

DELFT UNIVERSITY OF TECHNOLOGY

REPORT 03-18

SHIFTED LAPLACE PRECONDITIONERS FOR THE HELMHOLTZ
EQUATIONS

C. VUIK, Y.A. ERLANGGA AND C.W. OOSTERLEE

ISSN 1389-6520

Reports of the Department of Applied Mathematical Analysis

Delft 2003

Copyright © 2003 by Department of Applied Mathematical Analysis, Delft, The Netherlands.

No part of the Journal may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission from Department of Applied Mathematical Analysis, Delft University of Technology, The Netherlands.

Shifted Laplace preconditioners for the Helmholtz equations

C. Vuik * Y.A. Erlangga C.W. Oosterlee

December 19, 2003

1 Introduction

In this paper, we present a numerical method to solve the time-harmonic wave equation in 2D heterogeneous media. The underlying equation governs wave propagations and scattering phenomena arising in acoustic and optical problems. In particular, we look for solutions of the Helmholtz equation discretized by a finite difference method. Since the number of gridpoints per wavelength should be sufficiently large to result in acceptable solutions, for very high wavenumbers the discrete problem becomes extremely large, prohibiting the use of direct methods. Krylov subspace methods are an interesting alternative. However, these methods are not competitive without a good preconditioner. In this paper, we consider a class of preconditioners to improve the convergence of the Krylov subspace methods.

Various authors contributed to the development of powerful preconditioners for Helmholtz problems. The work in [1] can be considered as the start for the class of preconditioners we are interested in. A generalization has been recently proposed in [7]. In [1, 7], the preconditioners are constructed based on the Laplace operator. In [7], this operator is perturbed by a real-valued linear term. This surprisingly straightforward idea leads to very satisfactory convergence. Furthermore, the preconditioning matrix allows the use of SSOR, ILU, or multigrid to approximate the inversion within an iteration.

In this paper (see also [3]), we will generalize the approach in [1, 7] and give theoretical and numerical evidence that introducing a *complex* perturbation to the Laplace operator can result in a better preconditioner than using a real-valued perturbation. This class of preconditioners is simple to construct and is easy to extend to inhomogeneous media. Other preconditioners are proposed in [18, 4, 6, 8, 11].

This paper is organized as follows. In Section 2 we describe the mathematical model and

*Delft University of Technology, Mekelweg 4, 2628 CD Delft, Netherlands *c.vuik@math.tudelft.nl*, *Y.A.Erlangga@EWI.TUdelft.nl*, *C.W.Oosterlee@EWI.TUdelft.nl*

the discretization used for solving wave propagation problems. Iterative methods used for solving the resulting linear system and the preconditioner will be discussed in Section 3. In Section 4, we present the some spectral properties of the Shifted Laplace preconditioners. Numerical results are then presented in Section 5.

2 Mathematical model

In this paper, wave propagations are modelled in a two dimensional medium with inhomogeneous properties in a unit (scaled) domain governed by the Helmholtz equation

$$-\Delta\phi - k^2(x, y)\phi = f, \quad \Omega = [0, 1]^2, \quad (1)$$

where $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplace operator, and $k(x, y) \in \mathbb{R}$ is the wavenumber, which depends on the spatial position in the domain. We consider an open problem, i.e., outgoing waves penetrate at least at one boundary without (spurious) reflections. To satisfy this condition, a radiation-type condition is imposed. In this paper, the first order Sommerfeld condition is chosen of the form

$$\frac{\partial\phi}{\partial n} - ik\phi = 0, \quad \text{on a part of } \Gamma = \partial\Omega, \quad (2)$$

with n an outward direction normal to the boundary. Eventhough (2) may not be sufficiently accurate for inclined outgoing waves, it is state-of-the-art in industrial codes, easy to implement, and requires only a few gridpoints. We anticipate for possible reflections by considering a sufficiently large domain enabling any wave reflections to be immediately damped out and therefore be localized in the neighborhood of the boundaries.

