ESAIM: M2AN 46 (2012) 647–660
DOI: 10.1051/m2an/2011063

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

ITERATIVE SCHEMES FOR HIGH ORDER COMPACT DISCRETIZATIONS TO THE EXTERIOR HELMHOLTZ EQUATION *

Yogi Erlangga[1,2] and Eli Turkel[3]

Abstract. We consider high order finite difference approximations to the Helmholtz equation in an exterior domain. We include a simplified absorbing boundary condition to approximate the Sommerfeld radiation condition. This yields a large, but sparse, complex system, which is not self-adjoint and not positive definite. We discretize the equation with a compact fourth or sixth order accurate scheme. We solve this large system of linear equations with a Krylov subspace iterative method. Since the method converges slowly, a preconditioner is introduced, which is a Helmholtz equation but with a modified complex wavenumber. This is discretized by a second or fourth order compact scheme. The system is solved by BICGSTAB with multigrid used for the preconditioner. We study, both by Fourier analysis and computations this preconditioned system especially for the effects of high order discretizations.

Mathematics Subject Classification. 35J47, 65M06.

Received August 26, 2009. Revised January 3, 2011.
Published online January 11, 2012.

1. Introduction

Many physical phenomena such as acoustics, elasticity and electromagnetic waves are governed in the frequency domain by the Fourier transform of the wave equation, the Helmholtz equation, with appropriate boundary conditions. Examples include acoustic scattering about bodies and geophysical problems. In these problems, the boundaries of the computational domain are typically treated as absorbing ones to guarantee that the outgoing waves are not reflected back into the domain. The main challenge comes from the fact that the Helmholtz equation has to be solved for a wide range of frequencies, from low to high, especially if one is interested in the associated time-domain solution of the problem, obtained by an inverse Fourier-transform of the frequency-domain solutions.

In order to compute the numerical solution, a mesh needs to be defined, and an appropriate discretization is applied to the Helmholtz equation. It is well known [1,3] that the mesh length, h, required for a fixed accuracy increases faster than linear in the frequency. Babuška and coworkers named this the "pollution effect" [1]. In

Keywords and phrases. Helmholtz equation, high order compact schemes.

* *Dedicated to the memory of a dear friend, David Gottlieb.*

[1] Department of Earth and Ocean Sciences, The University of British Columbia, 2329 West Mall Vancouver, BC V6T 124, Canada.

[2] Currently at Mathematics Division, College of Sciences, Alfaisal University, P.O. Box 50927, Riyadh, 11533 Kingdom of Saudi Arabia.

[3] Department of Mathematics, Tel Aviv University, Tel Aviv, Israel. eliturkel@gmail.com

particular, the number of grid points N follows the relation

$$N = \frac{1}{h} \sim k^{\frac{p+1}{p}},$$

where p is the order of accuracy of the scheme. Thus, for a second order accurate scheme the grid grows like $k^{3/2}$, in order to maintain a constant accuracy. For a fourth order algorithm $N \approx k^{\frac{5}{4}}$, while for a sixth order accurate scheme $N \approx k^{\frac{7}{6}}$. Hence, for high wave numbers (frequencies) there is a need for very fine grids. This requirement is, however, less demanding if a higher order discretization is used. For finite differences, examples of high order accurate finite-difference discretizations on uniform meshes are introduced and analyzed by Harari and Turkel [12], and later by Singer and Turkel [13,14].

Given a discrete approximation to the Helmholtz equation, we need to solve the associated linear system. As the problem size increases, the efficiency, in terms of computational work and storage, of a direct method deteriorates rapidly, especially in 3D. For Helmholtz problems, iterative methods, however, are not an alternative to direct methods unless a special preconditioner is incorporated. In [9], Erlangga *et al.* introduce a preconditioner based on a discrete approximation to the damped wave equation. The idea of using an operator as a preconditioner for the Helmholtz equation is not new. A prototype of this can be traced back to the work by Bayliss *et al.* [2] in the early 1980s. In [9] it is shown that a Krylov method preconditioned by a damped wave equation can attain convergence, which depends linearly on the frequency (with a small proportional constant).

