SIAM J. SCI. COMPUT.
Vol. 21, No. 5, pp. 1927–1949
© 2000 Society for Industrial and Applied Mathematics

FIRST-ORDER SYSTEM LEAST-SQUARES FOR THE HELMHOLTZ EQUATION*

B. LEE†, T. A. MANTEUFFEL‡, S. F. MCCORMICK‡, AND J. RUGE‡

Abstract. This paper develops a multilevel least-squares approach for the numerical solution of the complex scalar exterior Helmholtz equation. This second-order equation is first recast into an equivalent first-order system by introducing several "field" variables. A combination of scaled L^2 and H^{-1} norms is then applied to the residual of this system to create a least-squares functional. It is shown that, in an appropriate Hilbert space, the homogeneous part of this functional is equivalent to a squared *graph* norm, that is, a product norm on the space of individual variables. This equivalence to a norm that decouples the variables means that standard finite element discretization techniques and standard multigrid solvers can be applied to obtain optimal performance. However, this equivalence is not uniform in the wavenumber k, which can signal degrading performance of the numerical solution process as k increases. To counter this difficulty, we obtain a result that characterizes the error components causing performance degradation. We do this by defining a finite-dimensional subspace of these components on whose orthogonal complement k-uniform equivalence is proved for this functional and an analogous functional that is based only on L^2 norms. This subspace equivalence motivates a nonstandard multigrid method that attempts to achieve optimal convergence uniformly in k. We report on numerical experiments that empirically confirm k-uniform optimal performance of this multigrid solver. We also report on tests of the error in our discretization that seem to confirm optimal accuracy that is free of the so-called *pollution* effect.

Key words. first-order system least-squares, Helmholtz equation, nonstandard multigrid

AMS subject classifications. 65F10, 65F30, 65N30, 65N55

PII. S1064827598339773

1. Introduction.

The scalar Helmholtz equation with exterior radiation boundary conditions describes a variety of wave propagation phenomena. One such phenomenon is the electromagnetic scattering of time-harmonic waves, which, since the advent of stealth technology, has helped propel an interest in efficient numerical solution techniques. However, such boundary value problems are challenging because they are both indefinite and non-self-adjoint. Hence, standard numerical procedures for solving them generally suffer from poor discretization accuracy and slow convergence of the algebraic solver. Accuracy can be improved by taking meshes that are refined enough for the given wavenumber, but such refinements are impractical for many problems. Also, convergence of the iterative solver can be improved by preconditioning the algebraic system, but finding an appropriate preconditioner is problematic and, in general, robust iterative solvers for indefinite and non-self-adjoint problems are difficult to design.

*Received by the editors May 29, 1998; accepted for publication (in revised form) February 16, 1999; published electronically April 28, 2000. This work was sponsored by the National Science Foundation under grant DMS-9706866 and the Department of Energy under grant DE-FG03-93ER25165.
http://www.siam.org/journals/sisc/21-5/33977.html

†Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-661, Livermore, CA 94551 (lee123@llnl.gov).

‡Applied Math Department, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526 (tmanteuf@boulder.colorado.edu, stevem@boulder.colorado.edu, jruge@boulder.colorado.edu).

Standard multigrid methods are no exception. Helmholtz problems tax multigrid methods by admitting certain highly oscillatory error components that yield relatively small residuals. Because these components are oscillatory, standard coarse grids cannot represent them well, so coarsening cannot eliminate them effectively. Because they yield small residuals, standard relaxation methods cannot effectively reduce them. Compounding these difficulties is the property that the dimension of the subspace of troublesome components increases with increasing wavenumber k. One approach to ameliorate these difficulties (cf. [9], [10], [21], and [23]) is to introduce ray-like basis functions on the coarser grids, via exponential interpolation and weighting, and to use multiple coarsening, where several coarse grids are used at a given discretization level to resolve some of these error components. Introduction of ray basis functions allows the oscillatory error components to be represented on coarser grids, and successive coarsening with an increasing but controlled number of grids allows a full range of these components to be resolved on the coarser levels. Since the magnitude of the wavenumber k dictates the oscillatory nature of the error components, then judicious introduction of exponential interpolation and weighting and of new coarse grid problems as a function of k leads to multigrid schemes whose convergence is k-uniform.

Direct application of such a multigrid algorithm to the scalar Helmholtz equation would not generally achieve optimal discretization accuracy that is uniform in k. Because k-uniform optimality of the discretization and multigrid solver is the central aim of this paper, the multigrid algorithm will be applied instead to a carefully designed first-order system least-squares (FOSLS) (cf. [11], [12], [21], and [22]) formulation of the Helmholtz problem. This FOSLS methodology involves recasting the scalar equation into a first-order system by introducing "field" variables, deriving boundary conditions for these new variables, and applying a least-squares approach to the resulting first-order system boundary value problem.

An important consideration in the numerical solution of Helmholtz problems is the so-called *pollution effect*. The results in [3], [5], and [19] show that, for large wavenumber k, even when $kh \ll 1$, the dispersive nature of the Galerkin finite element discretization introduces a large phase lead error term and, in fact, a *sharp* H^1 error estimate is contaminated by a pollution term unless $k^2h \ll 1$. A *stabilized* Galerkin least-squares finite element method was developed in [3], [5], and [19] to reduce the pollution effect and, hence, to ameliorate the effects of this restrictive condition.

In the work presented below, because the FOSLS methodology leads to a minimization principle, a Rayleigh–Ritz principle can be applied. Thus, standard bases can be used to achieve $O(kh)$ discretization error bounds in the *least-squares norm* (that is, the norm determined by the functional itself) for the FOSLS formulation. Our numerical results will suggest that this pollution-free performance is achieved by the FOSLS formulation—not only the least-squares norm but in a scaled H^1 norm as well.

Thus, the aim of this paper is two-fold: k-uniform multigrid convergence factors and pollution-free discretization accuracy. Since a fast k-uniformly convergent solver has been the most difficult to obtain historically, this will be the primary focus of our work. However, we will also show in our final numerical experiment that the FOSLS discretization appears to be pollution free.

Substantial literature is available on computational electromagnetics. Algorithms that apply to the time-dependent Maxwell equations or div-curl systems in general include those described in [13], [20], [24], [26], and [27].

Studies based on first-order system formulations of the Helmholtz equation include [16], [18], [21], and [23]. The approach described in [16] does not include a curl term in the functional, so its discretizations are somewhat restrictive and its standard multigrid methods could not perform well with large wavenumbers, as expected. The functional in [18] incorporates a curl expression but not in a way that achieves uniformity in the discretization (or the multigrid solver, had they analyzed it). In fact, except for [21] and [23] which is the basis of the work presented in the present paper, none of the methods cited above were shown to achieve optimal discretization accuracy and multigrid convergence.

For general literature on FOSLS, see [8], [11], [12], [21], and [22], which treat least-squares functionals based on L^2 and L^2-H^{-1} (that is, a combination of L^2 and H^{-1}) inner products. The least-squares functionals described in the next section are also based on these norms. In fact, the analytical techniques used in these papers are also used to derive some of the theoretical results presented below. These techniques enable the establishment of several uniform norm equivalence results that would not be possible by way of the more restrictive Agmon–Douglis–Nirenberg theory (cf. [2]).

This paper is organized as follows. Section 2 introduces the functional setting, the first-order system, and the L^2 and L^2-H^{-1} least-squares formulations. In section 3, a coercivity result for the L^2-H^{-1} formulation defined over an appropriate Hilbert space is established using a compactness argument. Unfortunately, this coercivity estimate does not hold uniformly with respect to the wavenumber, so in section 4 we introduce an appropriate subspace on which we do prove k-uniform coercivity. These estimates provide an intuitive basis for the nonstandard multigrid method developed in section 5. In the final section, we report on numerical tests that demonstrate k-uniform pollution-free optimality of the discretization technique and multigrid solver.

2. FOSLS formulation. Let $D \subset \Re^2$ be a bounded domain, which for technical reasons (e.g., scaled H^1 equivalence of the L^2 FOSLS functional; cf. Theorems 3.2 and 4.2) will be assumed to have a $C^{1,1}$ or convex polygonal boundary Γ_i of positive measure. Consider the exterior Helmholtz boundary value problem

$$\left\{\begin{array}{rcll} -\nabla\cdot\nabla p - k^2 p & = f & & \text{in } \Re^2\backslash\bar{D}, \\ p & = 0 & & \text{on } \Gamma_i, \\ \frac{\partial p}{\partial r} - \imath k p & = o\left(\|\mathbf{x}\|^{-\frac{1}{2}}\right) & & \text{as } \|\mathbf{x}\| \to \infty. \end{array}\right. \tag{2.1}$$

Here, k is the wavenumber, f is in $L^2_{\text{loc}}\left(\Re^2\backslash\bar{D}\right)$, and $\frac{\partial}{\partial r}$ denotes the derivative in the radial direction used in the asymptotic Sommerfeld radiation condition (cf. [14]). To solve (2.1) numerically, the unbounded domain is truncated and the Sommerfeld radiation condition is approximated. Thus, let B be a ball containing D ($D \subset\subset B$) with $C^{1,1}$ boundary Γ_e, so that the region bounded by $\Gamma := \Gamma_i \cup \Gamma_e$ is annulus-like. Assuming now that $f \in L^2(\Omega)$, consider the reduced boundary value problem

$$\left\{\begin{array}{rcll} -\nabla\cdot\nabla p - k^2 p & = f & & \text{in } \Omega, \\ p & = 0 & & \text{on } \Gamma_i, \\ \mathbf{n}\cdot\nabla p - \imath k p & = 0 & & \text{on } \Gamma_e. \end{array}\right. \tag{2.2}$$

Here, a first-order approximation to the Sommerfeld radiation condition is used. Assume henceforth that (2.2) has a unique solution in $H^2(\Omega)$.

Introducing the "field" variable $\mathbf{u} = \frac{1}{k}\nabla p$, (2.2) can be recast into the first-order system

$$\left\{\begin{array}{rcll} \mathbf{u} - \frac{1}{k}\nabla p & = \mathbf{0} & & \text{in } \Omega, \\ \frac{1}{k}\nabla \cdot \mathbf{u} + p & = -\frac{f}{k^2} & & \text{in } \Omega, \\ \frac{1}{k}\nabla \times \mathbf{u} & = 0 & & \text{in } \Omega, \\ p & = 0 & & \text{on } \Gamma_i, \\ \mathbf{n}\cdot\mathbf{u} - \imath p & = 0 & & \text{on } \Gamma_e, \\ \mathbf{n}\times\mathbf{u} & = 0 & & \text{on } \Gamma_i \end{array}\right. \tag{2.3}$$

or

$$L\begin{pmatrix} \mathbf{u} \\ p \end{pmatrix} = \mathbf{f},$$

where L is the first-order operator corresponding to (2.3). Note that (2.3) is essentially a scaled form of Maxwell's equations that has been reduced by assuming a transverse magnetic wave solution. Variable p gives the electric field and variable $\mathbf{u}$ gives the magnetic field.

