

A SADDLE POINT NUMERICAL METHOD FOR HELMHOLTZ EQUATIONS*

Russell B. Richins

Department of Mathematics and Computer Science, Thiel College, Pennsylvania, USA

Email: rrichins@thiel.edu

Abstract

In a previous work, the author and D.C. Dobson proposed a numerical method for solving the complex Helmholtz equation based on the minimization variational principles developed by Milton, Seppecher, and Bouchitté. This method results in a system of equations with a symmetric positive definite coefficient matrix, but at the same time requires solving simultaneously for the solution and its gradient. Herein is presented a method based on the saddle point variational principles of Milton, Seppecher, and Bouchitté, which produces symmetric positive definite systems of equations, but eliminates the necessity of solving for the gradient of the solution. The result is a method for a wide class of Helmholtz problems based completely on the Conjugate Gradient algorithm.

Mathematics subject classification: Primary 65N30, Secondary 35A15.

Key words: Helmholtz, Conjugate gradient, Saddle point, Finite element.

1. Introduction

The Helmholtz equation

$$\nabla \cdot L \nabla u = Mu,$$

is useful in modeling wave propagation in problems arising from many different physical situations. We will focus only on the homogeneous equation for simplicity and brevity, but the methods presented here can easily be extended to the non-homogeneous case. Suppose we wish to solve the Helmholtz equation in a domain $\Omega \subset \mathbb{R}^d$, and assume that L and M are complex-valued functions. A common source of numerical methods for solving this equation is the variational principle

$$\int_{\Omega} [-L \nabla u \cdot \nabla \bar{v} - Mu \bar{v}] dx = 0, \quad \forall v \in H_0^1(\Omega). \quad (1.1)$$

Since this is a stationary principle, the resulting system of equations is often indefinite, and indefinite systems are generally more difficult to solve than a system of equations having a positive definite coefficient matrix.

Because of the challenges in solving these indefinite systems, there has been much work devoted to solving the Helmholtz equation by replacing the indefinite systems with equivalent symmetric positive definite linear systems. Classical examples of such approaches are the CGNR and CGNE methods [1], based on solving normal equations associated with the original system. While such approaches produce positive definite systems, the normal equations are often poorly conditioned and preconditioning can be difficult. Another related approach

* Received November 20, 2014 / Accepted April 13, 2016 /
Published online January 18, 2017 /

is First Order System Least Squares (FOSLS) [2–4], which converts the second order equation into an equivalent system of first order equations and then solves a least squares problem for this system. The method presented here also produces positive definite systems of equations, but it does so without reformulation as a least squares problem.

When iterative methods are employed to solve a system of linear equations, it is usually necessary to precondition the original system in order to speed up convergence. A great deal of work has been dedicated to formulating effective preconditioning strategies for the linear systems resulting from discretizations of the Helmholtz equation [5]. One approach that has seen much success is the Shifted Laplacian preconditioner [6–8]. In this approach, the preconditioner for the system of equations corresponding to $\Delta u + k^2 u = 0$ is the matrix corresponding to the “shifted” equation $\Delta u + (\alpha + i\beta)u = 0$. If the imaginary shift β is large enough, multigrid methods are expected to be successful in solving the shifted problem, and if $\alpha \approx 1$, the shifted operator should be a good preconditioner for the original problem. While this approach is often effective, in [9] the authors point out the advantages in using a preconditioner that is symmetric positive definite. When the preconditioning matrix is not positive definite, the coefficient matrix of the preconditioned system is not symmetric with respect to any inner product, which limits the methods available for solving the resulting system. The solution suggested in [9] is to use an approximation of the absolute value of the original coefficient matrix as preconditioner. In the method proposed here, both the matrices and the suggested preconditioners are symmetric positive definite, and therefore a wide range of Krylov subspace methods is available. In particular, we shall demonstrate the results obtained with Conjugate Gradient, which has a short recurrence and is very simple to implement and parallelize.

As a background to this approach, we follow [10], where Milton, Seppecher, and Bouchitté developed variational principles that apply to the Helmholtz equation above, as well as the time-harmonic Maxwell equations and the equations of linear elasticity in lossy materials. To derive these variational principles, we first define the dual variable

$$v = iL\nabla u.$$

Then

$$\begin{pmatrix} L & 0 \\ 0 & M \end{pmatrix} \begin{pmatrix} \nabla u \\ u \end{pmatrix} = \begin{pmatrix} L\nabla u \\ Mu \end{pmatrix} = \begin{pmatrix} -iv \\ -i\nabla \cdot v \end{pmatrix},$$

or equivalently,

$$\mathcal{G} = Z\mathcal{F},$$

where

$$\mathcal{F} = \begin{pmatrix} \nabla u \\ u \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} -iv \\ -i\nabla \cdot v \end{pmatrix}, \quad Z = \begin{pmatrix} L & 0 \\ 0 & M \end{pmatrix}.$$

For a complex quantity z , we will write $z' = \text{Re}(z)$ and $z'' = \text{Im}(z)$. Taking real and imaginary parts, the constitutive relation becomes

$$\mathcal{G}' = Z'\mathcal{F}' - Z''\mathcal{F}'' \quad \text{and} \quad \mathcal{G}'' = Z'\mathcal{F}'' + Z''\mathcal{F}',$$

which can be written in matrix form as

$$\begin{pmatrix} \mathcal{G}'' \\ \mathcal{G}' \end{pmatrix} = \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix}. \quad (1.2)$$

Solving this relation for the imaginary parts of \mathcal{F} and \mathcal{G} , we find that

$$\begin{pmatrix} \mathcal{G}'' \\ \mathcal{F}'' \end{pmatrix} = \mathcal{L} \begin{pmatrix} \mathcal{F}' \\ -\mathcal{G}' \end{pmatrix}, \quad (1.3)$$

where

$$\mathcal{L} = \begin{pmatrix} Z'' + Z'(Z'')^{-1}Z' & Z'(Z'')^{-1} \\ (Z'')^{-1}Z' & (Z'')^{-1} \end{pmatrix}.$$

The matrix \mathcal{L} is positive definite as long as Z'' is positive definite (see [10]). In terms of L and M , this means that

$$L''(x) > 0 \quad \text{and} \quad M''(x) > 0 \quad \text{for all } x \in \Omega.$$

In practice it is only necessary that the values of L and M lie within any half plane of the complex plane. This half plane can then be rotated so that it becomes the upper half plane (see Section 2).

The approach in [11] was to use this constitutive relation and the corresponding energy functional

$$\int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ -\mathcal{G}' \end{pmatrix} \cdot \begin{pmatrix} Z'' + Z'(Z'')^{-1}Z' & Z'(Z'')^{-1} \\ (Z'')^{-1}Z' & (Z'')^{-1} \end{pmatrix} \begin{pmatrix} \mathcal{F}' \\ -\mathcal{G}' \end{pmatrix} dx$$

to formulate a numerical method. When this variational principle is discretized by the finite element method, the result is a system of equations whose coefficient matrix can be partitioned into $N \times N$ blocks, where N is the number of nodes in the computational grid. A similar system of equations must be solved to find approximations for u'' and v' . In all, to find u' and u'' , one must solve two positive definite systems of equations of size $3N \times 3N$.

