and Σ. (*b*) The mesh m.

FIG. 5.1. *Figure (a) shows the layer of nodes around F that are coupled with the nodes on Σ. Figure (b) shows the coarse mesh denoted by m.*

mesh	**triangles**	**nodes**	**nonzero entries**	***F* and Σ coupling**	**percent**
m	688	388	8073	5632	70
m1	688	388	8073	5632	70
m2	1540	840	17756	12288	69
m3	2752	1464	31205	21504	69
m4	4300	2660	48417	33280	69
m5	6192	3228	69393	47616	69
m6	8428	4368	94133	70272	75

TABLE 5.1
Comparison of the number of mesh nodes and nonzero entries in the resulting system matrix.

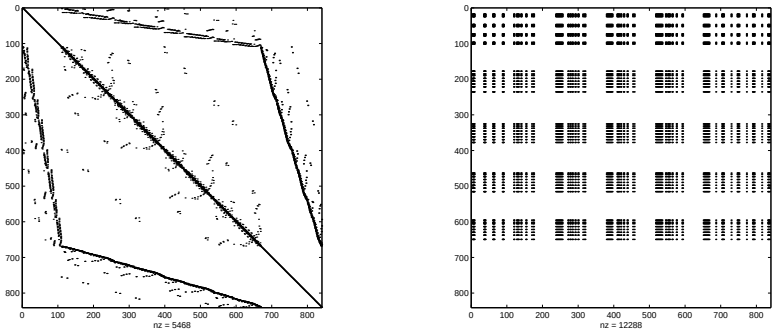
where $p\,(j)$ is zero if node j is not on F and 1 if node j is on F, and ϕ_j is the finite element basis function associated with $\boldsymbol{y}_j$.

Consider, for example, scattering by an impenetrable disc of radius 1. Using the Modulef mesh generating package, a coarse ($h = 0.8123$) triangular mesh, denoted by m, was created where Γ, F, and Σ are taken to be circles of radius 1, 2, and 3, respectively (see Figure 5.1 b). This mesh was refined four times, each by dividing number of triangles in the original mesh n^2 times, where $n = 2, ..., 5$. These refinements are denoted $m3, ..., m6$, respectively. Two additional meshes, $m1$ and $m2$, were generated separately with h values that are between the h values for m and $m3$ (see Table 6.1 for the h values). Each of the meshes cover the domain $\hat{\Omega}$. The resulting total number of nonzero entries in the system matrix corresponding to $a_{r,h}(\cdot,\cdot)$, as well as the number due the to coupling between F and Σ, are shown in Table 5.1 and in Figure 5.2.

6. Numerical examples. We present four computational examples, the first two of which are of an impenetrable scatterer and a penetrable scatterer, respectively. In both cases, the focus is on computing the near field and verifying the rate of convergence suggested in (5.16).

Since one of the goals of computational scattering is often to predict the far-field pattern of the scattered wave, in the third example it is shown that the far-field can be found easily once the near-field has been determined.

Finally, we have avoided singular integrals in the coupling of F and Σ which have



(*a*) No coupling of between F and Σ. (*b*) Coupling between F and Σ only.

FIG. 5.2. *Nonzero entries in the system matrix for the mesh m1. Figure (a) shows the nonzero entries without the coupling of F and Σ and (b) shows only the nonzero entries due to coupling of F and Σ.*

no point in common. The best choice of the distance between these two curves and the effect of this distance on the accuracy of the solution is not so clear. We offer some insight to these questions in the fourth example.

6.1. Impenetrable scatterer: near-field. The specific problem here is to compute the field scattered from an impenetrable circular scatterer of radius 1, denoted by D. The artificial boundary Σ is taken to be the circle of radius 3 and F the circle of radius 2. The meshes from the previous section are used in the computations. We choose $\lambda = k = 4.0$, which provides at least ten nodes per wavelength in $m4 - m6$, and the mesh m is not used because it is too coarse. We take $\mathcal{A} \equiv I$ and $n \equiv 1$ everywhere, and the incident field $u^i = e^{ik\boldsymbol{x}}$ (i.e., $\mathbf{d} = (1, 0)$).