To analyze (2.3) and the corresponding least-squares formulations, several functional spaces are required. With s being a nonnegative integer, let $H^s(\Omega)$ denote the usual Sobolev space of order s with norm $\|\cdot\|_s$. With $t \in \Re^+$, let $H^t(\Gamma)$ denote the usual Sobolev trace space with norm $\|\cdot\|_t$ (cf. [1]). Let $\langle\cdot,\cdot\rangle$ denote the $L^2(\Gamma)$ inner product. The following spaces and scaled H^1 norm will be needed:

$$\begin{aligned} H_0^1(\Omega) &= \left\{w \in H^1(\Omega) : w = 0 \text{ on } \Gamma\right\}, \\ H_{0,d}^1(\Omega) &= \left\{w \in H^1(\Omega) : w = 0 \text{ on } \Gamma_i\right\}, \\ \mathbf{H}(\mathrm{div};\Omega) &= \left\{\mathbf{u} \in [L^2(\Omega)]^2 :\ \nabla\cdot\mathbf{u} \in L^2(\Omega)\right\}, \\ \mathbf{H}_0(\mathrm{div};\Omega) &= \left\{\mathbf{u} \in \mathbf{H}(\mathrm{div};\Omega) :\ \mathbf{n}\cdot\mathbf{u} = 0 \text{ on } \Gamma_e\right\}, \\ \mathbf{H}(\mathbf{curl};\Omega) &= \left\{\mathbf{u} \in [L^2(\Omega)]^2 :\ \nabla\times\mathbf{u} \in L^2(\Omega)\right\}, \\ \mathbf{H}_0(\mathbf{curl};\Omega) &= \left\{\mathbf{u} \in \mathbf{H}(\mathbf{curl};\Omega) :\ \mathbf{n}\times\mathbf{u} = 0 \text{ on } \Gamma_i\right\}, \\ \mathbf{W}(\Omega) &= \left\{(\mathbf{u},p) \in \left[(\mathbf{H}(\mathrm{div};\Omega)\cap\mathbf{H}_0(\mathbf{curl};\Omega))\times H_{0,d}^1(\Omega)\right] : \right. \\ &\qquad\qquad \left. \mathbf{n}\cdot\mathbf{u} = \imath p \text{ on } \Gamma_e\right\}, \\ \|\cdot\|_{1,k} &= \sqrt{\|\cdot\|^2 + \frac{1}{k^2}\|\nabla\cdot\|^2}. \end{aligned}$$

The boundary condition imposed in the definition of $\mathbf{W}(\Omega)$ is to be taken in the trace sense. Finally, let $H_0^{-1}(\Omega)$ and $H_{0,d}^{-1}(\Omega)$ denote the completion of $L^2(\Omega)$ in the respective inverse norms

$$\|\cdot\|_{-1} = \sup_{w\in H_0^1(\Omega)} \frac{(\cdot,w)}{\|w\|_{1,k}} \tag{2.4}$$

and

$$\|\cdot\|_{-1,d} = \sup_{w\in H^1_{0,d}(\Omega)} \frac{(\cdot, w)}{\|w\|_{1,k}}\,. \tag{2.5}$$

It is well known (cf. [17]) that $[\mathbf{H}(\mathrm{div};\Omega)\cap \mathbf{H}_0(\mathbf{curl};\Omega)]\times H^1(\Omega)$ is a Hilbert space in the norm

$$|||(\mathbf{u},p)||| := \|\mathbf{u}\|_{\mathrm{div}} + \|\mathbf{u}\|_{\mathbf{curl}} + \|p\|_1,$$

where

$$\|\mathbf{u}\|_{\mathrm{div}} = \|\mathbf{u}\| + \|\nabla\cdot\mathbf{u}\|$$

and

$$\|\mathbf{u}\|_{\mathbf{curl}} = \|\mathbf{u}\| + \|\nabla\times\mathbf{u}\|\,.$$

Thus, it is also a Hilbert space in the norm

$$|||(\mathbf{u},p)|||_k := \|\mathbf{u}\|_{\mathrm{div}} + \|\mathbf{u}\|_{\mathbf{curl}} + \|p\|_{1,k}.$$

Using the continuity of the trace operators from $H^1(\Omega)$ onto $H^{\frac12}(\Gamma)$ and $\mathbf{H}(\mathrm{div};\Omega)$ into $H^{-\frac12}(\Gamma)$, it is easy to see that $\mathbf{W}(\Omega)$ is also a Hilbert space in the $|||\cdot|||_k$ norm (cf. [22]).

Define the L^2 and L^2-H^{-1} least-squares functionals by

$$F(\mathbf{u},p;f) = \left\|\mathbf{u}-\frac1k\nabla p\right\|^2 + \left\|\frac1k\nabla\cdot\mathbf{u}+p+\frac{f}{k^2}\right\|^2 + \left\|\frac1k\nabla\times\mathbf{u}\right\|^2 \tag{2.6}$$

and

$$G(\mathbf{u},p;f) = \left\|\mathbf{u}-\frac1k\nabla p\right\|^2 + \left\|\frac1k\nabla\cdot\mathbf{u}+p+\frac{f}{k^2}\right\|^2_{-1,d} + \left\|\frac1k\nabla\times\mathbf{u}\right\|^2_{-1}, \tag{2.7}$$

respectively. Both functionals incorporate curl terms that, with essentially no increase in complexity for F and only moderate increase for G, allow for simplified finite element discretization and multigrid solution processes.

For brevity, the equivalence results of the next two sections focus on G; we will state but not prove the analogous $H^{1,k}$ equivalence results for F (cf. [21] for detailed proofs of similar results).

3. Nonuniform coercivity: $\mathbf{W}(\Omega)$. A compactness argument (cf. [7] and [8]) is used here to establish equivalence between the homogeneous part of the L^2-H^{-1} least-squares functional, $G(\mathbf{u},p;0)$, and the square of the product norm $[\|\cdot\|]^2\times\|\cdot\|_{1,k}$. In what follows, C denotes a generic constant that may change meaning at each occurrence but is independent of k. The notation C_k will be used (in this section only) when dependence on k may occur.

THEOREM 3.1. *There exist constants C and C_k such that*

$$\frac{1}{C_k}\left[\|\mathbf{u}\|^2+\|p\|^2_{1,k}\right] \le G(\mathbf{u},p;0) \le C\left[\|\mathbf{u}\|^2+\|p\|^2_{1,k}\right]$$

for all $(\mathbf{u},p)\in\mathbf{W}(\Omega)$.

Proof. Let $(\mathbf{u},p) \in \mathbf{W}(\Omega)$. To prove the *upper bound*, note that the triangle inequality yields

$$G(\mathbf{u},p;0) \le 2\left(\|\mathbf{u}\|^2 + \frac{1}{k^2}\|\nabla p\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}\right\|_{-1,d}^2 + \|p\|_{-1,d}^2 + \left\|\frac{1}{k}\nabla\times\mathbf{u}\right\|_{-1}^2 \right). \tag{3.1}$$

Integration by parts, the triangle and Cauchy–Schwarz inequalities, the trace theorem, and the ϵ-inequality give

$$\begin{aligned}
\left\|\frac{1}{k}\nabla\cdot\mathbf{u}\right\|_{-1,d} &= \sup_{w\in H^1_{0,d}(\Omega)} \frac{\left|\left(\frac{1}{k}\nabla\cdot\mathbf{u},w\right)\right|}{\|w\|_{1,k}} \\
&= \sup_{w\in H^1_{0,d}(\Omega)} \frac{\left|\left(-\frac{1}{k}\mathbf{u},\nabla w\right)+\left\langle\frac{1}{k}\mathbf{n}\cdot\mathbf{u},w\right\rangle\right|}{\|w\|_{1,k}} \\
&\le \sup_{w\in H^1_{0,d}(\Omega)} \frac{\left|\left(\mathbf{u},\frac{1}{k}\nabla w\right)\right|}{\|w\|_{1,k}} + \sup_{w\in H^1_{0,d}(\Omega)} \frac{\left|\left\langle\frac{1}{\sqrt{k}}p,\frac{1}{\sqrt{k}}w\right\rangle\right|}{\|w\|_{1,k}} \\
&\le \|\mathbf{u}\| + C \sup_{w\in H^1_{0,d}(\Omega)} \frac{\left(\|p\|\left\|\frac{1}{k}\nabla p\right\|\right)^{\frac{1}{2}}\left(\|w\|\left\|\frac{1}{k}\nabla w\right\|\right)^{\frac{1}{2}}}{\|w\|_{1,k}} \\
&\le \|\mathbf{u}\| + C\|p\|_{1,k}. \qquad (3.2)
\end{aligned}$$

Also, integration by parts and the boundary condition on $H^1_0(\Omega)$ give

$$\begin{aligned}
\left\|\frac{1}{k}\nabla\times\mathbf{u}\right\|_{-1} &= \sup_{w\in H^1_0(\Omega)} \frac{\left|\left(\frac{1}{k}\nabla\times\mathbf{u},w\right)\right|}{\|w\|_{1,k}} \\
&= \sup_{w\in H^1_0(\Omega)} \frac{\left|\left(\mathbf{u},\frac{1}{k}\nabla^\perp w\right)\right|}{\|w\|_{1,k}} \\
&\le \sup_{w\in H^1_0(\Omega)} \frac{\|\mathbf{u}\|\left\|\frac{1}{k}\nabla w\right\|}{\|w\|_{1,k}} \\
&\le \|\mathbf{u}\|. \qquad (3.3)
\end{aligned}$$

Finally,

$$\begin{aligned}
\|p\|_{-1,d} &\le \|p\| \\
&\le \|p\|_{1,k}. \qquad (3.4)
\end{aligned}$$

The upper bound then follows from (3.1)–(3.4).