It is well known that classical multigrid methods cannot be used to solve the Helmholtz equation with the main difficulty being intermediate frequencies. Instead the multigrid can be used only for very fine grids that still maintain some accuracy. Elman and O'Leary [6] suggested a combination of multigrid and Krylov strategies on various grids. Alternatively, Brandt and Livshitz [5] developed a new multigrid strategy based on waves.

In this study we solve the Helmholtz equation with a preconditioned Krylov method. The preconditioner is based on a modified Helmholtz equation which is then solved by classical multigrid techniques utilizing all coarser levels of grids. In contrast to [9], which focused only on second-order accurate schemes, we shall concentrate on the effect of high-order finite-difference schemes on the convergence.

The paper is organized as follows. In Section 2, we shall briefly discussed the Helmholtz equation and the construction of some high-order finite difference schemes. Discretization of boundary conditions is described in Section 3. In Section 4, the iterative method for solving the resultant linear system is discussed, followed by Section 5, which presents numerical results. Some concluding remarks are drawn in Section 6.

2. Discretization

A dissipative term, for the wave equation, translates to either a plus or minus sign in the exponent in the frequency domain and similarly for the radiation condition. However, the two have to be consistent. Hence, depending on how the Sommerfeld radiation condition is implemented it will affect the sign of an additional imaginary dissipative term to the Helmholtz equation. Note that, with zero dissipation, we have

$$v_{xx} + v_{yy} + k^2 v = 0, \tag{2.1}$$

the standard Helmholtz equation. In general, the right-hand side may not be equal to zero. In this paper we only consider a zero right hand side. In [13] the formulae for including the right hand is given.

Let $\phi_{i,j}$ be a numerical approximation to $v(x_i, y_j)$. We wish to use symmetric stencil in both directions x and y. A nine point scheme having these properties has the form

$$A_0\phi_{i,j} + A_s\sigma_s + A_c\sigma_c = 0, \tag{2.2}$$

where

$$\sigma_s = \phi_{i,j+1} + \phi_{i+1,j} + \phi_{i,j-1} + \phi_{i-1,j}$$

is the sum of the values of the mid-side points, while

$$\sigma_c = \phi_{i+1,j+1} + \phi_{i+1,j-1} + \phi_{i-1,j-1} + \phi_{i-1,j+1}$$

is the sum of the values at the corner points.

The standard second order accurate approximation is given by

$$A_0 = -4 + (kh)^2, \quad A_s = 1, \quad A_c = 0. \tag{2.3}$$

A fourth order accurate scheme is given by the relations

$$\begin{aligned} A_0 &= -\frac{10}{3} + (kh)^2 \left(\frac{2}{3} + \frac{\gamma}{36}\right), \\ A_s &= \frac{2}{3} + (kh)^2 \left(\frac{1}{12} - \frac{\gamma}{72}\right), \quad A_c = \frac{1}{6} + (kh)^2 \frac{\gamma}{144}, \end{aligned} \tag{2.4}$$

with γ an arbitrary constant. For variable k the scheme needs to be slightly modified. Instead of $(kh)^2$ appearing in formula (2.4) multiplying the corresponding A in (2.2) now for each node of the stencil the $(kh)^2$ of that node is used. So, for example, $\phi_{i+1,j+1}$ is multiplied by the modified $A_c = \frac{1}{6} + (kh)^2_{i+1,j+1} \frac{\gamma}{144}$. This is then fourth-order accurate also for variable $k(x, y)$ as shown in [12]. By choosing $\gamma = \frac{14}{5}$ and adding $O((kh)^4)$ terms to the coefficients, one can achieve sixth order accuracy for constant k [14]. In this case, for arbitrary δ we have

$$\begin{aligned} A_0 &= -\frac{10}{3} + \frac{67}{90}(kh)^2 + \frac{\delta - 3}{180}(kh)^4, \\ A_s &= \frac{2}{3} + \frac{2}{45}(kh)^2 + \frac{3 - 2\delta}{720}(kh)^4, \\ A_c &= \frac{1}{6} + \frac{7}{360}(kh)^2 + \frac{\delta}{720}(kh)^4. \end{aligned} \tag{2.5}$$