The analytic solution outside of the scatterer, which is assumed to be centered at the origin, can be written as the series

$$u(r_o, \theta) = \sum_{n=-\infty}^{\infty} \frac{(i)^n J_n(ka)}{H_n^{(1)}(ka)} H_n^{(1)}(kr_o)\, e^{in\theta}, \tag{6.1}$$

where r_o is the distance from the origin of the observation point and θ is the azimuthal angle.

The finite element matrix equation is solved using the LU decomposition of the coefficient matrix. This was done using the IMSL Fortran subroutines dlftzg, which computes the LU factorization for a complex general sparse matrix, and dlfszg, which uses the LU factorization to solve the matrix equation.

We are then able to investigate the error estimate (5.16) and the L^2-error, as well as the maximum relative error

$$\frac{\max_j |u^j - u_h^j|}{\max_j |u^j|}, \tag{6.2}$$

where the superscript j is the index of the j^{th} grid point. Both the series (6.1) and its gradient were computed using Matlab, taking the sum from -20 to 20. Any additions past $|n| = 20$ were on the order of 10^{-10} or smaller.

The error results obtained using each mesh are given in Table 6.1. Table 6.2 shows the slopes of the lines joining the errors of consecutive meshes using a log-log

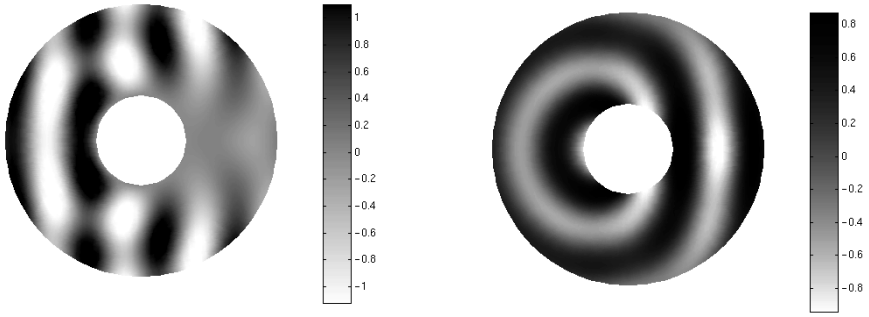
scale as well as the slope of the line that best fits all six data points. In the case where all six errors are used, slopes of 1.9093 and 1.3058 were obtained for rates of convergence in the L^2 and H^1-norms, respectively. These numerical results indicate a rate of convergence in the H^1-norm of order h as was predicted by Theorem 5.8. They also suggest a rate of h^2 in the L^2-norm which, although not proved in this paper, would be optimal for the L^2-norm convergence. The imaginary part of the computed total field and scattered field are shown in Figure 6.1.

mesh	h	**max rel. error**	L^2 **-error**	H^1 **-error**
m1	0.4062	0.4392	0.4058	0.4794
m2	0.2724	0.2208	0.1975	0.2771
m3	0.2054	0.1285	0.1149	0.1890
m4	0.1649	0.0846	0.0747	0.1424
m5	0.1377	0.0598	0.0524	0.1142
m6	0.1182	0.0442	0.0387	0.0954

TABLE 6.1
The h value and all three errors for the six meshes.

mesh	L^2 **-error**	H^1 **-error**
m1 to m2	1.8024	1.3483
m2 to m3	1.9179	1.3553
m3 to m4	1.9602	1.2859
m4 to m5	1.9712	1.2284
m5 to m6	1.9895	1.1845
m1 to m6	1.9093	1.3058

TABLE 6.2
Slopes between consecutive meshes.



(*a*) Real part of the total field. (*b*) Real part of the scattered field.