To prove the *lower bound*, let $\mathcal{W}(\Omega)$ be the completion of $\mathbf{W}(\Omega)$ in the $\|\mathbf{u}\| + \|p\|_{1,k} + \|\mathbf{n}\cdot\mathbf{u} - \imath p\|_{-\frac{1}{2},\Gamma_e}$ norm. To prove first that

$$\begin{aligned}
&\|\mathbf{u}\|^2 + \|p\|_{1,k}^2 + \|\mathbf{n}\cdot\mathbf{u}-\imath p\|_{-\frac{1}{2},\Gamma_e}^2 \\
&\qquad \le C\left(\left\|\mathbf{u}-\frac{1}{k}\nabla p\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|_{-1,d}^2 + \|p\|^2 \right.
\end{aligned}$$

$$+ \|\mathbf{n}\cdot\mathbf{u} - \imath p\|^2_{-\frac{1}{2},\Gamma_e}\Bigg) \tag{3.5}$$

on $\mathcal{W}(\Omega)$, note that the triangle inequality used twice and (3.4) yield

$$\begin{aligned}\|\mathbf{u}\|^2 &\le C\left(\left\|\mathbf{u}-\frac{1}{k}\nabla p\right\|^2 + \left\|\frac{1}{k}\nabla p\right\|^2\right)\\ &\le C\left(\left\|\mathbf{u}-\frac{1}{k}\nabla p\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|^2_{-1,d} + \|p\|^2_{1,k}\right). \qquad (3.6)\end{aligned}$$

Hence, (3.5) would follow if it could be shown that

$$\|p\|^2_{1,k} \le C\left(\left\|\mathbf{u}-\frac{1}{k}\nabla p\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|^2_{-1,d} + \|p\|^2\right). \tag{3.7}$$

To this end, note that integration by parts produces

$$\begin{aligned}\frac{1}{k^2}\|\nabla p\|^2 = \left(-\mathbf{u}+\frac{1}{k}\nabla p, \frac{1}{k}\nabla p\right) + \left(-\frac{1}{k}\nabla\cdot\mathbf{u}-p, p\right) + (p,p)\\ + \frac{\imath}{k}\langle p,p\rangle. \qquad (3.8)\end{aligned}$$

Moreover, since $\frac{1}{k^2}\|\nabla p\|^2$ is real, then the right-hand side of (3.8) must be real, which, together with the triangle inequality, complex modulus arithmetic, and the Cauchy–Schwarz and ϵ-inequalities, yields

$$\begin{aligned}\frac{1}{k^2}\|\nabla p\|^2 &= \mathrm{Re}\left[\left(-\mathbf{u}+\frac{1}{k}\nabla p, \frac{1}{k}\nabla p\right)\right] + \mathrm{Re}\left[\left(-\frac{1}{k}\nabla\cdot\mathbf{u}-p, p\right)\right] + (p,p)\\ &\le \left|\mathrm{Re}\left[\left(-\mathbf{u}+\frac{1}{k}\nabla p, \frac{1}{k}\nabla p\right)\right]\right| + \left|\mathrm{Re}\left[\left(-\frac{1}{k}\nabla\cdot\mathbf{u}-p, p\right)\right]\right| + (p,p)\\ &\le \left|\left(-\mathbf{u}+\frac{1}{k}\nabla p, \frac{1}{k}\nabla p\right)\right| + \left|\left(\frac{1}{k}\nabla\cdot\mathbf{u}+p, p\right)\right| \frac{\|p\|_{1,k}}{\|p\|_{1,k}} + \|p\|^2\\ &\le \left\|-\mathbf{u}+\frac{1}{k}\nabla p\right\| \left\|\frac{1}{k}\nabla p\right\| + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|_{-1,d}\|p\|_{1,k} + \|p\|^2\\ &\le \frac{1}{2\epsilon}\left\|-\mathbf{u}+\frac{1}{k}\nabla p\right\|^2 + \epsilon\left\|\frac{1}{k}\nabla p\right\|^2 + \frac{1}{2\epsilon}\left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|^2_{-1,d}\\ &\qquad + \left(1+\frac{\epsilon}{2}\right)\|p\|^2.\end{aligned}$$

Thus,

$$\frac{1}{k^2}\|\nabla p\|^2 \le C\left(\left\|-\mathbf{u}+\frac{1}{k}\nabla p\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|^2_{-1,d} + \|p\|^2\right),$$

so bound (3.5) holds on all of $\mathcal{W}(\Omega)$.

Now, to prove that

$$\begin{aligned}\|\mathbf{u}\|^2 + \|p\|^2_{1,k} + \|\mathbf{n}\cdot\mathbf{u}-\imath p\|^2_{-\frac{1}{2},\Gamma_e} \le C_k\Bigg(\left\|\mathbf{u}-\frac{1}{k}\nabla p\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}+p\right\|^2_{-1,d}\\ + \|\mathbf{n}\cdot\mathbf{u}-\imath p\|^2_{-\frac{1}{2},\Gamma_e}\Bigg) \qquad (3.9)\end{aligned}$$

on $\mathcal{W}(\Omega)$ by contradiction, suppose that (3.9) does not hold. Then there exists a sequence $\{(\mathbf{u}_j, p_j)\} \subset \mathcal{W}(\Omega)$ such that

$$\|\mathbf{u}_j\|^2 + \|p_j\|_{1,k}^2 + \|\mathbf{n}\cdot\mathbf{u}_j - \imath p_j\|^2_{-\frac{1}{2},\Gamma_e} = 1$$

and

$$\left\|-\mathbf{u}_j + \frac{1}{k}\nabla p_j\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}_j + p_j\right\|^2_{-1,d} + \|\mathbf{n}\cdot\mathbf{u}_j - \imath p_j\|^2_{-\frac{1}{2},\Gamma_e} \le \frac{1}{j}, \tag{3.10}$$

$j = 1, 2, \ldots$. Now, since $\|p_j\|^2_{1,k} \le 1$, then the Rellich selection theorem (cf. [7]) implies that $p_j \to p$ for some p in $L^2(\Omega)$, which, with (3.5) and (3.10), shows that $\{(\mathbf{u}_j, p_j)\}$ is a Cauchy sequence in the $\|\mathbf{u}\| + \|p\|_{1,k} + \|\mathbf{n}\cdot\mathbf{u} - \imath p\|_{-\frac{1}{2},\Gamma_e}$ norm, so it has a limit $\{(\mathbf{u}, p)\} \in \mathcal{W}(\Omega)$. Also, consider the scaled Helmholtz equation

$$\begin{cases} -\frac{1}{k^2}\nabla\cdot\nabla p - p &= \frac{f}{k^2} \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \Gamma_i, \\ \mathbf{n}\cdot\nabla p - \imath k p &= 0 \quad \text{on } \Gamma_e, \end{cases} \tag{3.11}$$

with the corresponding bilinear form

$$B(p,q) = \frac{1}{k^2}(\nabla p, \nabla q) - (p,q) - \frac{i}{k}\langle p, q\rangle, \qquad q \in H^1_{0,d}(\Omega).$$

Integration by parts gives

$$\begin{aligned} B(p,q) &= \frac{1}{k^2}(\nabla p, \nabla q) - (p,q) - \frac{i}{k}\langle p, q\rangle \\ &= \left(-\mathbf{u} + \frac{1}{k}\nabla p, \frac{1}{k}\nabla q\right) + \left(\mathbf{u}, \frac{1}{k}\nabla q\right) - (p,q) - \frac{i}{k}\langle p, q\rangle \\ &= \left(-\mathbf{u} + \frac{1}{k}\nabla p, \frac{1}{k}\nabla q\right) + \left(-\frac{1}{k}\nabla\cdot\mathbf{u} - p, q\right). \end{aligned} \tag{3.12}$$

Convergence of p_j to p in the $\|\cdot\|_{1,k}$ norm, the Cauchy–Schwarz and ϵ-inequalities, and (3.10) then imply that

$$\begin{aligned} |B(p,q)| &= \lim_{j\to\infty} |B(p_j, q)| \\ &\le \lim_{j\to\infty} \left(\left\|\mathbf{u}_j - \frac{1}{k}\nabla p_j\right\| \left\|\frac{1}{k}\nabla q\right\| + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}_j + p_j\right\|_{-1,d} \|q\|_{1,k} \right. \\ &\qquad \left. + \|\mathbf{n}\cdot\mathbf{u}_j - \imath p_j\|_{-\frac{1}{2},\Gamma_e}\|q\|_{1,k} \right) \\ &\le C\|q\|_{1,k} \\ &\quad \lim_{j\to\infty} \sqrt{\left\|\mathbf{u}_j - \frac{1}{k}\nabla p_j\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}_j + p_j\right\|^2_{-1,d} + \|\mathbf{n}\cdot\mathbf{u}_j - \imath p_j\|^2_{-\frac{1}{2},\Gamma_e}} \\ &= 0, \end{aligned} \tag{3.13}$$

or, equivalently, $p = 0$ because of the solution uniqueness of (3.11). Finally, by (3.5), it follows that

$$\begin{aligned}
\|\mathbf{u}\|^2 + \|\mathbf{n}\cdot\mathbf{u} - \imath p\|^2_{-\frac{1}{2},\Gamma_e} &= \lim_{j\to\infty}\left[\|\mathbf{u}_j\|^2 + \|\mathbf{n}\cdot\mathbf{u}_j - \imath p_j\|^2_{-\frac{1}{2},\Gamma_e}\right] \\
&\le C \lim_{j\to\infty}\left[\left\|\mathbf{u}_j - \frac{1}{k}\nabla p_j\right\|^2 + \left\|\frac{1}{k}\nabla\cdot\mathbf{u}_j + p_j\right\|^2_{-1,d}\right. \\
&\qquad\qquad \left. + \|p_j\|^2 + \|\mathbf{n}\cdot\mathbf{u}_j - \imath p_j\|^2_{-\frac{1}{2},\Gamma_e}\right] \\
&= C \lim_{j\to\infty}\frac{1}{j} \\
&= 0
\end{aligned}$$

so that

$$\|\mathbf{u}\|^2 + \|p\|^2_{1,k} + \|\mathbf{n}\cdot\mathbf{u} - \imath p\|^2_{-\frac{1}{2},\Gamma_e} = 0,$$

which is a contradiction. Thus, (3.9) must hold. The theorem now follows by restricting this bound to the subspace $\mathbf{W}(\Omega)$. ∎

The analogous result for F is essentially obtained using a Helmholtz decomposition of $\mathbf{u}$. Its proof is very similar to the proofs of Thorems 3.5 and 4.5 of [21], and, thus, will not be included here.