Here we develop a new method based on the saddle point variational principles in [10] that does not require that v be solved for in order to find u , but is still based on solving positive definite systems of equations. First, in Section 2, we will analyze the saddle point variational principles from [10] upon which our method is based. In Section 3, we will discuss the details of handling Dirichlet, Neumann, and Robin boundary conditions with these variational principles. Section 4 contains the derivation of a standard bound on the error incurred when the Helmholtz equation is solved using a finite element method that discretizes the saddle point variational principle. Section 5 outlines the numerical method and discusses the conditioning of the system. In Section 6, we provide several straightforward numerical explorations of the performance of the algorithm, as well as numerical verification of the error bound from Section 4.

2. The Saddle Point Variational Principle

The derivation of the saddle point variational principle from [10] follows the same steps presented in the introduction for the minimization principle, the difference being that instead of continuing to the constitutive relation (1.3), we stop at equation (1.2). Assuming that Z'' is positive definite, from (1.2) we define the functional

$$Y(u', u'') = \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} dx. \quad (2.1)$$

Let $u', u'' \in H^1(\Omega)$ be the real and imaginary parts of a solution to the Helmholtz equation. Let $s \in H_0^1(\Omega)$ and define

$$\mathcal{S} = \begin{pmatrix} \nabla s \\ s \end{pmatrix}.$$

Then we have

$$\begin{aligned} Y(u' + s, u'') &= \int_{\Omega} \begin{pmatrix} \mathcal{F}' + \mathcal{S} \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{F}' + \mathcal{S} \\ \mathcal{F}'' \end{pmatrix} dx \\ &= Y(u', u'') + 2 \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S} \\ 0 \end{pmatrix} dx + Y(s, 0). \end{aligned}$$

The integral in the line above can be rewritten as

$$\begin{aligned} \int_{\Omega} \begin{pmatrix} \mathcal{G}'' \\ \mathcal{G}' \end{pmatrix} \cdot \begin{pmatrix} \mathcal{S} \\ 0 \end{pmatrix} dx &= \int_{\Omega} \begin{pmatrix} -v' \\ -\nabla \cdot v' \end{pmatrix} \cdot \begin{pmatrix} \nabla s \\ s \end{pmatrix} dx \\ &= \int_{\Omega} [-v' \cdot \nabla s - \nabla \cdot v' s] dx = \int_{\Omega} -\nabla \cdot [v' s] dx = \int_{\partial\Omega} -v' \cdot ns dS = 0. \end{aligned} \quad (2.2)$$

Therefore,

$$Y(u' + s, u'') = Y(u', u'') + \int_{\Omega} \mathcal{S} \cdot Z'' \mathcal{S} dx,$$

and the last term must be nonnegative, since Z'' is assumed to be positive definite. A similar calculation yields

$$Y(u', u'' + s) = Y(u', u'') - \int_{\Omega} \mathcal{S} \cdot Z'' \mathcal{S} dx.$$

This shows that (u', u'') is at a saddle point of the functional Y .

Suppose that (u', u'') is a saddle point of the functional Y . Then the functional $Q(s', s'') = Y(u' + s', u'' + s'')$, defined for all $s', s'' \in H_0^1(\Omega)$, should have a saddle point at $s' = s'' = 0$. A necessary condition for this to happen is that the first variation of Q should vanish. If

$$\mathcal{S}' = \begin{pmatrix} \nabla s' \\ s' \end{pmatrix} \quad \text{and} \quad \mathcal{S}'' = \begin{pmatrix} \nabla s'' \\ s'' \end{pmatrix},$$

then we must have

$$0 = \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx.$$

After writing this equation out in terms of u , s , L , and M and integrating by parts, we find that the integrals

$$\int_{\Omega} \left(-\nabla \cdot L'' \nabla u' + M'' u' - \nabla \cdot L' \nabla u'' + M' u'' \right) s' dx,$$

and

$$\int_{\Omega} \left(-\nabla \cdot L' \nabla u' + M' u' + \nabla \cdot L'' \nabla u'' - M'' u'' \right) s'' dx$$

must add to zero for any choices of s' and s'' in $H_0^1(\Omega)$. The real and imaginary parts of the equation $\nabla \cdot L \nabla u = M u$ can be written as

$$\nabla \cdot L' \nabla u' - \nabla \cdot L'' \nabla u'' - M' u' + M'' u'' = 0,$$

and

$$\nabla \cdot L' \nabla u'' + \nabla \cdot L'' \nabla u' - M' u'' - M'' u' = 0.$$

Notice that the left-hand sides of these equations are just the opposites of the expressions multiplying s' and s'' in the integrals above. Since the result of the integral must be zero regardless of the choice of s' and s'' , the saddle point of Y must be a solution to the Helmholtz equation.

So far, we have assumed that Z'' is positive definite, but it is often possible to use this method even when L and M do not have positive imaginary parts. A solution of the equation

$$\nabla \cdot L \nabla u = Mu$$

is also a solution to

$$\nabla \cdot e^{i\theta} L \nabla u = e^{i\theta} Mu, \quad (2.3)$$

where θ is a constant. Therefore, to ensure that the imaginary part of Z'' is positive definite, we can apply a rotation so that the new coefficients $e^{i\theta}L$ and $e^{i\theta}M$ have positive imaginary parts. The necessary conditions on L and M for the method to apply are that their values lie within one open half-plane. That half-plane may then be rotated so that it becomes the upper half-plane.

3. Boundary Conditions

The calculations done above show that a saddle point of Y satisfying $u' = f'$ and $u'' = f''$ on $\partial\Omega$ is a solution of

$$\begin{cases} \nabla \cdot L \nabla u = Mu, & \text{in } \Omega, \\ u = f, & \text{on } \partial\Omega. \end{cases}$$

We can also solve the Neumann problem

$$\begin{cases} \nabla \cdot L \nabla u = Mu, & \text{in } \Omega, \\ v \cdot n = g, & \text{on } \partial\Omega. \end{cases} \quad (3.1)$$

Let $s', s'' \in H^1(\Omega)$ be arbitrary test functions. Then we have

$$\begin{aligned} 0 &= \int_{\Omega} \left((\nabla \cdot v' - \nabla \cdot v')s' + (-\nabla \cdot v'' + \nabla \cdot v'')s'' \right) dx \\ &= \int_{\Omega} \left(-v' \cdot \nabla s' - \nabla \cdot v' s' + v'' \cdot \nabla s'' + \nabla \cdot v'' s'' \right) dx + \int_{\partial\Omega} \left(s' v' \cdot n - s'' v'' \cdot n \right) dx \\ &= \int_{\Omega} \begin{pmatrix} \mathcal{G}'' \\ \mathcal{G}' \end{pmatrix} \cdot \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx + \int_{\partial\Omega} [s' v' \cdot n - s'' v'' \cdot n] dS \\ &= \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx + \int_{\partial\Omega} [s' v' \cdot n - s'' v'' \cdot n] dS. \end{aligned}$$

Therefore, in order to solve the Neumann problem, we solve the weak equation

$$\begin{aligned} &\int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx \\ &= \int_{\partial\Omega} [-s' g' + s'' g''] dS \quad \text{for all } s', s'' \in H^1(\Omega). \end{aligned}$$