The equation is discretized using the second-order difference scheme, in x -direction:

$$\frac{\partial^2\phi}{\partial x^2} = \frac{1}{h^2} (\phi_{i-1} - 2\phi_i + \phi_{i+1}) + \mathcal{O}(h^2), \quad (3)$$

and similar in y -direction. The first order derivative in (2) is discretized with the first order scheme

$$\frac{\partial\phi}{\partial n} = \frac{1}{\Delta n} (\phi_{i+1} - \phi_i). \quad (4)$$

Substituting (3) and (4) into (1) and (2), one obtains a linear system

$$Ap = b, \quad A \in \mathbb{C}^{N \times N}, \quad (5)$$

where A is a large, sparse symmetric matrix and N is the number of gridpoints. Matrix A is complex-valued and indefinite for large values of k .

3 Preconditioned Krylov methods

3.1 Krylov subspace methods

For a large, sparse matrix, Krylov subspace methods are very popular. The methods are developed based on a construction of iterants in the subspace

$$\mathcal{K}^j(A, r_0) = \text{span}\{r_0, Ar_0, A^2r_0, \dots, A^{j-1}r_0\}, \quad (6)$$

where $\mathcal{K}^j(A, r_0)$ is the j -th Krylov subspace associated with A and r_0 .

The basic algorithm within this class is the Conjugate Gradient method (CG) which has the nice properties that it uses only three vectors in memory and minimizes the error in the A -norm. However, the algorithm can only be used if the matrix A is symmetric and positive definite. In cases where one of these two properties is violated, CG may break down. For indefinite linear systems, CG can be applied to the normal equations since the resulting linear system becomes (positive) definite. Upon application of CG to the normal equations, CGNR results. Using CGNR, the iterations are guaranteed to converge. The drawback is that the condition number of the normal equations equals the square of the condition number of A , slowing down the convergence drastically.

Some algorithms with short recurrences but without the minimizing property are constructed based on the bi-Lanczos algorithm. Within this class, BiCG exists and its modifications: CGS [14] and Bi-CGSTAB [15]. In many cases, Bi-CGSTAB exhibits a smooth convergence behavior and often converges faster than BiCG and CGS. Also within this class are QMR [5] and COCG [16].

MINRES [10] can also be used to solve indefinite symmetric linear systems, as well as its generalization to the nonsymmetric case, GMRES [13]. Both algorithms have the minimization property but GMRES uses long recurrences so the amount of storage increases as the iteration number increases. A way to remedy the storage problem in GMRES is by including a so-called inner iteration as in GMRESR [17] and FGMRES [12]. For a comparison of these methods applied to Helmholtz problems, we refer to [2].

3.2 Preconditioners

To improve the convergence of iterative methods, a preconditioner M should be incorporated. By left preconditioning one solves the linear system

$$M^{-1}Ap = M^{-1}b. \quad (7)$$

The best choice for M^{-1} is the inverse of A , which is impractical. If A is SPD, one can approximate A^{-1} by one iteration of SSOR or multigrid. However, most practical wave

problems result in an indefinite linear system, for which SSOR or multigrid are not guaranteed to converge.

In general, one can distinguish two approaches for constructing preconditioners: matrix-based and operator-based. Within the first class lie, e.g., incomplete LU (ILU) factorizations. Examples of operator-based preconditioners are: separation of variables [11] and analytic ILU (AILU) [6].

An ILU preconditioner can be constructed by performing Gauss elimination and dropping elements which are smaller than a specified value, giving $ILU(tol)$. Preconditioners from this class are not effective for general indefinite problems. Reference [6] shows some results in which ILU-type preconditioners are used to solve the Helmholtz equation using QMR. For high wavenumbers k , $ILU(0)$ converges slowly, while $ILU(tol)$ encounters storage problems and also slow convergence. For sufficiently high wavenumbers k , the cost to construct the $ILU(tol)$ factors may become very high. Recently some shifted ILU preconditioners have been investigated [8].

The separation of variables preconditioner [11] is constructed by approximating k by a sum of two terms, one depending on x , and the other depending on the remaining coordinates. For smooth models and low frequencies the convergence rate is satisfactory, but it deteriorates when the roughness of the model or the frequency increases.