THEOREM 3.2. *There exist constants C and C_k such that*

$$\frac{1}{C_k}\left[\|\mathbf{u}\|^2_{1,k} + \|p\|^2_{1,k}\right] \le F(\mathbf{u},p;0) \le C\left[\|\mathbf{u}\|^2_{1,k} + \|p\|^2_{1,k}\right]$$

for all $(\mathbf{u},p) \in \mathbf{W}(\Omega)$.

These theorems show that the homogeneous parts of functionals F and G are bounded (with constant C) and coercive (with constant $\frac{1}{C_k}$) in $\mathbf{W}(\Omega)$. Unfortunately, the coercivity constant depends on the wavenumber k, which means that standard numerical solution processes can degrade in performance as k increases. In the next section, a subspace of $\mathbf{W}(\Omega)$ is introduced on which we prove the coercivity constant to be independent of k.

4. Uniform coercivity: $\mathbf{Z_\epsilon(\Omega)}$. A fundamental problem with the Helmholtz equation is that approximate Fourier components with wavenumbers near k produce relatively small residuals; that is, oscillatory components are in the *near nullspace* of (2.1). This degradation of the usual sense of ellipticity is also present in system (2.3), which means that the homogeneous parts of respective functionals G and F are not equivalent to the square of the product norms $[|||\cdot|||]^2 \times \|\cdot\|_{1,k}$ and $[|||\cdot\|_{1,k}]^3$ uniformly in k. However, as we will see, these troublesome functions can be treated specially in the multigrid coarsening and finite element approximation processes. To guide the design of these processes, this section will establish equivalence results on a subspace of $\mathbf{W}(\Omega)$ that excludes these near-nullspace components. To this end, note that

$$H^2_N(\Omega) := \left\{q \in H^2(\Omega) :\ q = 0 \text{ on } \Gamma_i, \mathbf{n}\cdot\nabla q = 0 \text{ on } \Gamma_e\right\}$$

is a dense subspace of $H^1_{0,d}(\Omega)$ (see [6]) that corresponds to the range of the inverse Laplace operator for

$$\left\{\begin{array}{rcl} -\nabla\cdot\nabla q &= f & \text{in } \Omega,\\ q &= 0 & \text{on } \Gamma_i,\\ \mathbf{n}\cdot\nabla q &= 0 & \text{on } \Gamma_e, \end{array}\right. \tag{4.1}$$

where $f \in L^2(\Omega)$ (that is, $[-\nabla\cdot\nabla]^{-1} : L^2(\Omega) \to H^2_N(\Omega)$). The operator in (4.1) is self-adjoint, so there exists an $L^2(\Omega)$-orthonormal eigensystem $\{(q_j,\lambda_j)\} \subset H^2_N(\Omega)\times\Re$ with $\lim_{j\to\infty}\lambda_j = \infty$:

$$\left\{\begin{array}{rcl} -\nabla\cdot\nabla q_j &= \lambda_j q_j & \text{in } \Omega,\\ q_j &= 0 & \text{on } \Gamma_i,\\ \mathbf{n}\cdot\nabla q_j &= 0 & \text{on } \Gamma_e. \end{array}\right. \tag{4.2}$$

To isolate the near-nullspace components of $H^1_{0,d}(\Omega)$, let $\epsilon \in (0,1)$ be given and define the spaces

$$\begin{aligned} S &:= \text{span}\left\{q_j\right\},\\ S_\epsilon &:= \text{span}\left\{q_j :\ \lambda_j^{\frac{1}{2}} \le \sqrt{\frac{1-\epsilon}{1+\epsilon}}\,k \ \text{ or } \ \lambda_j^{\frac{1}{2}} \ge \sqrt{\frac{1+\epsilon}{1-\epsilon}}\,k\right\},\\ \mathbf{Z}_\epsilon(\Omega) &:= \left\{(\mathbf{u},p)\in\mathbf{W}(\Omega) :\ p\in\bar{S_\epsilon}\right\}. \end{aligned}$$

Here we use *span* to mean *finite* linear combinations and overbar to denote closure in the $H^{1,k}$ norm.

The subspace $\mathbf{Z}_\epsilon(\Omega)$ of $\mathbf{W}(\Omega)$, which is closed in the $|||\cdot|||_k$ norm, avoids the oscillatory near-nullspace error components of system (2.3) that yield small residuals. Hence, in this subspace, uniform equivalence statements can be made between the least-squares norms corresponding to G and F (that is, the square roots of their homogeneous parts) and the respective product norms $[|||\cdot|||]^2 \times \|\cdot\|_{1,k}$ and $[|||\cdot|||_{1,k}]^3$.

THEOREM 4.1. *Let $\epsilon \in (0,1)$ be given. Then there exists a constant C, independent of ϵ and k, such that*

$$\frac{1}{C}\left[\epsilon^2\|\mathbf{u}\|^2 + \epsilon^2\|p\|^2_{1,k}\right] \le G(\mathbf{u},p;0) \le C\left[\|\mathbf{u}\|^2 + \|p\|^2_{1,k}\right]$$

for all $(\mathbf{u},p)\in\mathbf{Z}_\epsilon(\Omega)$.

Proof. The *upper bound* follows from that of Theorem 3.1 and the fact that $\mathbf{Z}_\epsilon(\Omega)\subset\mathbf{W}(\Omega)$.

The proof of the *lower bound* rests on the observation that $\mathbf{Z}_\epsilon(\Omega)$ is contained in the closure of

$$\mathbf{Z}_{0,\epsilon}(\Omega) := [\mathbf{H}_0(\text{div};\Omega)\cap\mathbf{H}_0(\mathbf{curl};\Omega)]\times S_\epsilon$$

in the product norm $[|||\cdot|||]^2\times\|\cdot\|_{1,k}$. A proof analogous to that of the upper bound in Theorem 3.1 shows that $G(\mathbf{u},p;0)$ is continuous on $\mathbf{Z}_\epsilon(\Omega)+\mathbf{Z}_{0,\epsilon}(\Omega)$ in this norm. It is therefore sufficient to establish the lower bound for $(\mathbf{u},p)\in\mathbf{Z}_{0,\epsilon}(\Omega)$. To this end, consider the Helmholtz decomposition (cf. [12])

$$\mathbf{u} = \frac{1}{k}\nabla q + \nabla^\perp\psi,$$

where $q \in H^2_N(\Omega)$, $\psi \in H^2(\Omega)$, and $\nabla^\perp \psi$ is divergence-free such that $\mathbf{n} \cdot \nabla^\perp \psi = 0$ on Γ_e. Substituting this decomposition into $G(\mathbf{u}, p; 0)$ and using the $L^2(\Omega)$ orthogonality property of this decomposition and the divergence-free property of $\nabla^\perp \psi$, we have

$$\begin{aligned} G(\mathbf{u}, p; 0) &= \left\| \nabla^\perp \psi + \frac{1}{k} \nabla (q - p) \right\|^2 + \left\| \frac{1}{k^2} \nabla \cdot \nabla q + \frac{1}{k} \nabla \cdot \nabla^\perp \psi + p \right\|^2_{-1,d} \\ &\quad + \left\| \frac{1}{k} \nabla \times \mathbf{u} \right\|^2_{-1} \\ &\geq \left\| \nabla^\perp \psi \right\|^2 + \left\| \frac{1}{k} \nabla (q - p) \right\|^2 + \left\| \frac{1}{k^2} \nabla \cdot \nabla q + p \right\|^2_{-1,d}. \end{aligned} \tag{4.3}$$

Also, integration by parts and the Cauchy–Schwarz inequality give

$$\begin{aligned} \left\| \frac{1}{k^2} \nabla \cdot \nabla (q - p) \right\|_{-1,d} &= \sup_{w \in H^1_{0,d}(\Omega)} \frac{\left(\frac{1}{k^2} \nabla \cdot \nabla (q - p), w \right)}{\|w\|_{1,k}} \\ &= \sup_{w \in H^1_{0,d}(\Omega)} \frac{\left(-\frac{1}{k^2} \nabla (q - p), \nabla w \right)}{\|w\|_{1,k}} \\ &\leq \left\| \frac{1}{k} \nabla (q - p) \right\|. \end{aligned} \tag{4.4}$$

Now, the triangle inequality and (4.4) yield

$$\left\| \frac{1}{k} \nabla (q - p) \right\|^2 + \left\| \frac{1}{k^2} \nabla \cdot \nabla p + p \right\|^2_{-1,d} \leq 3 \left\| \frac{1}{k} \nabla (q - p) \right\|^2 + 2 \left\| \frac{1}{k^2} \nabla \cdot \nabla q + p \right\|^2_{-1,d},$$

which gives

$$\left\| \frac{1}{k} \nabla (q - p) \right\|^2 + \left\| \frac{1}{k^2} \nabla \cdot \nabla q + p \right\|^2_{-1,d} \geq \frac{1}{2} \left\| \frac{1}{k^2} \nabla \cdot \nabla p + p \right\|^2_{-1,d}. \tag{4.5}$$

But $[-\nabla \cdot \nabla]$ is positive-definite and self-adjoint on S_ϵ and $p \in S_\epsilon$, so it is easy to verify the bound

$$\left\| \frac{1}{k^2} \nabla \cdot \nabla p + p \right\|^2_{-1,d} \geq \epsilon^2 \|p\|^2_{1,k}. \tag{4.6}$$

Hence, (4.3), (4.5), and (4.6) imply that

$$G(\mathbf{u}, p; 0) \geq \frac{\epsilon^2}{2} \|p\|^2_{1,k} \tag{4.7}$$

for any $(\mathbf{u}, p) \in \mathbf{Z}_{0,\epsilon}(\Omega)$. Finally, combining (4.7) with the triangle inequality and the restriction $\epsilon < 1$, we obtain

$$\begin{aligned} \epsilon^2 \|\mathbf{u}\|^2 &\leq C \left(\epsilon^2 \left\| \mathbf{u} - \frac{1}{k} \nabla p \right\|^2 + \epsilon^2 \left\| \frac{1}{k} \nabla p \right\|^2 \right) \\ &\leq C \left(\left\| \mathbf{u} - \frac{1}{k} \nabla p \right\|^2 + \left\| \frac{1}{k} \nabla \cdot \mathbf{u} + p \right\|^2_{-1,d} + \epsilon^2 \|p\|^2_{1,k} \right) \\ &\leq C G(\mathbf{u}, p; 0). \end{aligned} \tag{4.8}$$

The lower bound now follows from (4.7) and (4.8). ∎

We again state the analogous result for F without proof.