To solve the Robin problem

$$\begin{cases} \nabla \cdot L \nabla u = Mu, & \text{in } \Omega, \\ u + av \cdot n = g, & \text{on } \partial\Omega, \end{cases} \quad (3.2)$$

we begin with the weak form of the Neumann problem, which we will write as

$$0 = \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx + \int_{\partial\Omega} \begin{pmatrix} s' \\ s'' \end{pmatrix} \cdot \begin{pmatrix} v' \cdot n \\ -v'' \cdot n \end{pmatrix} dS.$$

We split the boundary condition into its real and imaginary parts as

$$\begin{aligned} u' + a'v' \cdot n - a''v'' \cdot n &= g', \\ u'' + a'v'' \cdot n + a''v' \cdot n &= g'', \end{aligned}$$

which we can write as

$$\begin{pmatrix} u' \\ u'' \end{pmatrix} + \begin{pmatrix} a' & a'' \\ a'' & -a' \end{pmatrix} \begin{pmatrix} v' \cdot n \\ -v'' \cdot n \end{pmatrix} = \begin{pmatrix} g' \\ g'' \end{pmatrix}.$$

If the matrix in the equation above is called W , then

$$\begin{pmatrix} v' \cdot n \\ -v'' \cdot n \end{pmatrix} = -W^{-1} \begin{pmatrix} u' \\ u'' \end{pmatrix} + W^{-1} \begin{pmatrix} g' \\ g'' \end{pmatrix},$$

so the weak form of the equation with Robin boundary conditions is

$$\begin{aligned} & \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx - \int_{\partial\Omega} \begin{pmatrix} u' \\ u'' \end{pmatrix} \cdot W^{-1} \begin{pmatrix} s' \\ s'' \end{pmatrix} dS \\ &= - \int_{\partial\Omega} \begin{pmatrix} g' \\ g'' \end{pmatrix} \cdot W^{-1} \begin{pmatrix} s' \\ s'' \end{pmatrix} dS. \end{aligned} \quad (3.3)$$

The inverse of W is

$$W^{-1} = \frac{1}{-(a')^2 - (a'')^2} \begin{pmatrix} -a' & -a'' \\ -a'' & a' \end{pmatrix} = \frac{1}{|a|^2} \begin{pmatrix} a' & a'' \\ a'' & -a' \end{pmatrix} = \frac{1}{|a|^2} W,$$

so if we require that a' be negative, the matrix that results from discretizing the left-hand side of (3.3) will have the same block form as those that result from the other boundary conditions. If $a' > 0$, we can instead rotate so that L'' and M'' are negative so that the necessary block structure of the matrices is preserved.

Care must be taken with solving the Neumann and Robin problems when rotation is used, to ensure that the correct boundary conditions are enforced. For example, if one desires to solve the Neumann problem

$$\begin{cases} \nabla \cdot L \nabla u = Mu, & \text{in } \Omega, \\ v \cdot n = g, & \text{on } \partial\Omega. \end{cases}$$

the rotated version of the problem is

$$\begin{cases} \nabla \cdot e^{i\theta} L \nabla u = e^{i\theta} Mu, & \text{in } \Omega, \\ \tilde{v} \cdot n = e^{i\theta} g, & \text{on } \partial\Omega, \end{cases}$$

where $\tilde{v} = ie^{i\theta} L \nabla u$.

4. Error Bound

We will make the following assumptions on Z'' :

- a. there is a constant γ_1 such that $|[Z'']_{ij}(x)| < \gamma_1$ for all i, j , and $x \in \Omega$
 - b. there is γ_2 such that $Z''(x) > \gamma_2 I$ for all $x \in \Omega$.
- (4.1)

The requirements on Z'' are equivalent to requiring similar bounds on L'' and M'' . Define the space $V = [H^1(\Omega)]^2$, endowed with the norm

$$\|(u', u'')\|_V = (\|u'\|_{H^1(\Omega)}^2 + \|u''\|_{H^1(\Omega)}^2)^{\frac{1}{2}}. \quad (4.2)$$

Also, we will assume that V_{N1} and V_{N2} are finite dimensional subspaces of $H^1(\Omega)$, and that $V_N = V_{N1} \times V_{N2}$ is the space in which we seek our numerical solution.

Define a functional $f(s')$ for $s' \in H^1(\Omega)$ as

$$\begin{aligned} f(s') &= \frac{1}{2} Y(s', u'') + Q(s', u'') \\ &= \frac{1}{2} \int_{\Omega} \begin{pmatrix} S' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} S' \\ \mathcal{F}'' \end{pmatrix} dx + Q(s', u''), \end{aligned}$$

where, in practice, $Q(s', u'')$ would contain terms that arise from the enforcement of boundary conditions and any inhomogeneous terms. We will further split the terms as

$$f(s') = \frac{1}{2} B(s', s') - F(s', u''),$$

where

$$B(s', s'') = \int_{\Omega} S' \cdot Z'' S'' dx$$

and $F(s', u'')$ contains the rest of the terms in $f(s')$. If u' is a minimizer of $f(s')$, then u' must satisfy the Euler-Lagrange equation

$$B(u', s') = F(s', u'') \quad \text{for all } s' \in H^1(\Omega). \quad (4.3)$$

Therefore, we can write

$$\begin{aligned} f(s') &= \frac{1}{2} B(s', s') - F(s', u'') = B(u', u') - F(u', u'') + \frac{1}{2} B(s', s') - F(s', u'') \\ &= B(u', u') - F(u', u'') + \frac{1}{2} B(s', s') - B(u', s') \\ &= \frac{1}{2} B(u', u') - F(u', u'') + \frac{1}{2} B(u', u') - B(u', s') + \frac{1}{2} B(s', s') \\ &= \frac{1}{2} B(u', u') - F(u', u'') + \frac{1}{2} B(u' - s', u' - s'). \end{aligned}$$

Suppose that $u'_N \in V_{N1}$ is such that

$$f(u'_N) = \min_{s \in V_{N1}} f(s').$$

Then

$$B(u' - u'_N, u' - u'_N)^{\frac{1}{2}} = \min_{s' \in V_{N1}} B(u' - s', u' - s')^{\frac{1}{2}}, \quad (4.4)$$

and the inequalities (4.1) imply that

$$\sqrt{\gamma_2} \|s'\|_{H^1(\Omega)} \leq \sqrt{B(s', s')} \leq C\sqrt{\gamma_1} \|s'\|_{H^1(\Omega)} \quad \text{for all } s' \in H^1(\Omega).$$

Applying these inequalities to both sides of (4.4) yields

$$\sqrt{\gamma_2} \|u' - u'_N\|_{H^1(\Omega)} \leq \min_{s' \in V_{N1}} C\sqrt{\gamma_1} \|u' - s'\|_{H^1(\Omega)}.$$

Here and in what follows, C will represent a constant that does not depend on u' , u'' , or the grid spacing h .

In order to get the necessary bound, we must choose s' properly. Let F_1 be the orthogonal projection from $H^1(\Omega)$ onto V_{N1} . Then $\|F_1\|_{B(H^1(\Omega), H^1(\Omega))} = 1$, where $B(H^1(\Omega), H^1(\Omega))$ is the set of all bounded linear functions from $H^1(\Omega)$ to itself. We then take $s' = F_1 u'$ to obtain the inequality

$$\|u' - u'_N\|_{H^1(\Omega)} \leq C \|u' - F_1 u'\|_{H^1(\Omega)}. \quad (4.5)$$

If instead we use

$$\begin{aligned} \tilde{f}(s'') &= \frac{1}{2} Y(u', s'') + Q(u', s'') \\ &= \frac{1}{2} \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{S}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{F}' \\ \mathcal{S}'' \end{pmatrix} dx + Q(u', s''), \end{aligned}$$

and perform calculations similar to those above, we obtain the bound

$$\|u'' - u''_N\|_{H^1(\Omega)} \leq C \|u'' - F_2 u''\|_{H^1(\Omega)}, \quad (4.6)$$

where F_2 is the orthogonal projection from $H^1(\Omega)$ onto V_{N2} .