Instead of constructing the ILU factors from A , the Helmholtz operator $\mathcal{L}_h = -\Delta - k^2$ can be used to set up *ILU-like* factors in so-called analytic ILU (AILU) [6]. Starting with the Fourier transform of the analytic operator in one direction, one constructs parabolic factors of the Helmholtz operator consisting of a first order derivative in one direction and a non-local operator. To remove the non-local operator, a localized approximation is proposed, involving optimization parameters. Finding a good approximation for inhomogeneous problems is the major difficulty in this type of preconditioner. This is because the method is sensitive with respect to small changes in these parameters. The optimization parameters depend on $k(x, y)$.

3.3 Shifted Laplace preconditioner

Another approach is found in *not* looking for an approximate inverse of the discrete indefinite operator A , but merely looking for a form of M , for which $M^{-1}A$ has satisfactory properties for Krylov subspace acceleration. A first effort to construct a preconditioner in such a way is in [1]. An easy-to-construct $M = -\Delta_h$ preconditioner is incorporated for CGNR. One SSOR iteration is used whenever operations involving M^{-1} are required.

Instead of the Laplace operator as the preconditioner, [7] investigates possible improvements if an extra term k^2 is added to the Laplace operator $-\Delta_h$. So, the Helmholtz

equation with reversed sign is proposed as the preconditioner M . This preconditioner is then used in CGNR. One multigrid iteration is employed whenever M^{-1} must be computed. Instead of the normal equations, our findings suggest that GMRES or Bi-CGSTAB can solve the preconditioned linear system efficiently in less arithmetic operations. We propose a generalization of this preconditioner. In this preconditioner a term $(\alpha + \beta i)k^2$ is added to the Laplace operator $-\Delta_h$, with $\alpha, \beta \in \mathbb{R}$ and $\alpha \geq 0$.

In the next section, we concentrate on this type of preconditioners.

4 Spectral properties of the preconditioned matrix

In this section we investigate the spectral properties of the preconditioned matrices for the class of shifted Laplace preconditioners. In Section 4.1 we compare various choices of the shift parameter, whereas the dependence on the gridsize is investigated in Section 4.2.

4.1 A comparison of the eigenvalues

We analyze the spectral properties of the discrete formulation of (1). Suppose that the Helmholtz equation is discretized, we arrive at the linear system $Ap = b$. Matrix A can be splitted into two parts: the Laplace component B and the additional diagonal term k^2I so that $A = B - k^2I$ and therefore

$$(B - k^2I) p = b. \quad (8)$$

In this analysis, we use only Dirichlet or Neumann conditions at the boundaries in order to keep the matrix A real-valued. We precondition (8) using $M = B + (\alpha + i\beta)k^2I$, constructed with the same boundary conditions as for A . This gives

$$(B + (\alpha + i\beta)k^2I)^{-1} (B - k^2I) p = (B + (\alpha + i\beta)k^2I)^{-1} b. \quad (9)$$

The generalized eigenvalue problem of (9) is accordingly

$$(B - k^2I) p_v = \lambda_v (B + (\alpha + i\beta)k^2I) p_v. \quad (10)$$

Both systems (9) and (10) are indefinite if k^2 is larger than the smallest eigenvalue of B . In such a case, the convergence is difficult to estimate. Therefore, the subsequent analysis will be based on the normal equations formulation of the preconditioned matrix system (as in [7]).

Denote the eigenvalues of B as $0 < \mu_1 \leq \mu_2 \leq \dots \leq \mu_N$ and $Q_{\alpha+\beta i} = (M^{-1}A)^* (M^{-1}A)$, where $M = B + (\alpha + \beta i)k^2I$. We find for the eigenvalues of the preconditioners the following expressions:

Bayliss and Turkel	$\lambda_j(Q_0) = \left(1 - \frac{k^2}{\mu_j}\right)^2,$
Laird	$\lambda_j(Q_1) = \left(1 - \frac{2k^2}{\mu_j + k^2}\right)^2,$
Complex	$\lambda_j(Q_i) = 1 - \frac{2\mu_j k^2}{\mu_j^2 + k^4}.$

We consider two cases: $k^2 < \mu_1$ and $\mu_1 \leq k^2 \leq \mu_N$.