THEOREM 4.2. *Let $\epsilon \in (0,1)$ be given and $k \geq 1$. Then there exists a constant C, independent of ϵ and k, such that*

$$\frac{1}{C}\left[\epsilon^2\|\mathbf{u}\|_{1,k}^2 + \epsilon^2\|p\|_{1,k}^2\right] \leq F(\mathbf{u},p;0) \leq C\left[\|\mathbf{u}\|_{1,k}^2 + \|p\|_{1,k}^2\right]$$

for all $(\mathbf{u},p) \in \mathbf{Z}_\epsilon(\Omega)$.

Remark 4.1. While the coercivity constant in Theorem 4.2 is uniform in k, the dimension of $\mathbf{Z}_\epsilon^\perp(\Omega)$ grows with k (for fixed ϵ). However, these troublesome components can still be specially treated in the multigrid coarsening process without increasing the order of complexity, as we will show in the next section.

5. Nonstandard multigrid. Some of the basic concepts of this section were taken from or inspired by the work of Achi Brandt (cf. [9] and Brandt and Livshitz [10]).

The L^2 functional F, when it applies, is usually more practical than the $L^2 - H^{-1}$ functional G, which involves the somewhat cumbersome negative norms. For this reason, the remaining two sections will focus on minimizing $F(\mathbf{u},p;f)$. A Rayleigh–Ritz finite element method is used for the discretization process. Let $\mathcal{T}_h$ be a triangulation of domain Ω into finite elements of maximal length $h = \max\{\text{diam}(K) : K \in \mathcal{T}_h\}$, and let $\mathbf{W}_1$ be a finite-dimensional subspace of $\mathbf{W}(\Omega)$ having the approximation property

$$\inf_{(\mathbf{v}_1,q_1)\in\mathbf{W}_1}[\|\mathbf{v}-\mathbf{v}_1\|_1 + \|q-q_1\|_1] \leq Ch[\|\mathbf{v}\|_2 + \|q\|_2] \tag{5.1}$$

for $(\mathbf{v},q) \in \left[\left[H^2(\Omega)\right]^3 \cap \mathbf{W}(\Omega)\right]$. The discrete fine grid minimization problem is the following.

- Find $(\mathbf{u}_1,p_1) \in \mathbf{W}_1$ such that

$$F(\mathbf{u}_1,p_1;f) = \min_{(\mathbf{w}_1,q_1)\in\mathbf{W}_1} F(\mathbf{w}_1,q_1;f).$$

Equivalently, defining the bilinear form

$$\mathbf{D}(\mathbf{u},p;\mathbf{v},q) := \left(\mathbf{u} - \frac{1}{k}\nabla p, \mathbf{v} - \frac{1}{k}\nabla q\right) + \left(\frac{1}{k}\nabla\cdot\mathbf{u} + p, \frac{1}{k}\nabla\cdot\mathbf{v} + q\right) + \left(\frac{1}{k}\nabla\times\mathbf{u}, \frac{1}{k}\nabla\times\mathbf{v}\right)$$

for $(\mathbf{u},p),(\mathbf{v},q) \in \mathbf{W}(\Omega)$, the discrete fine grid variational problem is the following.

- Find $(\mathbf{u}_1,p_1) \in \mathbf{W}_1$ such that

$$\mathbf{D}(\mathbf{u}_1,p_1;\mathbf{v}_1,q_1) = \left(-\frac{f}{k^2}, \frac{1}{k}\nabla\cdot\mathbf{v}_1 + q_1\right)$$

  for all $(\mathbf{v}_1,q_1) \in \mathbf{W}_1$.

A standard multilevel scheme for solving either of these discrete problems is fairly straightforward. Let

$$\mathcal{T}_{2^{m-1}h} \subset \mathcal{T}_{2^{m-2}h} \subset \cdots \subset \mathcal{T}_{2h} \subset \mathcal{T}_h$$

be a conforming sequence of coarsenings of triangulation $\mathcal{T}_h$,

$$\mathbf{W}_m \subset \mathbf{W}_{m-1} \subset \cdots \subset \mathbf{W}_2 \subset \mathbf{W}_1$$

be a set of nested coarse grid subspaces of $\mathbf{W}_1$, the finest subspace, and

$$\mathbf{B}_j = \{(\mathbf{b}_j, c_j)_\nu : \ \nu = 1, 2, \ldots, n_j\}, \qquad j = 1, 2, \ldots, m,$$

be a suitable (generally local) basis for $\mathbf{W}_j$, referred to here as *level* j. Given an initial approximation $(\mathbf{u}_j, p_j)$ on level j, the level j relaxation sweep consists of the following cycle:

- For each $\nu = 1, 2, \ldots, n_j$, set

$$(\mathbf{u}_j, p_j) \leftarrow (\mathbf{u}_j, p_j) + \boldsymbol{\alpha}(\mathbf{b}_j, c_j)_\nu,$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ is chosen to minimize

$$F\left((\mathbf{u}_j, p_j) + \boldsymbol{\alpha}(\mathbf{b}_j, c_j)_\nu; f\right). \tag{5.2}$$

(By $\boldsymbol{\alpha}(\mathbf{b}, c)$ is meant $(\boldsymbol{\beta}, \gamma)$, where $\beta_1 = \alpha_1 b_1$, $\beta_2 = \alpha_2 b_2$, and $\gamma = \alpha_3 c$.) Since $F\left((\mathbf{u}_j, p_j) + \boldsymbol{\alpha}(\mathbf{b}_j, c_j)_\nu; f\right)$ is a quadratic function in $\boldsymbol{\alpha}$, this local minimization procedure is simple and very inexpensive (see [23] and [25]). In fact, this minimization process is just a block Gauss–Seidel iteration with blocks determined by the choice of $(\mathbf{b}_j, c_j)_\nu$; that is, recalling that L is the first-order operator corresponding to (2.3), the block system is expressed as

$$[A_j \boldsymbol{\alpha}]_\nu = -\left(L \begin{pmatrix} \mathbf{u}_j \\ p_j \end{pmatrix} - \mathbf{f}, L \begin{pmatrix} \mathbf{b}_j \\ c_j \end{pmatrix}_\nu\right), \tag{5.3}$$

where

$$[A_j]_{\mu,\nu} = \left(L \begin{pmatrix} \mathbf{b}_j \\ c_j \end{pmatrix}_\mu, L \begin{pmatrix} \mathbf{b}_j \\ c_j \end{pmatrix}_\nu\right). \tag{5.4}$$

Now, given a fine grid approximation $(\mathbf{u}_1, p_1)$ on level 1, the level 2 coarse grid problem is to find a correction $(\mathbf{u}_2, p_2) \in \mathbf{W}_2$ such that

$$F\left(\mathbf{u}_1 + \mathbf{u}_2, p_1 + p_2; f\right) = \min_{(\mathbf{v}_2, q_2) \in \mathbf{W}_2} F\left(\mathbf{u}_1 + \mathbf{v}_2, p_1 + q_2; f\right). \tag{5.5}$$

Having obtained $(\mathbf{u}_2, p_2)$, the fine grid approximation is corrected according to

$$(\mathbf{u}_1, p_1) \leftarrow (\mathbf{u}_1 + \mathbf{u}_2, p_1 + p_2).$$

Applying this process recursively yields a multilevel scheme in the usual way.

Unfortunately, if the coarser level basis functions are chosen in the natural way, then this coarse grid correction process may not be very effective for the FOSLS formulation of the Helmholtz equation. Difficulties can arise with error components for elements in $\mathbf{Z}_\epsilon^\perp$. Specifically, trouble arises from error components that have the approximate form

$$\begin{pmatrix} \imath \cos\theta \\ \imath \sin\theta \\ 1 \end{pmatrix} e^{\imath k(x\cos\theta + y\sin\theta)}, \tag{5.6}$$

where $\theta \in [0, 2\pi)$. The main problem is that, on a grid where $kh = O(1)$, error components of this type are highly oscillatory and, hence, poorly approximated by standard

coarse grids, yet they produce small residuals and are therefore poorly reduced by relaxation. A compounding problem is that there are, in principle, infinitely many of these components since θ can be virtually any angle in $[0, 2\pi)$. More precisely, as k increases, so does the dimension of $\mathbf{Z}_\epsilon^\perp$, which consists of elements that are approximately of the form in (5.6). Fortunately, these problems can be circumvented by introducing exponential interpolation and multiple coarsening into the multilevel scheme ([10] and [23]).

Exponential interpolation. Exponential interpolation is used to approximate the troublesome oscillatory error components of approximate form (5.6). But the *ray* $e^{\imath k(x\cos\theta+y\sin\theta)}$ is the problematic factor of this form, so a more useful characterization of these error components is given by

$$s(x,y)e^{\imath k(x\cos\theta+y\sin\theta)}, \tag{5.7}$$

where $s(x,y)$ is a smooth function satisfying appropriate boundary conditions.

Let W_j^γ be the components of the product space $\mathbf{W}_j$, $\gamma = 1,2,3$, which is assumed for concreteness to consist only of continuous piecewise bilinear functions. (These spaces generally differ over γ only at the boundary.) Also, let level j correspond to a grid that is fine enough relative to k that components (5.7) can be adequately approximated with bilinear functions, although only marginally so in the sense that this approximation just begins to deteriorate on level $j+1$. Now, exact coarse level correction from level j is assumed to be effective at eliminating the smooth components left by relaxation on level $j-1$. At level $j+1$, however, the ability of bilinear functions to approximate components (5.7) deteriorates enough to contaminate two-level performance between levels $j+1$ and j. Some heuristics show that this occurs when $2^j hk = O(1)$. To approximate these exponential components from levels $j+1, j+2, \ldots, m$, their basis functions are simply rescaled by the ray elements $e^{\imath k(x\cos\theta+y\sin\theta)}$. To be more specific, fix $\theta \in [0, 2\pi)$. For each $\gamma = 1,2,3$ and $l = j+1, \ldots, m$, let $d_{l,\nu}^\gamma$ denote entry number γ of element number ν of the standard basis $\mathbf{B}_l$ and, on level l, define the *level j ray* element by

$$\rho_{l,\nu}^{\theta,\gamma} = \mathcal{E}_j^\theta\left(d_{l,\nu}^\gamma\right),$$

where $\mathcal{E}_j^\theta(d)$ for $d \in W_l^\gamma$ is the element of W_j^γ that agrees with $e^{\imath k(x\cos\theta+y\sin\theta)}d$ at the level j node values; that is, $\mathcal{E}_j^\theta(d)$ is the level j nodal interpolant of $e^{\imath k(x\cos\theta+y\sin\theta)}d$. This has the effect of altering the usual bilinear interpolation formula that relates coarse to fine node values so that the rays are better approximated. In fact, this ensures that the level j interpolant of the given ray is in the range of interpolation on all coarser levels determined by the corresponding θ. Figure 5.1 sketches the real part of a one-dimensional level 3 basis function d_3 and a typical ray basis element $\mathcal{E}_1^\theta(d_3)$ determined by an oscillatory level 1 ray (assuming for simplicity that $j = 1$).