3. Boundary condition

We next wish to construct approximations to the boundary conditions that maintain the high order accuracy. Suppose that along the boundary $x = 1$ we have a simplified radiation condition of the form

$$\frac{\partial u}{\partial x} + i\beta u = 0. \tag{3.1}$$

We approximate the derivative by (dropping the index in the y direction for simplicity)

$$\frac{u_{N+1} - u_{N-1}}{2h} = \frac{\partial u}{\partial x} + \frac{h^2}{6}\frac{\partial^3 u}{\partial x^3} + \frac{h^4}{120}\frac{\partial^5 u}{\partial x^5} + O(h^6), \tag{3.2}$$

where the $N + 1$-th grid point lies outside the computational domain, and hence is a virtual point.

Differentiating (3.1) twice we obtain a relation for the third-order and fifth-order derivatives in (3.2):

$$\frac{\partial^3 u}{\partial x^3} = (-i\beta)^3 u = i\beta^3 u \quad \frac{\partial^5 u}{\partial x^5} = -i\beta^5 u$$

Substitution of the above relation into (3.2) yields

$$\frac{u_{N+1} - u_{N-1}}{2h} = \frac{\partial u}{\partial x} + \frac{i\beta^3 h^2}{6}\left(1 - \frac{\beta^2 h^2}{20}\right) u_N + O(h^6),$$

or

$$\frac{\partial u}{\partial x} = \frac{u_{N+1} - \frac{i\beta^3 h^3}{3}\left(1 - \frac{\beta^2 h^2}{20}\right)\ u_N - u_{N-1}}{2h} + O(h^6)$$

and hence from (3.1),

$$0 = \frac{\partial u}{\partial x} + i\beta u = \frac{u_{N+1} + 2i\beta h\left(1 - \frac{\beta^2 h^2}{6} + \frac{\beta^4 h^4}{120}\right)\ u_N - u_{N-1}}{2h} + O(h^6). \tag{3.3}$$

This is combined with the discretization of the Helmholtz equation (2.2) at the N-th grid point to eliminate the u_{N+1} term.

4. Iterative methods

The resultant linear system of equations has a matrix of coefficients, which is sparse, has complex quantities, is not self adjoint and is not positive definite. Furthermore, the resultant systems are too large, especially in three dimensions to directly solve. In order to solve this linear system, we shall concentrate on iterative methods. In particular we shall consider Krylov space methods.

Numerous authors have found that BICGSTAB [18] is an effective iterative method to solve the linear system arising from the Helmholtz equation. However, this method, along with other Krylov space methods, converges reasonably well only if the equations are preconditioned. Hence, we replace

$$Av = b$$

by the right preconditioning

$$AP^{-1}w = b \qquad Pv = w. \tag{4.1}$$

For the preconditioner to be efficient we require two conditions. The first is that P is an approximation to A and second that $Pv = w$ is much easier to solve than $Av = b$. Over the years two types of preconditioners have been developed [16,17]. In the first kind, the preconditioner utilizes algebraic properties of the matrix. A typical example is ILU, where lower and upper factors are derived which are sparse and LU agrees with A for some set of elements of A.

A different class of preconditioners is based on properties of the differential system, *i.e.* an operator based preconditioner. An example of this is basing the preconditioner on the solution of the Laplace equation [2]. Erlangga *et al.* [7–10] extended this idea by considering a Helmholtz equation with a complex wave number as the preconditioner, namely

$$\Delta v + k_{\text{prec}}^2 v = 0, \qquad k_{\text{prec}}^2 = (1 + i\alpha)k^2, \tag{4.2}$$

with the sign of α depending on formulation 1 or formulation 2 as noted in Section 2, and similarly for the sign in the radiation boundary condition. After discretization, the linear system associated with the preconditioner is then approximately inverted by one multigrid iteration, with point damped Jacobi as a smoother [9]. This will make the preconditioning step cheap, *i.e.*, it is an $O(N)$ work.