Comparison for $k^2 < \mu_1$.

In this case matrix A is positive definite. After some analysis, the following inequalities are derived:

$$\lambda_{\min}(Q_0) > \lambda_{\min}(Q_1), \text{ and } \lambda_{\min}(Q_0) > \lambda_{\min}(Q_i).$$

Furthermore the following limits can be obtained:

$$\lim_{\mu_N \rightarrow \infty} \lambda_{\max}(Q_0) = \lim_{\mu_N \rightarrow \infty} \lambda_{\max}(Q_1) = \lim_{\mu_N \rightarrow \infty} \lambda_{\max}(Q_i) = 1.$$

The convergence of CGNR is well described by the condition number of Q : $\kappa(Q) = \frac{\lambda_{\max}}{\lambda_{\min}}$. Using the relations given above we conclude that the Bayliss and Turkel preconditioner converges faster than the other choices.

Comparison for $\mu_1 \leq k^2 \leq \mu_N$.

In this case matrix A is indefinite. We first consider the Bayliss and Turkel preconditioner. Using the expression for the eigenvalues it appears that

$$\lim_{k \rightarrow \infty} \lambda_{\max}(Q_0) = \lim_{k \rightarrow \infty} \left(1 - \frac{k^2}{\mu_1}\right)^2 = \infty.$$

Therefore $\lambda_{\max}(Q_0)$ can become very large, which makes this preconditioner less favorable in this situation. With respect to the other preconditioners we note that the largest eigenvalues remain bounded by 1. In order to compare these preconditioners we have to consider the smallest eigenvalue in more detail. Assuming that $\lambda_{\min} \approx 0$ implies that there is an m such that $\mu_m = k^2 + \epsilon$, and $|\epsilon| \ll k^2$. After substituting this relation into the expressions of the smallest eigenvalues and neglecting higher order terms one obtains:

$$\lambda_{\min}(Q_1) = \frac{\epsilon^2}{4k^4} < \frac{\epsilon^2}{2k^4} = \lambda_{\min}(Q_i).$$

This implies that $\kappa(Q_i) < \kappa(Q_1)$, so we expect that M_i is the best preconditioner.

4.2 Eigenvalue dependence on the grid size

From the previous section it appears that the difference in convergence behavior for the shifted Laplace preconditioners is mainly determined by the value of the smallest eigenvalue. In order to analyze the dependence of the smallest eigenvalue on the stepsize h we consider the following simple Helmholtz problem:

$$-\frac{d^2\phi}{dx^2} - k^2\phi = 0, \quad 0 < x < 1, \quad \phi(0) = 1 \text{ and } \phi(1) = 0. \tag{11}$$

The eigenvalues μ_j^c of problem (11) with $k = 0$ are well known: $\mu_j^c = (j\pi)^2$, with $j = 1, 2, \dots$. Using the standard central difference method for the Laplace operator, with $N + 1$ grid points and $h = \frac{1}{N}$, it appears that the eigenvalues of the matrix B are given by

$$\mu_j = \frac{4}{h^2} \left(\sin \frac{\pi h j}{2} \right)^2, \text{ with } j = 1, \dots, N.$$

If \hat{j} is such that $\frac{\pi h \hat{j}}{2} \ll 1$ it follows that $|\mu_j - \mu_j^c| = O(h^2)$ for $j \leq \hat{j}$. So, the smallest eigenvalues of the matrix B are good approximations of the eigenvalues of the continuous problem. Suppose that $k^2 \neq \mu_j^c$ for all j . Then we have that

$$\lim_{h \rightarrow 0} \min_j |\mu_j - k^2| = |\mu_m^c - k^2| \neq 0,$$

where $|\mu_m^c - k^2| = \min_j |\mu_j^c - k^2|$. Combining this limit with the analysis given in Section 4.1 shows that

$$\lim_{h \rightarrow 0} \lambda_{\min}(Q_1) = \frac{|\mu_m^c - k^2|^2}{4k^4} \text{ and } \lim_{h \rightarrow 0} \lambda_{\min}(Q_i) = \frac{|\mu_m^c - k^2|^2}{2k^4}.$$

Since the maximal eigenvalues of Q_1 and Q_i are bounded by 1, we conclude that the convergence behavior of both preconditioners is independent of h (see also [9]). Only initially there can be some dependence of the smallest eigenvalue on h .