Combining inequalities (4.5) and (4.6), we find that

$$\begin{aligned} \|(u' - u'_N, u'' - u''_N)\|_V^2 &= \|u' - u'_N\|_{H^1(\Omega)}^2 + \|u'' - u''_N\|_{H^1(\Omega)}^2 \\ &\leq C \left(\|u' - F_1 u'\|_{H^1(\Omega)}^2 + \|u'' - F_2 u''\|_{H^1(\Omega)}^2 \right) \\ &= C \|(u' - F_1 u', u'' - F_2 u'')\|_V^2 \end{aligned}$$

and consequently,

$$\|(u', u'') - (u'_N, u''_N)\|_V \leq C \|(u' - F_1 u', u'' - F_2 u'')\|_V.$$

We partition Ω into subregions e_l , each of which can be viewed as a suitably shifted and rotated version of a reference element \hat{e} , so that there exist affine changes of variables $F_l(x) = B_l x + x_l$ such that $F_l(\hat{e}) = e_l$. In what follows, a hat over a function will denote the corresponding function defined over the reference element \hat{e} obtained by a change of variables.

We define the seminorm $|\cdot|_s$ by

$$|u|_s^2 = [u, u]_s,$$

where

$$[u, w]_s = \sum_{|\alpha|=s} \int_{\hat{e}} D^\alpha u \cdot D^\alpha w \, dx \quad (4.7)$$

and α is a multi-index. From [12] we get the inequality

$$c^{-1} h^{s-\frac{d}{2}} |w|_{s, e_l} \leq |\hat{w}|_s \leq c h^{s-\frac{d}{2}} |w|_{s, e_l}, \quad (4.8)$$

where c is a constant, $w = \hat{w} \circ F_l^{-1}$, and the subscript e_l denotes (4.7) with e_l in place of \hat{e} .

We now recall the following lemma from [13]:

Lemma 4.1 (Bramble-Hilbert Lemma) *For some region $\Omega \subset \mathbb{R}^d$ and some integer $k \geq -1$, let there be given a bounded linear functional*

$$f : H^{k+1}(\Omega) \rightarrow \mathbb{R},$$

satisfying $|f(u)| \leq \delta \|u\|_{H^{k+1}(\Omega)}$ for all $u \in H^{k+1}(\Omega)$ for some δ independent of u . Suppose that $f(u) = 0$ for all $u \in P_k(\bar{\Omega})$. Then there exists a constant C , dependent only on Ω such that

$$|f(u)| \leq C\delta |u|_{k+1}, \quad u \in H^{k+1}(\Omega).$$

Let $s \in \{0, 1\}$ and fix $w \in H^s(\hat{e})$. Define the functionals

$$L_1(\hat{u}) = [\hat{u} - F_1 \hat{u}, w]_s \quad \text{and} \quad L_2(\hat{u}) = [\hat{u} - F_2 \hat{u}, w]_s.$$

Since

$$\begin{aligned} |L_j(\hat{u})| &\leq |\hat{u} - F_j \hat{u}|_s |w|_s \leq \left(|\hat{u}|_s + |F_j \hat{u}|_s \right) |w|_s \\ &\leq \left(\|\hat{u}\|_{H^1(\hat{e})} + \|F_j \hat{u}\|_{H^1(\hat{e})} \right) |w|_s \\ &\leq 2 \|\hat{u}\|_{H^1(\hat{e})} |w|_s \leq 2 \|u\|_{H^{k+1}(\hat{e})} |w|_s, \end{aligned}$$

and $F_j u = u$ for polynomial functions u in V_{N_j} ($j = 1, 2$), we see that the Bramble-Hilbert Lemma applies, and there exist constants such that

$$|L_1(\hat{u}')| \leq C |w|_s |\hat{u}'|_{k+1} \quad \text{and} \quad |L_2(\hat{u}'')| \leq C |w|_s |\hat{u}''|_{k+1},$$

as long as k is small enough so that all polynomials of degree less than or equal to k are contained in the span of the basis functions representing \hat{u}' and \hat{u}'' . Taking $w = \hat{u}' - F_1 \hat{u}'$ in the first inequality and $w = \hat{u}'' - F_2 \hat{u}''$ in the second yields

$$|\hat{u}' - F_1 \hat{u}'|_s \leq C |\hat{u}'|_{k+1} \quad \text{and} \quad |\hat{u}'' - F_2 \hat{u}''|_s \leq C |\hat{u}''|_{k+1}. \quad (4.9)$$

Assuming that $h \leq 1$ and using inequality (4.8), we see that

$$\begin{aligned} |u' - F_1 u'|_{s, e_l} &\leq C h^{\frac{d}{2}-s} |\hat{u}' - F_1 \hat{u}'|_s \leq C h^{\frac{d}{2}-s} |\hat{u}'|_{k+1} \leq C h^{k-s+1} |u'|_{k+1, e_l}, \\ |u'' - F_2 u''|_{s, e_l} &\leq C h^{\frac{d}{2}-s} |\hat{u}'' - F_2 \hat{u}''|_s \leq C h^{\frac{d}{2}-s} |\hat{u}''|_{k+1} \leq C h^{k-s+1} |u''|_{k+1, e_l}. \end{aligned}$$

Consequently, the overall error satisfies

$$\begin{aligned}
& \|(u', u'') - (u'_N, u''_N)\|_V^2 \leq C \|(u' - F_1 u', u'' - F_2 u'')\|_V^2 \\
& \leq C \sum_l \left(|u' - F_1 u'|_{0,e_l}^2 + |u' - F_1 u'|_{1,e_l}^2 + |u'' - F_2 u''|_{0,e_l}^2 + |u'' - F_2 u''|_{1,e_l}^2 \right) \\
& \leq C \sum_l \left(h^{2k+2} |u'|_{k+1,e_l}^2 + h^{2k} |u'|_{k+1,e_l}^2 + h^{2k+2} |u''|_{k+1,e_l}^2 + h^{2k} |u''|_{k+1,e_l}^2 \right) \\
& \leq Ch^{2k} \left(|u'|_{k+1,\Omega}^2 + |u''|_{k+1,\Omega}^2 \right).
\end{aligned}$$

We have now proved:

Theorem 4.1. *Under the assumptions (4.1) on Z'' , if the solution $(u', u'') \in [H^{k+1}(\Omega)]^2$ and the finite element subspace used in the numerical method contains $[P_k(\bar{\Omega})]^2$, then there exists a constant C such that the error satisfies*

$$\|(u', u'') - (u'_N, u''_N)\|_V^2 \leq Ch^{2k} \left(|u'|_{k+1,\Omega}^2 + |u''|_{k+1,\Omega}^2 \right),$$

where $h \leq 1$ is the grid spacing.