FIG. 6.1. *The real part of the total and scattered fields in the case of an impenetrable object.*

6.2. Penetrable scatterer: near-field. This problem differs from the previous one in that the scatterer, D, is a penetrable isotropic object with boundary that is the circle of radius 1. In this case, $\mathcal{A}$ takes the form

$$\mathcal{A} = \begin{pmatrix} \hat{a} & 0 \\ 0 & \hat{a} \end{pmatrix}.$$

Again, the artificial boundaries Σ and F are taken to be the circles of radius 3 and 2, respectively. Six mesh, $pm1, ...pm6$, were generated in the same manner as in previous example. They differ from $m1 - m6$ in that the meshes are now discs without the hole in the middle. In this case, $pm2 - pm6$ are all n^2 refinements of $pm1$, where $n = 2, ..., 6$, see Table 6.3. We choose $\lambda = k = 3.0$ and take the isotropy to be $\hat{a} = 2 - \frac{1}{2}i$ in D, $\mathcal{A} = I$ outside of D and $n = 1$ everywhere.

The analytic solution inside D can be written as the series

$$u(r_o, \theta) = \sum_{n=-\infty}^{\infty} a_n J_n(Kr_o) e^{in\theta},$$

where r_o is the distance from the origin of the observation point, θ is the azimuthal angle and

$$K = \frac{k^2}{\hat{a}} = \frac{9}{2 - \frac{1}{2}i}.$$

For this example, we take the source point to be located at $(4, 0)$. Outside D the incident field (which is a point source) and scattered field are known to be

$$u^i(r, \theta) = \sum_{n=-\infty}^{\infty} H_n^{(1)}(kr_s) J_n(kr_o) e^{in\theta} \tag{6.3}$$

and

$$u^s(r_o, \theta) = \sum_{n=-\infty}^{\infty} b_n H_n^{(1)}(kr_o) e^{in\theta}, \tag{6.4}$$

respectively.

Using these series representations and the following conditions at the boundary of D :

$$\left(2 - \frac{i}{2}\right) \frac{\partial u}{\partial r} = \frac{\partial u^i}{\partial r} + \frac{\partial u^s}{\partial r},$$

$$u = u^i + u^s,$$

we can evaluate the coefficients a_n and b_n. We can then analyze the finite element errors as in the previous example. The error results obtained using each mesh are given in Table 6.3 and, as before, Table 6.2 shows the slopes of the lines joining the errors of consecutive meshes using a log-log scale as well as the slope of the line that best fits all six data points. Using a log-log plot, slopes of 1.9533 and 1.2873 were obtained for rates of convergence in the L^2 and H^1-errors, respectively, using all six meshes. The results again indicate a rate of convergence in the H^1-norm of order h, as was predicted by Theorem 5.8, and a rate of h^2 for the L^2-norm convergence. The imaginary part of the computed total field and scattered field are shown in Figure 6.2.

6.3. Penetrable scatterer: far-field. One advantage of the method is that the $I^R[F, \cdot]$ operator defined in (4.9) provides a way to compute the scattered field outside F, using only the knowledge of the scattered field and its gradient near F. This has

mesh	triangles	h	max error	L^2 error	H^1 -error
pm1	200	0.8388	0.4279	0.5447	0.6364
pm2	800	0.4534	0.1556	0.1773	0.2814
pm3	1800	0.3134	0.0759	0.0843	0.1718
pm4	3200	0.2396	0.0443	0.0487	0.1231
pm5	5000	0.1940	0.0228	0.0316	0.0969
pm6	7200	0.1630	0.0202	0.0221	0.0789

TABLE 6.3
The h value and all three errors for the six meshes.

mesh	L^2 -error	H^1 -error
pm1 to pm2	1.8615	1.3265
pm2 to pm3	1.9514	1.3362
pm3 to pm4	2.0435	1.2414
pm4 to pm5	2.0488	1.1336
pm5 to pm6	2.0538	1.1803
pm1 to pm6	1.9533	1.2873

TABLE 6.4
Slopes between consecutive meshes.