In order to illustrate this we show the smallest eigenvalue of Q for $k = 6.5$ and $k = 8$ in Figure 1 and 2 respectively as a function of the number of grid points. Note that for $k = 6.5$ the smallest eigenvalue decreases somewhat, and for $k = 8$ the smallest eigenvalue increases.

If $k^2 = \mu_j^c$ we have a resonance. In such a case the solver should be adapted because

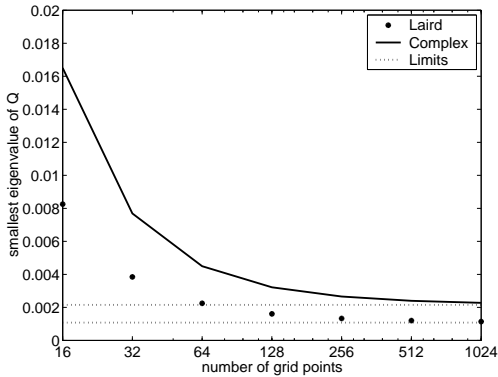


Figure 1: The smallest eigenvalue of Q for $k = 6.5$

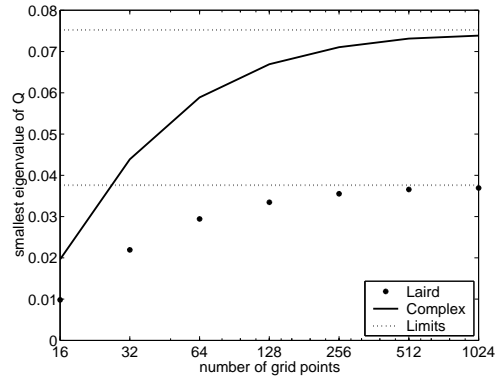


Figure 2: The smallest eigenvalue of Q for $k = 8$

the original matrix A is singular. The observations and analysis given above also holds for more general problems and discretizations.

Table 1: Number of GMRES iterations

k	M_0			M_1			M_i		
	h^{-1}	h^{-1}	h^{-1}	h^{-1}	h^{-1}	h^{-1}	h^{-1}	h^{-1}	h^{-1}
10	13	14	14	17	17	17	15	15	16
20	26	26	28	29	30	32	25	26	26
30	52	57	57	54	59	60	44	47	49
40	99	96	98	96	96	99	79	78	80

5 Numerical results

Closed-off problem

We consider a problem in a rectangular homogeneous medium governed by

$$\begin{aligned}
 -(\Delta + k^2)\phi &= -(k^2 - 5\pi^2)\sin(\pi x)\sin(2\pi y), & x = [0, 1], y = [0, 1], \\
 \phi &= 0, & \text{at the boundaries.}
 \end{aligned} \tag{12}$$

In Table 1, the numerical performance is shown for the preconditioners for different grid resolutions. If h is small enough (depending on k) the convergence does not depend on the grid size. For most cases, M_i outperforms the other preconditioners.

2-D open inhomogeneous problem

The second problem represents an open problem allowing waves to penetrate the boundaries. We consider an inhomogeneous medium: the wavenumber varies inside the domain

$$k = \begin{cases} k_{\text{ref}} & 0 \leq y \leq 1/3, \\ 1.5k_{\text{ref}} & 1/3 \leq y \leq 2/3, \\ 2.0k_{\text{ref}} & 2/3 \leq y \leq 1.0. \end{cases} \tag{13}$$

The number of gridpoints used is $5 \times k_{\text{ref}}$. Numerical results are presented in Table 2. In this harder problem, M_i again outperforms M_0 and M_1 indicated by the smaller number of iterations required to reach convergence. From Table 2, we also see that the preconditioned Bi-CGSTAB does not perform well for M_0 and M_1 , as already indicated in [7]. However, the convergence with M_i as the preconditioner is still satisfactory.