5. The Numerical Method

To fix ideas, we will examine the numerical solution of the Dirichlet problem

$$\begin{cases} \nabla \cdot L \nabla u = Mu, & \text{in } \Omega, \\ u = f, & \text{on } \Omega. \end{cases} \quad (5.1)$$

The first step in solving the problem is to select a set of finite element basis functions. The numerical examples presented here will use a rectangular grid with bilinear basis functions.

Regardless of how the basis is chosen, we will assume that the basis functions are labeled as $\{\psi_k\}$ and we assume that the solution has the form

$$\begin{pmatrix} u' \\ u'' \end{pmatrix} = \begin{pmatrix} \psi'_0 + \sum \alpha'_k \psi_k \\ \psi''_0 + \sum \alpha''_k \psi_k \end{pmatrix},$$

where ψ'_0 and ψ''_0 are auxiliary functions satisfying the boundary conditions $\psi'_0 = f'$ and $\psi''_0 = f''$ on $\partial\Omega$. The weak form of the Euler-Lagrange equation for the saddle point variational principle is

$$0 = \int_{\Omega} \begin{pmatrix} \mathcal{F}' \\ \mathcal{F}'' \end{pmatrix} \cdot \begin{pmatrix} Z'' & Z' \\ Z' & -Z'' \end{pmatrix} \begin{pmatrix} \mathcal{S}' \\ \mathcal{S}'' \end{pmatrix} dx \quad \text{for all } s', s'' \in H_0^1(\Omega),$$

where, as usual,

$$\mathcal{S}' = \begin{pmatrix} \nabla s' \\ s' \end{pmatrix} \quad \text{and} \quad \mathcal{S}'' = \begin{pmatrix} \nabla s'' \\ s'' \end{pmatrix}.$$

We make the substitution above for u' and u'' and let s' and s'' be equal to each of the basis functions in turn. In doing so, we arrive at a system of equations which has the block form

$$\begin{pmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{pmatrix} \begin{pmatrix} \alpha' \\ \alpha'' \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad (5.2)$$

where A_1 is positive definite. The entries of the blocks in the coefficient matrix satisfy

$$[A_1]_{kj} = \int_{\Omega} \nabla \psi_k \cdot L'' \nabla \psi_j \, dx + \int_{\Omega} \psi_k M'' \psi_j \, dx,$$

$$[A_2]_{kj} = \int_{\Omega} \nabla \psi_k \cdot L' \nabla \psi_j \, dx + \int_{\Omega} \psi_k M' \psi_j \, dx.$$

The elements of the vector $b = (b_1, b_2)^T$ satisfy

$$[b_1]_j = - \int_{\Omega} \nabla \psi'_0 \cdot L'' \nabla \psi_j \, dx - \int_{\Omega} \psi'_0 M'' \psi_j \, dx - \int_{\Omega} \nabla \psi''_0 \cdot L' \nabla \psi_j \, dx - \int_{\Omega} \psi''_0 M' \psi_j \, dx,$$

$$[b_2]_j = - \int_{\Omega} \nabla \psi'_0 \cdot L' \nabla \psi_j \, dx - \int_{\Omega} \psi'_0 M' \psi_j \, dx + \int_{\Omega} \nabla \psi''_0 \cdot L'' \nabla \psi_j \, dx + \int_{\Omega} \psi''_0 M'' \psi_j \, dx.$$

This system of equations (5.2) is of saddle point type, and therefore there is a wide array of numerical methods that apply [14]. Among the simplest is the following, based on Schur complements. By using this approach, we reduce the problem from solving an indefinite $2N \times 2N$ system to solving two $N \times N$ positive definite systems. We solve the second equation in (5.2) for α'' and substitute into the first to obtain

$$A_1 \alpha'' = -b_2 + A_2 \alpha',$$

$$(A_1 + A_2 A_1^{-1} A_2) \alpha' = b_1 + A_2 A_1^{-1} b_2. \quad (5.3)$$

Because A_1 is positive definite and A_2 is symmetric, the coefficient matrices in both these systems of equations are positive definite. Equivalently, we can solve the second equation for α' and make the corresponding substitution into the first equation to obtain the system of equations

$$A_2 \alpha' = b_2 + A_1 \alpha'',$$

$$(A_2 + A_1 A_2^{-1} A_1) \alpha'' = b_1 - A_1 A_2^{-1} b_2. \quad (5.4)$$

The methods below can be adapted to this second system of equations under the assumption that A_2 is positive definite, which corresponds to L and M having positive real parts. If the real parts of L and M are both positive, the problem can be rotated so that the imaginary parts become positive, so we will focus primarily on equations (5.3).

While the matrix $A_1 + A_2 A_1^{-1} A_2$ is positive definite, it is also costly to store and to compute. For this reason, we use the preconditioned conjugate gradient (PCG) method to compute the solution to the system with this coefficient matrix, since this method only requires the ability to perform matrix-vector multiplication with the coefficient matrix. As a preconditioner for $A_1 + A_2 A_1^{-1} A_2$, we use the matrix A_1 . In this case, the preconditioned system has coefficient matrix

$$A_1^{-1} (A_1 + A_2 A_1^{-1} A_2) = I + (A_1^{-1} A_2)^2.$$

We can expect our system of equations to have the best conditioning when A_1^{-1} and A_2^{-1} are approximately the same, or alternatively when $\|A_1\|_2$ is much larger than $\|A_2\|_2$.

Systems with coefficient matrix A_1 appear explicitly in the algorithm, but must also be solved at each step when PCG is applied to the matrix $A_1 + A_2 A_1^{-1} A_2$, and there are many different ways in which this system can be solved. In the numerical examples that follow, all the systems of equations of the form $A_1 x = b$ are solved using PCG with an incomplete

Cholesky factorization of A_1 as a preconditioner. In essence, this introduces an inner and an outer PCG iteration in step 4 below. The following section illustrates how the total number of PCG iterations performed solving systems $A_1 x = b$ is related to the size of the computational grid and the coefficients in the Helmholtz equation.

The algorithm used here is as follows, though details such as the iterative solver or preconditioning method may be modified as desired:

1. Form the matrices A_1 and A_2 .
2. Compute the right-hand side vectors b_1 and b_2 .
3. Compute $w_1 = b_1 + A_2 A_1^{-1} b_2$.
4. Solve $(A_1 + A_2 A_1^{-1} A_2) \alpha' = w_1$ using PCG with the preconditioner A_1 .
5. Compute $w_2 = -b_2 + A_2 \alpha'$.
6. Solve $A_1 \alpha'' = w_2$ by PCG with an incomplete Cholesky factorization of A_1 as preconditioner.

This algorithm is completely implicit, and therefore is well suited for large-scale problems. Because all that is required are sparse matrix-vector multiplications, parallel implementations of this algorithm can produce a significant speedup. In particular, this algorithm could be implemented on a GPU cluster, where many graphics processing units (each of which contains many processing cores) are used in parallel to perform very fast computations.

In some situations, particularly those involving high frequency, $\|A_2\|_2$ is much larger than $\|A_1\|_2$, suggesting that we use formulation (5.4). However, A_2 is not positive definite and therefore neither is $A_2 + A_1 A_2^{-1} A_1$. The basic algorithm outlined above can still be used in this case, provided that PCG is replaced by an iterative method that does not require positivity, such as GMRES.

As pointed out in [15], many solver packages are focused mainly on solving systems of equations with real matrices. This method can be considered an equivalent real formulation of the usual complex system of equations.