Note that exponential interpolation is not much more costly than bilinear interpolation: a given function d on level $l > j$ is bilinearly interpolated up to level j, then simply rescaled by $e^{\imath k(x\cos\theta+y\sin\theta)}$ at the level j nodes.

Multiple coarsening. Consider now the case that level j is where approximation of components (5.7) just begins to deteriorate. It may be enough to introduce one ray (e.g., with $\theta = 0$) on level $j+1$ in the sense that an exact correction from level $j+1$ might then adequately correct "smooth" level j error. However, it is perhaps more effective here to introduce a few coarse grids, one for each of, say, four rays (e.g., with

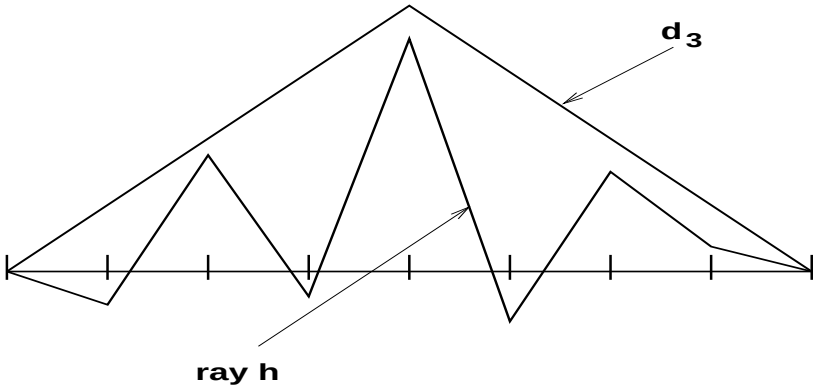

FIG. 5.1. *One-dimensional real parts of a typical standard d_3 basis function and a $\mathcal{E}_1^{\theta}(d_3)$ ray basis element.*

$\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$). More critically, as each level is coarsened, scale $k2^{j-1}h$ doubles and essentially twice as many rays become oscillatory. To be more specific, suppose that θ_1 and θ_2 are angles corresponding to two neighboring level j rays in the sense that these two rays were used to coarsen level $j+1$, and so perhaps others, but none were needed with angles between θ_1 and θ_2. Now, on any of the level $j+2$ grids (there will be one for each ray, that is, for each coarsening), the ray with angle $\frac{\theta_1+\theta_2}{2}$ will be poorly reduced in relaxation and poorly approximated by any of the level $j+3$ grids. This means that twice as many rays must somehow be used in coarsening to level $j+2$, which in turn will again require ray doubling to level $j+3$.

This can be accomplished effectively by introducing two separate ray coarsenings on a level that may itself have come from ray coarsening. Suppose that level $j+1$ was created from a level j ray with angle θ, and that the task is now to coarsen level $j+1$ using rays of angles $\theta \pm \phi$. One way to accomplish this is to introduce the level j ray $\rho_{j+2,\nu}^{\theta+\phi,\gamma} = \mathcal{E}_j^{\theta+\phi}\left(d_{j+2,\nu}^{\gamma}\right)$. But now $\rho_{j+2,\nu}^{\theta+\phi,\gamma}$ oscillates on level j scale, so applying an operator to this function will generally require level j evaluations. Alternatively, these ray coarsenings can be efficiently introduced using intermediate level $j+1$ rays. Suppose the level j rays $\rho_{j+1,\nu}^{\theta,\gamma}$ have already been introduced on level $j+1$. Since

$$e^{\imath k[x\cos(\theta+\phi)+y\sin(\theta+\phi)]} = e^{\imath k(x\cos\theta+y\sin\theta)}e^{\imath k\{x[\cos(\theta+\phi)-\cos\theta]+y[\sin(\theta+\phi)-\sin\theta]\}},$$

then letting $\mathcal{E}_j^{\theta,\phi}$ denote the level j interpolant of the ray perturbation

$$e^{\imath k\{x[\cos(\theta+\phi)-\cos\theta]+y[\sin(\theta+\phi)-\sin\theta]\}}$$

means that the perturbed rays

$$\sigma_{j+2,\nu}^{\phi,\gamma} = \mathcal{E}_j^{\theta}\left[\mathcal{E}_{j+1}^{\theta,\phi}\left(d_{j+2,\nu}^{\gamma}\right)\right]$$

can be computed directly from level $j+1$ without appealing to level j. In particular, since

$$\mathcal{E}_{j+1}^{\theta,\phi}(d_{j+2,\nu}^{\gamma}) \tag{5.8}$$

is a level $j+1$ piecewise bilinear function, then $\sigma_{j+2,\nu}^{\phi,\gamma}$ can be constructed by exponentially interpolating the level $j+1$ ray basis functions. (See Figure 5.2 for a sketch

of this process for creating a level 3 ray basis function from levels 1 and 2.) Hence, given an operator M and assuming that

$$\left\{M\rho_{j+1,\mu}^{\theta,\gamma} : \ \mu = 1, 2, \ldots, n_{j+1}\right\} = \left\{M\mathcal{E}_j^{\theta}(d_{j+1,\mu}^{\gamma}) : \ \mu = 1, 2, \ldots, n_{j+1}\right\} \tag{5.9}$$

has been constructed for the ray basis of level $j+1$, then $M\sigma_{j+2,\nu}^{\phi,\gamma}$ can be constructed by exponentially weighting the $M\rho_{j+1,\nu}^{\theta,\gamma}$'s:

$$\begin{aligned} M\sigma_{j+2,\nu}^{\phi,\gamma} &= M\left\{\mathcal{E}_j^{\theta}\left[\mathcal{E}_{j+1}^{\theta,\phi}(d_{j+2,\nu}^{\gamma})\right]\right\} \\ &= M\left\{\mathcal{E}_j^{\theta}\left[\sum_{\mu} c_\mu d_{j+1,\mu}^{\gamma}\right]\right\} \\ &= \sum_{\mu} c_\mu M\left\{\mathcal{E}_j^{\theta}(d_{j+1,\mu}^{\gamma})\right\} \\ &= \sum_{\mu} c_\mu M\rho_{j+1,\mu}^{\theta,\gamma}, \end{aligned}$$

which is a weighting of elements of (5.9) with exponential weights c_μ. Recursively applying this exponential weighting process, analogous coarser level calculations can be computed without appealing to level j.

Now, having solved a coarse grid problem, its solution can be exponentially interpolated up to level j using the angle perturbations in a recursive way. For example, if s_{j+2}^{ϕ} is the solution of the level $j+2$ coarse grid problem corresponding to perturbation ϕ, then s_{j+2}^{ϕ} can be interpolated up to level j by first exponentially interpolating up to level $j+1$ to obtain

$$\mathcal{E}_{j+1}^{\theta,\phi}(s_{j+2}^{\phi}),$$

then exponentially interpolating this result to level j using θ. Since exponential interpolation consists of a standard bilinear interpolation and an exponential scaling, then this successive interpolation process is easy to implement. Moreover, since the exponential interpolation operator is the adjoint of the exponential weighting operator, then exponential weighting is also easy to implement.

Angles (ϕ). The angles cannot be arbitrary, but must be chosen such that an *optimal* number of exponential components are *adequately* approximated by the perturbed ray elements. To describe this selection procedure, assume that rays are first introduced on level $j+1$. Then, to obtain adequate approximation, a simple but tedious analysis [21] shows that the perturbation angle ϕ must be chosen so that the following approximation is within the level j discretization error:

$$\mathcal{E}_j^{\theta}\left[\mathcal{E}_{j+1}^{\theta,\phi}\left(d_{j+2,\nu}^{\gamma}\right)\right] \approx \mathcal{E}_j^{\theta+\phi}\left(d_{j+2,\nu}^{\gamma}\right).$$

For example, for $j = 1$ with level 2 angles $\theta = 0, \frac{\pi}{2}, \pi$, and $\frac{3\pi}{2}$, the level 3 angle perturbations must be $\phi = \pm\frac{\pi}{8}$, as depicted in Figure 5.4. The angle perturbations are simply halved at each successively finer level.

Nonstandard multigrid scheme and computational cost. Exponential interpolation/weighting and multiple coarsening are essentially all that is required to

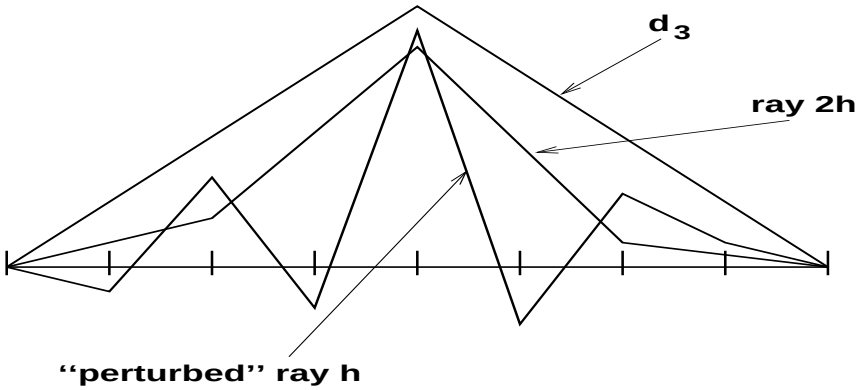

FIG. 5.2. *Level* 3 *ray element determined from a perturbation corresponding to angle* ϕ *of the level* 2 *ray element.*