already been implicitly demonstrated in the set up of the variational formulation. This equation also provides a way to compute the far-field pattern, again with only the knowledge of the scattered field and its gradient near F. The operator $I^R[F,\cdot]$ needs only to be modified using the asymptotic properties of the fundamental solution. In this case, the result, denoted $I^R_\infty[F,\cdot]$, is

$$I^R_\infty[F,u] = \int_F u^s(\boldsymbol{y}) \frac{\partial}{\partial \nu_{\boldsymbol{y}}} e^{-ik\hat{\boldsymbol{x}}\cdot\boldsymbol{y}} ds_{\boldsymbol{y}} - k^2 \int_{\hat{\Omega}_o} Re^{-ik\hat{\boldsymbol{x}}\cdot\boldsymbol{y}} u^s(\boldsymbol{y})\, dA_{\boldsymbol{y}}$$

$$+ \int_{\hat{\Omega}_o} \nabla_{\boldsymbol{y}} u^s(\boldsymbol{y}) \cdot \nabla_{\boldsymbol{y}} Re^{-ik\hat{\boldsymbol{x}}\cdot\boldsymbol{y}} dA_{\boldsymbol{y}}. \tag{6.5}$$

The setting for this example will be the same as for the previous example with the following exceptions. The isotropy is given by $\hat{a} = 2$ and the incident field is taken to be a plane wave changing (6.3) to

$$u^i(r_o,\theta) = \sum_{n=-\infty}^{\infty} i^n J_n(kr_o) e^{in\theta}.$$

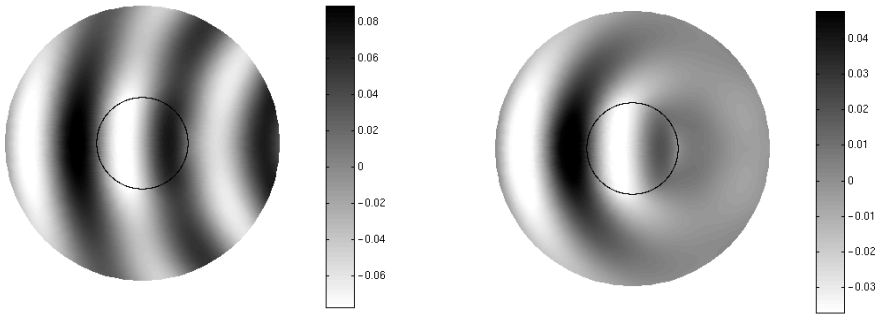
Using the properties of the Hankel function for large arguments,

$$H^{(1)}_n(z) = \sqrt{\frac{2}{\pi z}} e^{i\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right)},$$

we obtain the following series representation for the far-field pattern

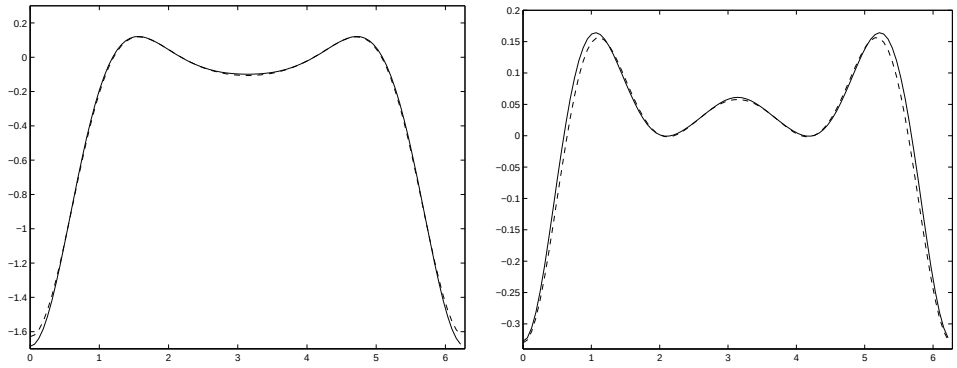
$$u_\infty = \sum_{n=-\infty}^{\infty} b_n \sqrt{\frac{2}{\pi k}} e^{-i\left(\frac{n\pi}{2} + \frac{\pi}{4}\right)} e^{in\theta}.$$

The scattering data from the finite element code (using the mesh *pm*6) was used and the far-field was computed at 100 evenly spaced points on the unit circle. The results are shown in Figure 6.3 where the maximum relative error is 0.0301.