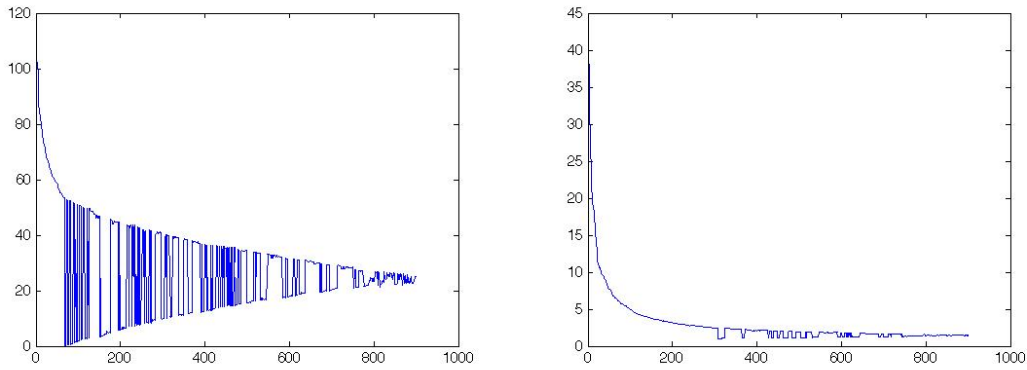


Fig. 5.1. The distribution of the eigenvalues of $A_1 + A_2 A_1^{-1} A_2$ (left) and the eigenvalues of $A_1^{-1}(A_1 + A_2 A_1^{-1} A_2)$ (right) for an example with 30×30 computational grid.

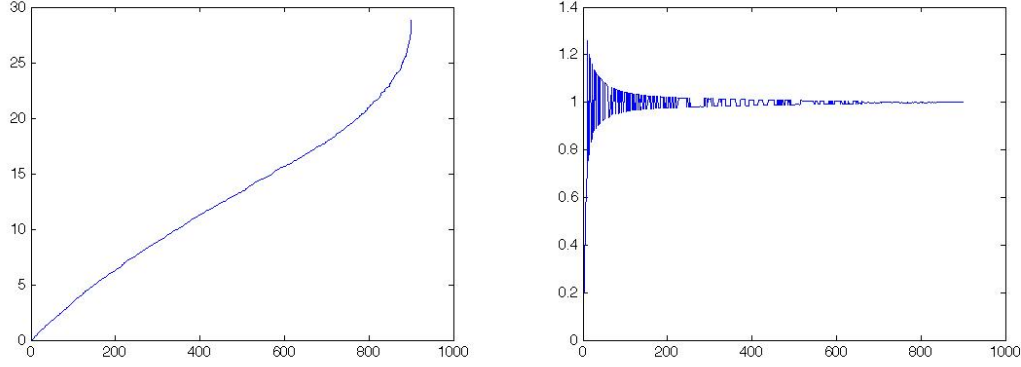


Fig. 5.2. The distribution of the eigenvalues of A_1 (left) and the eigenvalues of $(P^T P)^{-1} A_1$ (right), where P is an incomplete Cholesky factorization of A_1 for an example with a 30×30 computational grid.

5.1. Conditioning

In the numerical algorithm outlined above, we suggest that A_1 be used as a preconditioner for the system with matrix $A_1 + A_2 A_1^{-1} A_2$. In Fig. 5.1, we see the distribution of the eigenvalues of $A_1 + A_2 A_1^{-1} A_2$ and $A_1^{-1}(A_1 + A_2 A_1^{-1} A_2)$ for an example where the real and imaginary parts of L and M take on random values in the range $(0, 10)$.

Because the bulk of the work in this method comes from solving systems with matrix A_1 , it is important that such systems can be effectively preconditioned. A simple and effective choice is to use an incomplete Cholesky factorization of A_1 as the preconditioner, but there are many other preconditioning strategies that might be used. If the algorithm is being implemented in parallel, a particularly useful strategy would be to use a sparse approximate inverse [16], which avoids the “serial bottleneck” caused by having to perform back substitutions at each step in the PCG algorithm.

Fig. 5.2 shows the distribution of eigenvalues of A_1 before and after preconditioning. The preconditioner used here is an incomplete Cholesky factorization of A_1 with drop tolerance 0.01 and the real and imaginary parts of L and M take on random values in the range $(0, 10)$.

6. Numerical Results

In this section, we provide some demonstrations of the numerical solution of the Helmholtz problem

$$\nabla \cdot L \nabla u = M u.$$

Fig. 6.1 demonstrates the application of this method to a non-homogeneous problem with variable coefficients. The coefficients are chosen so that the frequency of the solution is higher in the lower right half of the unit square than in the upper left half, and the non-homogeneous term is a point source. In the following numerical examples we will solve homogeneous problems with constant coefficients and attempt to quantify the convergence of the algorithm for different values of the coefficients.

Fig. 6.2 shows how the fill-in depends on the drop tolerance in the incomplete Cholesky factorization of A_1 (which is used in this section as preconditioner for systems with coefficient

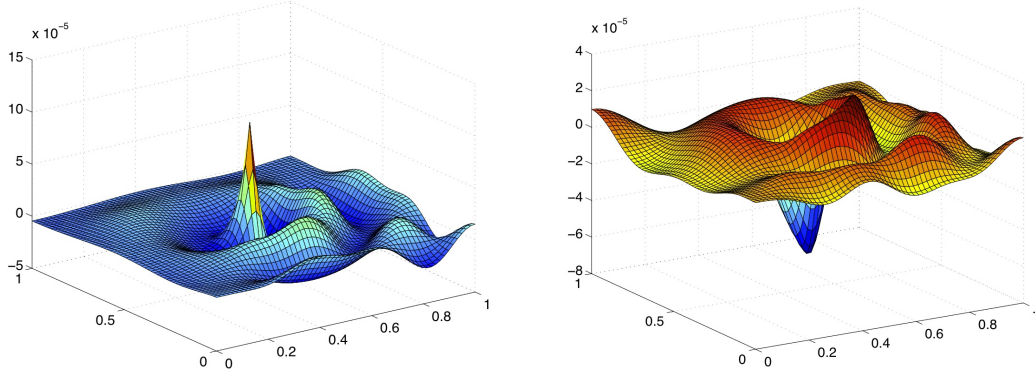


Fig. 6.1. The real (left) and imaginary (right) parts of a solution to the Helmholtz equation with a point source in the center of the domain and homogeneous Neumann boundary conditions. The two phase material is chosen so that the frequency is higher below the line $y = x$ than above it.

matrix A_1) for the problem with $L = 1$ and $M = 30 - 90i$ (before rotation) and the number of PCG iterations necessary to solve the same problem with a tolerance of 1×10^{-4} on the relative residual.