6 Conclusions

In this paper, a class of preconditioners based on the Shifted Laplace operator for the Helmholtz equation has been presented and analyzed. We find that the complex Shifted-

Table 2: Number of iterations for GMRES, CGNR, and Bi-CGSTAB

k_{ref}	GMRES			CGNR			Bi-CGSTAB		
	M_0	M_1	M_i	M_0	M_1	M_i	M_0	M_1	M_i
10	39	47	31	189	88	66	150	56	22
15	73	85	54	647	175	126	685	113	40
20	120	>150	82	>1000	268	194	>1000	177	60
30	>150	>150	141	>1000	502	361	>1000	344	105

Laplace operator leads to the most effective preconditioning matrix within this class of preconditioners.

References

- [1] A. Bayliss, C.I. Goldstein, and E. Turkel. An iterative method for Helmholtz equation. *J. Comput. Phys.*, 49:443–457, 1983.
- [2] M. Clemens and T. Weiland. Comparison of Krylov-type methods for complex symmetric linear systems applied to high-voltage problems. *IEEE Transactions on Magnetics*, 34:3335–3338, 1998.
- [3] Y.A. Erlangga, C. Vuik, and C.W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation. Report 03-01, Delft University of Technology, Department of Applied Mathematical Analysis, Delft, 2003. <http://ta.twi.tudelft.nl/nw/users/vuik/papers/Erl03VO.pdf>.
- [4] R.W. Freund. Preconditioning of symmetric, but highly indefinite linear systems. Num. Anal. Manus. 97-3-03, Bell Labs., Murray Hill, NJ, 1997.
- [5] R.W. Freund and N.M. Nachtigal. QMR: a Quasi-Minimal method for non-Hermitian linear systems. *Num. Math.*, 60:315–339, 1991.
- [6] M.J. Gander and F. Nataf. AILU for Helmholtz problems: a new preconditioner based on the analytic parabolic factorization. *J. Comp. Acoustics*, 9:1499–1509, 2001.
- [7] A.L. Laird. Preconditioned iterative solution of the 2d helmholtz equation. First year’s report, Oxford University, St. Hugh’s College, Oxford, 2001.
- [8] M.M.M. Made. Incomplete factorization-based preconditionings for solving the Helmholtz equation. *Int. J. Numer. Meth. Engng.*, 50:1077–1101, 2001.

- [9] T.A. Manteuffel and S.V. Parter. Preconditioning and boundary conditions. *SIAM J. Numer. Anal.*, 27:656–694, 1990.
- [10] C.C. Paige and M.A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.
- [11] R.E. Plessix and W.A. Mulder. Separation-of-variables as a preconditioner for an iterative Helmholtz solver. *Appl. Numer. Math.*, 44:385–400, 2003.
- [12] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM J. Sci. Stat. Comp.*, 14:461–469, 1993.
- [13] Y. Saad and M.H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856–869, 1986.
- [14] P. Sonneveld. CGS, a fast Lanczos-type solver for nonsymmetric linear systems. *SIAM J. Sci. Stat. Comp.*, 10:36–52, 1989.
- [15] H.A. van der Vorst. Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for solution of non-symmetric linear systems. *SIAM J. Sci. Stat. Comp.*, 13:631–644, 1992.
- [16] H.A. van der Vorst and J.B.M. Melissen. A Petrov-Galerkin type method for solving $Ax = b$, where A is symmetric complex. *IEEE Trans. on Magnetics*, 26:706–708, 1990.
- [17] H.A. van der Vorst and C. Vuik. GMRESR: a family of nested GMRES methods. *Num. Lin. Alg. Appl.*, 1:369–386, 1994.
- [18] H. Yserentant. Preconditioning indefinite discretization matrices. *Numer. Math.*, 54:719–734, 1988.