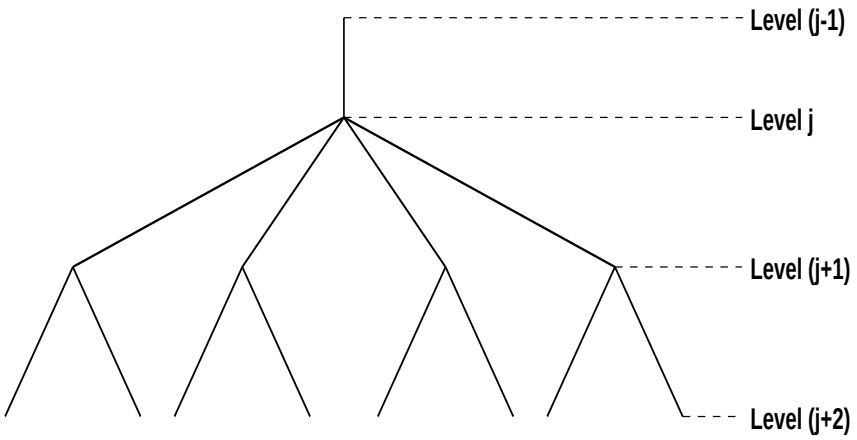

FIG. 5.3. *Multiple coarsening.*

develop an effective nonstandard multigrid scheme for the Helmholtz equation. With these components, ray basis elements are naturally introduced, and this nonstandard multigrid scheme has the same form as the standard multigrid scheme defined by (5.2) and (5.5). To see this, again assume that rays are first introduced on level $j+1$ and that all perturbation angles are determined before the multigrid cycling begins. Here, we choose j to be the largest integer satisfying

$$2^{j-1}kh \leq \frac{1}{8},$$

since this guarantees essentially 8 grid points per wavelength in the piecewise bilinear approximation on level j. Now, for levels $l \leq j$, the nested spaces

$$\mathbf{W}_l \subset \mathbf{W}_{l-1} \subset \cdots \subset \mathbf{W}_1$$

still consist of products of piecewise bilinear functions, but, for levels $l \geq j+1$, we must introduce the product ray element spaces

$$\mathbf{W}_l^{\theta_r} := \text{span}\left\{\boldsymbol{\sigma}_{l,\nu}^{\theta_r} : \ \nu = 1, 2, \ldots, n_l\right\}, \qquad r = 1, \ldots, t_l,$$

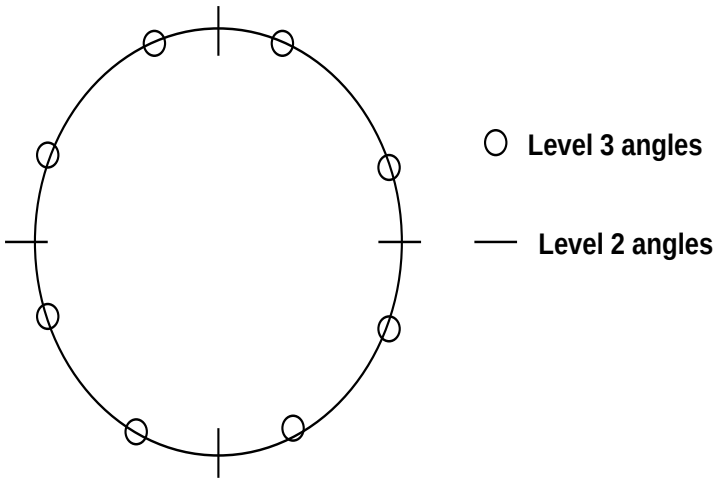

FIG. 5.4. *Angle choices (uniformly distributed) for level* 2 *and level* 3 *rays.*

where t_l is the total number of rays introduced on level l, θ_r is the angle of the rth ray, and $\left\{\boldsymbol{\sigma}_{l,\nu}^{\theta_r} : \nu = 1, 2, \ldots, n_l\right\}$ is the corresponding perturbed vector ray basis introduced on level l. We first introduce four rays on level $j+1$ and double that number for each successively coarser level, so we have $t_l = 2^{l+1-j}$. Note that

$$\mathbf{W}_{l+1}^{\theta_s} \subset \mathbf{W}_l^{\theta_r}$$

only if θ_s is a perturbation of angle θ_r (that is, only if $s = 2r-1$ or $2r$). As before, for levels $l \leq j$, relaxation is defined by (5.2) with $(\mathbf{b}_l, c_l)_\nu$ the standard basis for $\mathbf{W}_l$ and, for levels $l < j$, the level $(l+1)$ coarse grid problem is to find $(\mathbf{u}_{l+1}, p_{l+1}) \in \mathbf{W}_{l+1}$ such that

$$F\left(\mathbf{u}_l + \mathbf{u}_{l+1}, p_l + p_{l+1}; f\right) = \min_{(\mathbf{v}_{l+1}, q_{l+1}) \in \mathbf{W}_{l+1}} F\left(\mathbf{u}_l + \mathbf{v}_{l+1}, p_l + q_{l+1}; f\right),$$

where $(\mathbf{u}_l, p_l)$ is the level l approximation. However, for the ray levels $l > j$, relaxation is again defined by (5.2); but now with $(\mathbf{b}_l, c_l)_\nu$ the product ray basis element $\boldsymbol{\sigma}_{l,\nu}^{\theta_r}$, and, for levels $l \geq j$, the two level $(l+1)$ coarse grid problems corresponding to each θ_r, are to find $(\mathbf{u}_{l+1}, p_{l+1}) \in \mathbf{W}_{l+1}^{\theta_s}$, $s = 2r-1$, and $s = 2r$ such that

$$F\left(\mathbf{u}_l + \mathbf{u}_{l+1}, p_l + p_{l+1}; f\right) = \min_{(\mathbf{v}_{l+1}, q_{l+1}) \in \mathbf{W}_{l+1}^{\theta_s}} F\left(\mathbf{u}_l + \mathbf{v}_{l+1}, p_l + q_{l+1}; f\right),$$

where $(\mathbf{u}_l, p_l) \in \mathbf{W}_l^{\theta_r}$ is now the approximation on level l with angle θ_r. Note that each level l with angle θ_r is corrected by two level $(l+1)$ problems (actually, four when $l = j$). This gives the structure of the multigrid process. The scheduling that determines how the various levels are visited is best described by studying the schematic in Figure 5.5. We will continue to refer to this process as a V-cycle, although it looks very different from a standard V-cycle below level j.

For this nonstandard multigrid scheme (which can be interpreted as a multilevel multiplicative Schwarz method; cf. [28]), the level j coarse grid problem involves subspace corrections from the standard finite element space $\mathbf{W}_j$ and the ray subspaces $\mathbf{W}_l^{\theta_r}, l > j$. Also, relaxation on the ray spaces is simply a block Gauss–Seidel iteration

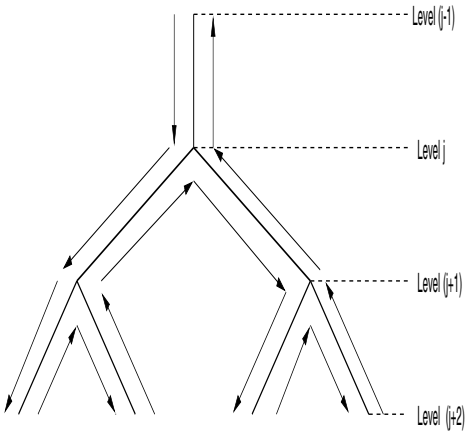

FIG. 5.5. *V-cycle for the nonstandard multigrid scheme.*

(5.3) with $(\mathbf{b}_l, c_l)_\nu = \boldsymbol{\sigma}_{l,\nu}^{\theta_r}$. Since $L\boldsymbol{\sigma}_{l,\nu}^{\theta_r}$ can be computed efficiently using exponential weighting, so can the elemental block matrices or stencils

$$[A_l]_{\mu,\nu}^{\theta_r} = \left(L\boldsymbol{\sigma}_{l,\mu}^{\theta_r}, L\boldsymbol{\sigma}_{l,\nu}^{\theta_r}\right).$$

In fact, these stencils can be computed and stored prior to any multigrid cycling so that the same relaxation routine can be used for all levels and angles.

Now, assuming that the stencils have been computed and stored before any multigrid cycling occurs, the computational cost of such a multiple coarsening algorithm is generally not excessive. In fact, one such multilevel V-cycle requires only several fine grid work units (that is, the cost of a fine grid relaxation). For example, consider the worst case, when four rays are introduced on level 2. (The need for more rays would signal a level h so coarse relative to k that the discretization accuracy would be almost meaningless; in fact, the need for four rays on level 2 is already a sign that level 1 accuracy is quite poor.) Then, for level $j \geq 2$, there would be a total of 2^j level j problems. Also, for each j, if it is assumed that the stepsize is doubled at each coarsening, then the number of fine grid work units required for relaxation is 2^{-2j+2}. Omitting the cost of ray interpolation (which is substantial but proportional to the number of points) and other more minor operations, the total number of fine level work units for a nonstandard $V(1,0)$-cycle is then

$$\begin{aligned}
&1 + \left[2^2\left(2^{-2}\right) + 2^3\left(2^{-4}\right) + \cdots + 2^j\left(2^{-2j+2}\right) + \cdots + 2^m\left(2^{-2m+2}\right)\right] \\
&\quad = 1 + \left[1 + 2^{-1} + 2^{-2} + \cdots + 2^{-j+2} + \cdots + 2^{-m+2}\right] \\
&\quad = 1 + \left[2 - 2^{-m+2}\right] \\
&\quad \approx 3,
\end{aligned}$$

which is only about double the cost of a standard $V(1,0)$-cycle.

6. Numerical examples. For simplicity and to facilitate comparisons, the examples treated here are only for the unit square. The lack of a Dirichlet boundary is in fact a more severe test of the methodology in some sense. The results will include multigrid performance for both the standard and the nonstandard schemes described

TABLE 6.1
Standard coarsening: $V(1,0)$-cycles.

$k\backslash h$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$
1	.40	.44	.46	.45
2	-	.52	.53	.50
4	-	-	.78	.78
8	-	-	-	.94
16	-	-	-	.96

TABLE 6.2
Standard coarsening: two-level cycles.

$k\backslash h$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
1	.31	.34	.36	.36	.36
2	.31	.34	.36	.36	.36
4	.38	.34	.37	.37	.37
8	.66	.49	.37	.37	.37
16	.73	.73	.63	.42	.37

in the previous section. We also assess accuracy of the discretization based on ray-type elements.