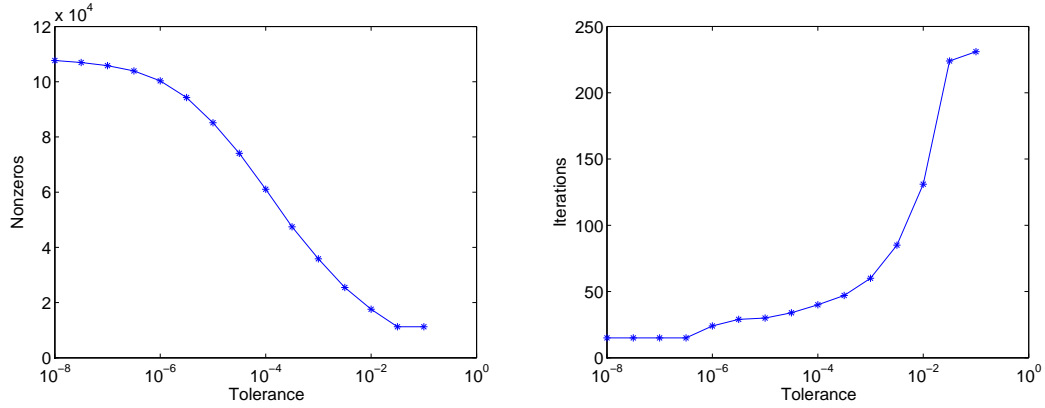


Fig. 6.2. Number of iterations and fill-in as a function of the drop tolerance used in the incomplete Cholesky factorization of A_1 .

In Fig. 6.3, we see the total number of PCG iterations necessary to solve all the systems with coefficient matrix A_1 for several different values of the coefficients in the problem as the size of the computational grid increases. The tolerances for the PCG algorithm is 1×10^{-6} and the drop tolerance for the incomplete Cholesky factorization of A_1 is 1×10^{-4} . It should be noted that in this case the growth in iterations happens entirely within the inner PCG iterations. The number of outer iterations required was either 2 or 3 in every instance.

Fig. 6.4 shows how the number of iterations is related to the frequency for problems where M is in the left-hand side of the upper half plane. In this situation, we cannot take advantage of the fact that $\|A_2\|_2 > \|A_1\|_2$ if we wish to solve only positive definite systems because A_2 is not positive definite. The number of grid points per wavelength is held approximately constant at 10 points per wavelength as ω grows.

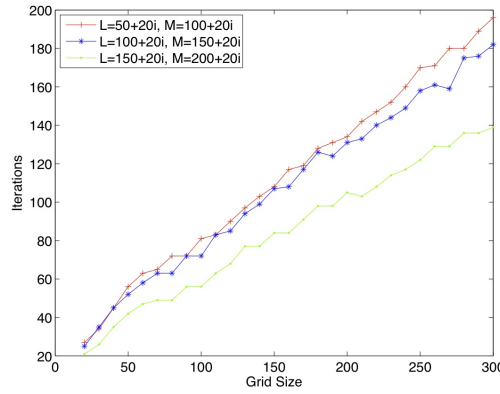


Fig. 6.3. The total number of inner PCG iterations required to solve the Helmholtz equation for several values of L and M as the size of the computational domain increases.

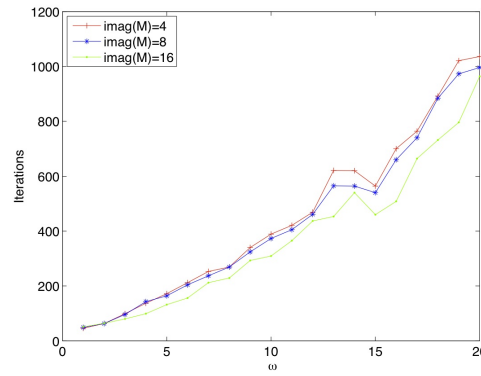


Fig. 6.4. The total number of inner PCG iterations required to solve the Helmholtz equation with $L = 1$ and $M = -\omega^2 + M''i$ for several values of M'' (after rotation). The number of grid points per wavelength is held approximately constant at 10 as ω increases.

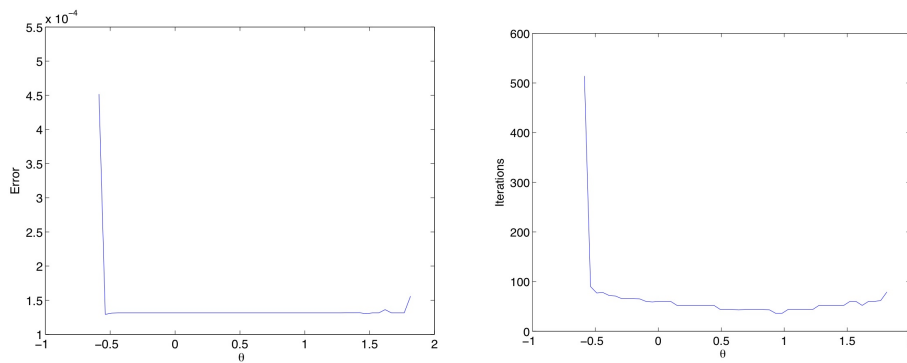
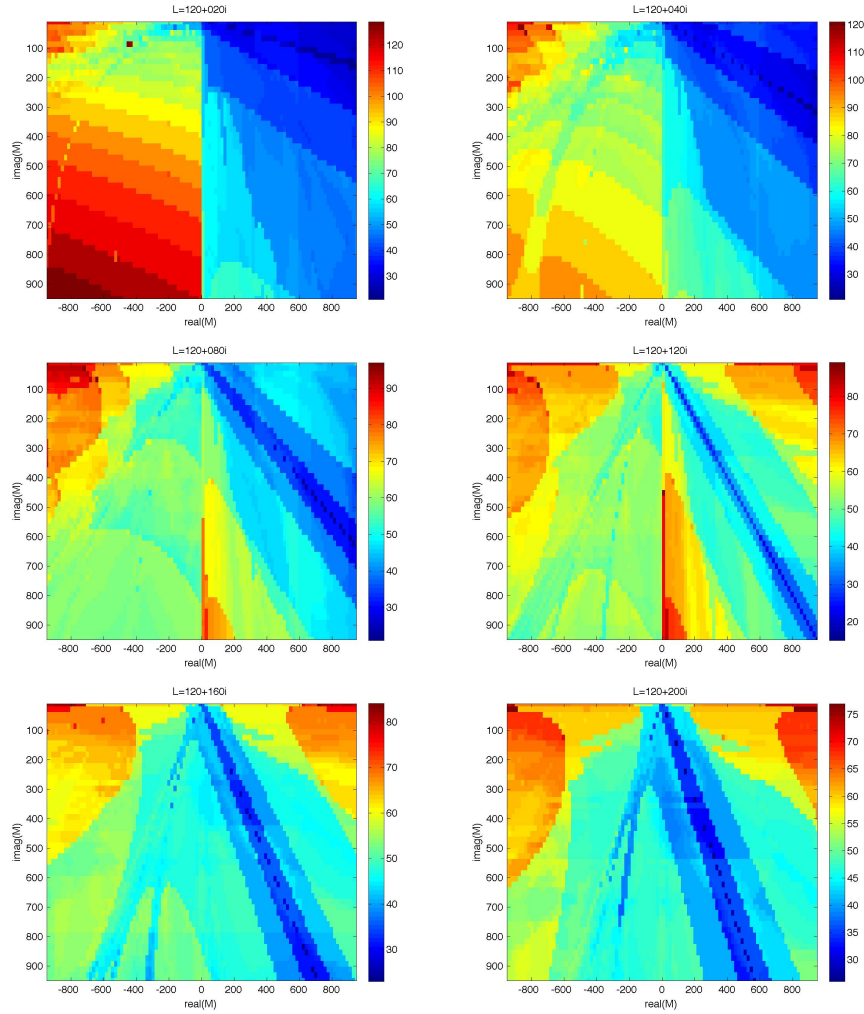


Fig. 6.5. The error and number of inner iterations in an example problem versus the value of θ used to rotate the problem.