(*a*) Imaginary part of the total field. (*b*) Imaginary part of the scattered field.

FIG. 6.2. *The imaginary part of the total and scattered fields. The boundary of the scatterer is a circle of radius* 1 *outlined in black.*



(*a*) Real part. (*b*) Imaginary part.

FIG. 6.3. *The computed and series representation of the far-field. Figure (a) shows the real part and (b) shows the imaginary part. The exact series solution is the solid line and the far-field computed using the scattered field generated from the finite element code is the dashed line.*

7. Distance between F and Σ. Table 5.1 shows a significant increase in the number of nonzero entries in the system matrix due to the coupling between F and Σ that results simply from the refinements of the original mesh. However, Tables 6.1 and 6.3 suggest that the refinement is necessary to obtain a reasonable degree of accuracy. This example demonstrates the influence that the distance between F and Σ has on the accuracy of the computed solution.

To demonstrate this interaction, we use the first example of an impenetrable scatterer given in §7.3. A mesh was made with Σ a circle of radius 3, F a circle of radius ranging from 1.20 to 2.80 and the scatterer a disc of radius 1. Since a major factor depending on the distance between F and Σ is the argument of the fundamental solution $k|\boldsymbol{x}-\boldsymbol{y}|$, several possible values of the wave number k, $1.0, 2.0, 4.0, 6.0$ and 8.0, were chosen. With these choices, the value of $k|\boldsymbol{x}-\boldsymbol{y}|$ ranges from 0.2 to 46.4. As in each of the previous examples, we take $\lambda = k$. The value of the resulting maximum relative error from each computation is shown in Table 7.1.

Note that, even though F varies in the above example, the same mesh was used each time. For this mesh, $h = 0.1344$. It is necessary to have at least one layer of elements between the object and F as well as between F and Σ. In practice, this is

not a restriction because at least ten nodes per wavelength are needed to obtain an accurate solution. Also, because we can choose the discretized $R\Phi$ using a function that is nonzero only on the first layer of elements exterior to Ω_i, we are not required to have more than two layers of elements between F and Σ. In fact, Table 7.1 indicates that it is desirable to have F close to Σ.

Radius of F	**Wave number**				
	1.0	2.0	4.0	6.0	8.0
1.2	0.0008	0.0048	0.0430	0.1698	0.4477
1.6	0.0008	0.0048	0.0429	0.2693	0.4474
2.0	0.0007	0.0047	0.0423	0.1667	0.4431
2.4	0.0007	0.0047	0.0392	0.1484	0.3850
2.8	0.0007	0.0047	0.0392	0.1484	0.3850

TABLE 7.1

Cross reference between the maximum relative error and the radius of F. The radius of Σ is fixed at $r = 3$.

8. Concluding remarks. While the method used here is successful, there are several questions about methods of this type that remain to be investigated, both in the context presented in this paper as well as their use in other problems. For example, the optimal positioning of Σ and F for accuracy and stability is not known. In addition, the best approach to solving the matrix problem, taking into account the coupling between Σ and F, needs to be investigated. We finally note that the method appears to be quite useful in computing scattering from more complicated anisotropic objects as can be seen in [22], where an example of two different anisotropic objects generating the same exterior scattered field is shown.

Acknowledgment and Disclaimer. The effort of Peter Monk was sponsored by the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under grant number F49620-96-1-0039. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the US Government. The effort of Joseph Coyle was sponsored in part by the NSF under grant number DMS-9631287. The authors would like to thank the editor and referees for suggesting many important improvements to the paper.