To facilitate assessment of the iterative solver, we first restrict ourselves to the *homogeneous* boundary value problem (2.2) with $f = 0$. The exact solution is of course $p = 0$, but this is useful for measuring asymptotic algebraic convergence factors of stationary linear iterative methods such as multigrid methods since it avoids the limiting stagnation caused by machine representation. The unit square is partitioned into rectangles, and the functional F is minimized over the space of piecewise continuous bilinear functions (possibly involving rays) that satisfy the boundary condition $\mathbf{n} \cdot \mathbf{u} - \imath p = 0$ on Γ. In most cases, the finest grid is chosen to satisfy $kh = \frac{1}{8}$, and rays are first introduced on level $j+1$ when $k2^j h > \frac{1}{8}$ (that is, typically when $j = 1$). The relaxation scheme is nodal Gauss–Seidel relaxation (that is, block Gauss–Seidel where a block consists of all of the unknowns $(\mathbf{u}_\nu^h, p_\nu^h)$ at each node ν). To assess the worst-case convergence factors, twenty multigrid V(1,0)-cycles (that is, one relaxation before and none after coarsening) were performed starting from a random initial guess $(\mathbf{u}_h^{(0)}, p_h^{(0)})$. The convergence measure ω is then defined as

$$\omega = \max_{1 \le n \le 20} \sqrt{\frac{F(\mathbf{u}_h^{(n)}, p_h^{(n)}; 0)}{F(\mathbf{u}_h^{(n-1)}, p_h^{(n-1)}; 0)}},$$

where the superscript in parentheses denotes the iteration number. Tables 6.1 and 6.2, respectively, summarize the results of twenty $V(1,0)$- and two-level cycles for standard bilinear elements. Table 6.1 confirms that convergence degrades as a function of k due to the troublesome oscillatory error components present on the coarser levels. Table 6.2 shows that Poisson-like (that is, $k = 0$) factors are obtained when kh is not too large and, comparing it with Table 6.1, indicates that rays should be introduced as soon as $kh > \frac{1}{8}$.

Tables 6.3 and 6.4 show that multigrid convergence factors improve substantially by introducing ray functions. Here, "exact" rays (that is, $\mathcal{E}_j^\theta(d_{l,\nu}), l > j$) are intro-

TABLE 6.3
Exact ray ($\rho_{l,\nu}^{\theta,\gamma}$) coarsening: $V(1,0)$-cycles.

$k\backslash h$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$
1	.30	.32	.32	.32
2	-	.33	.33	.33
4	-	-	.39	.38
8	-	-	-	.39
16	-	-	-	.40

TABLE 6.4
Exact ray ($\rho_{l,\nu}^{\theta,\gamma}$) coarsening: two-level cycles.

$k\backslash h$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$
1	.30	.33	.33	.3
2	.31	.34	.34	.34
4	.33	.35	.38	.37
8	.39	.35	.38	.38
16	.61	.46	.38	.39

TABLE 6.5
Approximate ray ($\sigma_{l,\nu}^{\phi,\gamma}$) coarsening: $V(1,0)$-cycles.

$k\backslash h$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
1	.39	.44	.45	.45
2	.34	.39	.44	.46
4	-	.39	.42	.45
8	-	-	.40	.43
16	-	-	-	.40

TABLE 6.6
Approximate ray ($\sigma_{l,\nu}^{\phi,\gamma}$) coarsening, delayed until $kh > \frac{1}{4}$: $V(1,0)$-cycles.

$k\backslash h$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
1	.45	.46	.45	.45
2	.40	.45	.45	.45
4	-	.40	.46	.47
8	-	-	.44	.55
16	-	-	-	.55

duced when $kh > \frac{1}{8}$ using the algorithm described in the previous section. These factors are fairly uniform in k and compare favorably to the two-level factors for standard coarsening. Moreover, they degrade only slightly when the exact rays are approximated using the recursive scheme described in the previous section (see Table 6.5).

To test the effect of introducing the rays late in the multigrid solver, we reran these examples but delayed the use of rays until $kh > \frac{1}{2}$. As expected, the rates degrade noticeably, as Tables 6.6 and 6.7 show.

As a measure of discretization accuracy, we ran tests to see if our FOSLS scheme has any pollution effects. We specified the exact solution $p = e^{\imath k\left(x\cos\frac{\pi}{5}+y\sin\frac{\pi}{5}\right)}$, which solves (2.2) with $f = 0$, except that the boundary condition is inhomogeneous:

$$\mathbf{n}\cdot\nabla p - \imath k p = \mathbf{n}\cdot\begin{pmatrix} \imath k\cos\frac{\pi}{5} \\ \imath k\sin\frac{\pi}{5} \end{pmatrix} e^{\imath k\left(x\cos\frac{\pi}{5}+y\sin\frac{\pi}{5}\right)} - \imath k e^{\imath k\left(x\cos\frac{\pi}{5}+y\sin\frac{\pi}{5}\right)} \quad \text{on } \Gamma.$$

TABLE 6.7
Approximate ray $(\sigma_{l,\nu}^{\phi,\gamma})$ coarsening, delayed until $kh > \frac{1}{2}$: $V(1,0)$-cycles.

$k\backslash h$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
1	.44	.46	.45	.45
2	.48	.50	.48	.47
4	-	.47	.48	.47
8	-	-	.58	.54
16	-	-	-	.79

TABLE 6.8
Discretization error in the least-squares functional norm.

$kh\backslash h$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
$\frac{1}{16}$	$1.89e-2$	$1.89e-2$	$1.89e-2$	$1.89e-2$
$\frac{1}{32}$	$9.44e-3$	$9.44e-3$	$9.44e-3$	$9.44e-3$
$\frac{1}{64}$	$4.72e-3$	$4.72e-3$	$4.72e-3$	$4.72e-3$

TABLE 6.9
Relative discretization error in the scaled $H^{1,k}$ norm.

$kh\backslash h$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
$\frac{1}{16}$	$9.46e-3$	$9.46e-3$	$9.55e-3$	$1.05e-2$
$\frac{1}{32}$	$4.72e-3$	$4.72e-3$	$4.74e-3$	$4.73e-3$
$\frac{1}{64}$	$2.36e-3$	$2.36e-3$	$2.36e-3$	$2.36e-3$

We then used our FOSLS scheme to approximate p and the scaled gradient $\mathbf{u} = \frac{1}{k}\nabla p$. To isolate possible pollution, we chose various h, then varied k so that kh is constant ($kh = 1/16, 1/32, 1/64$). The errors were measured by comparing the discrete approximation (obtained after applying several multigrid cycles) to the exact solution in a relative sense in the least-squares and product $H^{1,k}$ norms. Tables 6.8 and 6.9 essentially show constant discretization errors for constant kh, with about a factor of 2 decrease as kh is halved, which is consistent with the assertion that our approach is pollution-free and $O(kh)$, respectively.

REFERENCES

[1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] S. AGMON, A. DOUGLIS, AND L. NIRENBERG, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions* II, Comm. Pure Appl. Math., 17 (1963), pp.35–92.

[3] I.M. BABUSKA, F. IHLENBURG, E.T. PAIK, AND S.A. SAUTER, *A generalized finite element method for solving the Helmholtz equation in two dimensions with minimal pollution*, Comput. Methods Appl. Mech. Engrg., 128 (1995), pp. 325–359.

[4] I. BABUSKA AND J. M. MELENK, *The partition of unity method*, Internat. J. Numer. Methods Engrg., 40 (1997), pp. 727–758.

[5] I.M. BABAUSKA AND S.A. SAUTER, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers*, SIAM J. Numer. Anal., 34 (1997), pp. 2392–2423.

[6] J.M. BEREZANSKII, *Expansions in Eigenfunctions of Selfadjoint Operators*, American Mathematical Society, Providence, RI, 1968.

[7] D. BRAESS, *Finite Elements Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, UK, 1997.

[8] J.H. Bramble, R.D. Lazarov, and J.E. Pasciak, *A least-squares approach based on a discrete minus one inner product for first order systems*, Math. Comp., 66 (1997), pp 935–955.
[9] A. Brandt, *Stages in developing multigrid solutions*, in Proceedings of the Second International Congress on Numerical Methods for Engineering, E. Absi, R. Glowinski, P. Lascaux, and H. Veysseyre, eds., Dunod, Paris, 1980, pp. 23–43.
[10] A. Brandt and I. Livshitz, *Ray-wave multigrid method for Helmholtz equation*, Electron. Trans. Numer. Anal., 6 (1997), pp. 162–181.
[11] Z. Cai, R. Lazarov, T.A. Manteuffel, and S.F. McCormick, *First-order system least squares for second-order partial differential equations: Part* I, SIAM J. Numer. Anal., 31 (1994), pp. 1785–1799.
[12] Z. Cai, T.A. Manteuffel, and S.F. McCormick, *First-order system least squares for second-order partial differential equations: Part* II, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.
[13] C. Chang, *Finite element method for the solution of Maxwell's equations in multiple media*, Appl. Math. Math. Comput., 25 (1988), pp. 89–99.
[14] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, Berlin, 1992.
[15] L.C. Evans, *Partial Differential Equations*, Vol. 3b, Berkeley Mathematics Lecture Notes, University of California, Berkeley, 1994.
[16] G.J. Fix and M.D. Gunzburger, *On numerical methods for acoustic problems*, Comput. Math. Appl., 6 (1980), pp. 265–278.
[17] V. Girault and P.A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer-Verlag, New York, 1979.
[18] J. Haslinger and P. Neittaanamäki, *On different finite element methods for approximating the gradient of the solution to the Helmholtz equation*, Comput. Methods Appl. Mech. Engrg., 42 (1984), pp. 131–148.
[19] F. Ihlenburg and I.M. Babuska, *Finite element solution of the Helmholtz equation with high wave number.* I. *The h-version of the FEM*, Comput. Math. Appl., 39 (1995), pp. 9–37.
[20] M. Křížek and P. Neittaanamäki, *On time-harmonic Maxwell's equations with nonhomogeneous conductivities: Solvability and FE-approximation*, Appl. Math., 6 (1989), pp. 480–499.
[21] B. Lee, *First-Order Systems Least Squares for Electromagnetics*, Ph.D. thesis, University of Colorado at Boulder, Boulder, 1996.
[22] B. Lee, *First-order systems least squares for elliptic problems with Robin boundary conditions*, SIAM J. Numer. Anal., to appear.
[23] B. Lee, T. Manteuffel, S. McCormick, and J. Ruge, *Multilevel first-order system least-squares (FOSLS) for Helmholtz equations*, manuscript, 1993.
[24] J.S. Li, Z.Y. Yu, X.Q. Xiang, W.P. Ni, and C.L. Chang, *Least-square finite element method for electromagnetic fields in* 2-D, Appl. Math. Comput., 58 (1993), pp. 143–167.
[25] S. McCormick, *Multilevel Projection Methods for Partial Differential Equations*, SIAM, Philadelphia, 1992.
[26] P.B. Monk, *A mixed method for approximating Maxwell's equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1610–1634.
[27] J. Saranen, *On electric and magnetic problems for vector fields in anisotropic nonhomogeneous media*, J. Math. Anal. Appl., 91 (1983), pp. 254–275.
[28] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.