Table 6.1: The error in the finite element solution for various grid sizes.

Grid	h	$\ (u' - u'_N, u'' - u''_N)\ _V^2$
32×32	0.032258	1.4013×10^{-3}
40×40	0.025641	8.8176×10^{-4}
50×50	0.020408	5.5822×10^{-4}
64×64	0.015873	3.3785×10^{-4}
70×70	0.014493	2.8107×10^{-4}
80×80	0.012658	2.1502×10^{-4}
90×90	0.011236	1.6936×10^{-4}
100×100	0.010101	1.3663×10^{-4}
128×128	0.007874	8.3172×10^{-5}
256×256	0.003923	2.0663×10^{-5}
512×512	0.001957	5.1762×10^{-6}

Fig. 6.6. The number of inner iterations required to solve the Helmholtz equation as the real and imaginary parts of M vary for several values of L . The grid size is fixed at 30×30 .

The graphs in Fig. 6.6 show in more detail how the number of iterations required to solve the Helmholtz equation depends on the coefficients in the problem. In order to get maximum advantage from the preconditioning strategy outlined in Section 5.1, when both coefficients L and M are in the first quadrant we choose formulation (5.3) when $\|A_1\|_\infty > \|A_2\|_\infty$, and we choose formulation (5.4) when the reverse inequality holds. Unfortunately, we must use formulation (5.3) when L is in the first quadrant and M is in the second quadrant because in this situation only A_1 is positive definite.

Table 6.1 shows the relationship between the error and the grid spacing in a problem with $L = -0.25 + 0.25i$ and $M = 0.1 + 0.3i$ and Dirichlet boundary conditions. In Fig. 6.5, the result of rotation on an example with $L = 3 + 2i$ and $M = 1 + 4i$ is shown. The error and number of iterations remain nearly constant until θ is such that one of the imaginary parts of the rotated coefficients approaches zero.

7. Conclusion

By formulating a finite element method with the saddle-point variational principles of Milton, Seppecher, and Bouchitté, we are able to solve boundary value problems for the complex Helmholtz equation through solving symmetric positive definite systems of equations. The method is based on using elimination on the block structure of the finite element matrix to produce two smaller systems of equations, both of which have positive definite coefficient matrices. The systems can then be solved using purely iterative methods. This method applies to a large class of problems, especially in light of the ability to “rotate” the coefficients of a given problem to fit the assumptions of the algorithm.

It should be emphasized that the method developed here does not only apply to the Helmholtz equation. In [10], there are similar variational principles given for the time-harmonic Maxwell equations and the equations of linear elasticity in lossy materials. The ideas presented here can easily be adapted to these situations. Also, the original variational principles of this type, developed by Cherkaev and Gibianski in [17], can be used to apply this numerical method to the complex Poisson equation.

As with the previous minimization-based method, the variational principles upon which this method is based remain valid as long as L and M have positive imaginary part, but the conditioning of the system deteriorates and the error incurred increases as L and M come close to violating this condition.

There is still more study necessary to determine the conditions under which this approach is competitive with other methods already in use. Also, it is worthwhile to consider other boundary conditions in addition to the ones presented herein, such as a PML [18]. Also, the application of this method to problems with a non-local boundary condition, such as those considered in [19] may also be explored.

In Section 6, the preconditioning method used in the inner iterations was simply an incomplete Cholesky factorization. To the extent that the growth in iterations in the inner iterations can be controlled, this method will become more attractive. Future work in this direction will be to compare potential preconditioning methods and their performance in the overall algorithm, including multigrid, sparse approximate inverse [16], and sweeping preconditioners [20], and also to compare the amount of work required when this method is implemented in parallel to more standard methods of solving Helmholtz equations.

Acknowledgements. The author would like to thank David Dobson for his helpful comments

during the preparation of the manuscript and Gang Bao for emphasizing the need for the improvements contained in this paper.

References

- [1] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, 2003.
- [2] T. Manteuffel and S. McCormick, First-order system least squares for second-order partial differential equations: Part I, *SIAM J. Numer. Anal.*, **31** (1994), 1785–1799.
- [3] T. Manteuffel and S. McCormick, First-order system least squares for second-order partial differential equations: Part II, *SIAM J. Numer. Anal.*, **34** (1997), 425–454.
- [4] S. McCormick and J. Ruge, First-order system least squares for the Helmholtz equation, *SIAM J. Sci. Comput.*, **21** (2000), 1927–1949.
- [5] O. Ernst and M. Gander, Why it is difficult to solve Helmholtz problems with classical iterative methods, I.G.T.H.O. Lakkis and R. Scheichl, editors, *Numerical Analysis of Multiscale Problems*, Springer, 2012.
- [6] C. Vuik and C. Oosterlee, On a class of preconditioners for the Helmholtz equation, *Appl. Numer. Math.*, **50** (2005), 409–425.
- [7] C. Oosterlee and C. Vuik, Comparison of the multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation, *Appl. Numer. Math.*, **56** (2006), 648–666.
- [8] M. van Gijzen, Y.A. Erlangga and C. Vuik, Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian, *SIAM J. Sci. Comput.*, **29** (2007), 1942–1958.
- [9] E. Vecharynski and A. Knyazev, Absolute value preconditioning for symmetric indefinite linear systems, *SIAM J. Sci. Comput.*, **35** (2013), A696–A718.
- [10] G. Milton, P. Seppecher and G. Bouchitté, Minimization variational principles for acoustics, elastodynamics, and electromagnetism in lossy inhomogeneous bodies at fixed frequency, *Proc. R. Soc. A*, **465** (2009), 367–396.
- [11] R. Richins and D. Dobson, A numerical minimization scheme for the complex Helmholtz equation, *ESAIM, Math. Model. Numer. Anal.*, **46** (2012), 39–57.
- [12] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, NY, 1991.
- [13] O. Axelsson and V.A. Barker, *Finite element solution of boundary value problems, theory and computation*, SIAM, Philadelphia, PA, 2001.
- [14] G. Golub and J. Liesen, Numerical solution of saddle point problems, *Acta Numer.*, **14** (2005), 1–137.
- [15] D. Day and M. Heroux, Solving complex-valued linear systems via equivalent real formulations, *SIAM J. Sci. Comput.*, **23** (2001), 480–498.
- [16] C. Meyer and M. Tuma, A sparse approximate inverse preconditioner for the conjugate gradient method, *SIAM J. Sci. Comput.*, **17** (1996), 1135–1149.
- [17] A. Cherkaev and L. Gibiansky, Variational principles for complex conductivity, viscoelasticity, and similar problems in media with complex moduli, *J. Math. Phys.*, **35** (1994), 127–145.
- [18] M. Slavutin and E. Turkel, Analytical and numerical studies of a finite element PML for the Helmholtz equation, *J. Comp. Acoust.*, **8** (2000), 121–137.
- [19] G. Bao and W. Sun, A fast algorithm for the electromagnetic scattering from a large cavity, *SIAM J. Sci. Comput.*, **27** (2005), 553–574.
- [20] B. Engquist and L. Ying, Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation, *Commun. Pur. Appl. Math.*, **64** (2011), 697–735.