REFERENCES

[1] J. BÉRENGER, *A perfectly matched layer for the absorbtion of electromagnetic waves*, J. Comput. Phys., 114 (1994), pp. 110–117.

[2] S. BRENNER AND R. SCOTT, *The Mathematical Theory of Finite Element Methods*, vol. 15 of Texts in Applied Mathematics Ser., Springer-Verlag, New York, 1996.

[3] K. CHADAN, D. COLTON, L. PÄIVÄRINTA, AND W. RUNDELL, *An Introduction to Inverse Scattering and Inverse Spectral Problems*, Society for Industrial and Applied Mathematics, 1997.

[4] G. CHEN AND J. ZHOU, *Boundary Element Methods*, Academic Press, 1992.

[5] W. CHEW AND F. TEIXERIA, *Analytical derivation of a conformal perfectly matched absorber for electromagnetic waves*, Microwave and Optical technology letters, 17 (1998), pp. pp 231–236.

[6] P. G. CIARLET, *Numerical Analysis of the Finite Element Method*, North-Holland, 1974.

[7] F. Collino and P. Monk, *The perfectly matched layer in curvilinear coordinates*, SIAM J. Sci. Computing, 19 (1998), pp. pp 2061 – 2090.
[8] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory*, John Wiley and Sons, Inc., 1983.
[9] D. Colton and P. Monk, *A linear sampling method for the detection of leukemia using microwaves*, SIAM Journal of Applied Math., 58 (1998), pp. pp 926 – 941.
[10] B. Engquist and A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp., 31 (1977), pp. 629–651.
[11] D. Gilbarg and N. Trudinger, *Elliptic Partial Differentail Equations of Second Order*, Springer-Verlag, 1985.
[12] D. Givoli and J. Keller, *Exact non-reflecting boundary conditions*, J. Comp. Phys., 82 (1989), pp. 172–192.
[13] C. Hazard and M. Lenoir, *On the solutions of time-harmonic scattering problems for Maxwell's equations*, SIAM Journal on Mathematical Analysis, (1996).
[14] L. Hörmander, *The Analysis of Linear Partial Differential Operations, III*, Springer-Verlag, 1985.
[15] G. C. Hsiao, *The coupling of BEM and FEM - a brief review*, Boundary Elements V., 1 (1988), pp. 431–445.
[16] ———, *The coupling of boundary element and finite element methods*, Math. Mech., 6 (1990), pp. 493–503.
[17] A. Jami and M. Lenoir, *A variational formulation for exterior problems in linear hydrodynamics*, Computational Methods in Applied Mechanical Enginering, 16 (1978), pp. 341–359.
[18] J. Jin, *The Finite Element Method in Electromagnetics*, Wiley, 1993.
[19] A. Kirsch and P. Monk, *Convergence analysis of the coupled finite element method and spectral method in acoustic scattering*, IMA J. on Numerical Analysis, 9 (1990), pp. 425–447.
[20] ———, *An analysis of the coupling of finite element and Nyström methods in acoustic scattering*, IMA J. on Numerical Analysis, 14 (1994), pp. 523–544.
[21] M. Masmoudi, *Numerical solution for exterior problems*, Numer. Math., 51 (1987), pp. 87–101.
[22] M. Piana, *On uniqueness for anisotropic inhomogeneous inverse scattering problems*, Inverse Problems, 14 (1998), pp. 1565–1579.
[23] R. Potthast, *Electromagnetic scattering from an orthotropic medium*, Journal of Integral Equations, 11 (1999).
[24] M. Renardy and R. Rogers, *An Introduction to Partial Differential Equations*, Springer-Verlag, 1993.
[25] A. H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Mathematics of Computation, 28 (1974), pp. 959–962.
[26] B. Stupfel and R. Mittra, *Numerical absorbing boundary conditions for the scalar and vector wave equations*, IEEE Transactions on Antennas and Propagation, 44 (1996), p. 1015.
[27] J. Wloka, *Partial Differential Equations*, Cambridge University Press, 1987.