In this paper, in addition to the use of damped Jacobi smoother, we also consider a (red-black (RB)) point Gauss-Seidel (GS) smoother with a damping factor $0 < \omega_R \leq 1$. This should be referred to as successive underrelaxation. Nevertheless, we shall use the more standard notation of either SOR or GS. In both case, one pre- and post- smoothing step is employed. The coarsest grid goes to a 2×2 grid independent of k. We compare V, F and W cycles, with bilinear interpolation and full weighting as the restriction and prolongation operator, respectively.

Since the original Helmholtz equation (without a dissipative term) is discretized by a high order method we study what is the appropriate discretization for (4.2).

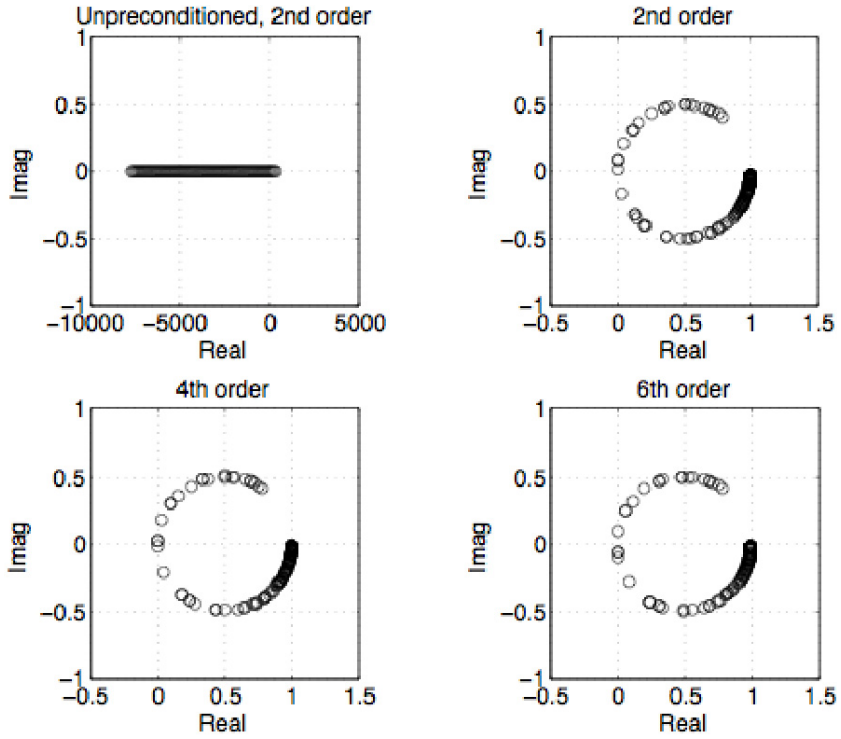

FIGURE 1. Spectrum of the unpreconditioned system (top left) and preconditioned systems with second-order (top right), fourth order (bottom left) and sixth order (bottom right) discretization used in A and P. $k = 20$, grid $= 32^2$, and $\alpha = 0.5$.

4.1. Spectra of the preconditioned system

For Krylov subspace iterative methods, spectrum of the matrix in the linear system to be solved often provides information about convergence. In general, a clustered spectrum is favorable for fast convergence, and this clustering should be sufficiently far from zero. While spectral information is important, this information is difficult to obtain theoretically for general problems. In this section, we present some spectral information, numerically computed to show the effectiveness of the preconditioner (4.2).

Figure 1 shows spectra of AP^{-1} (4.1) for matrices A and P obtained with the same order of discretization, with $k = 20$, 32^2 mesh ($kh = 0.625$), and $\alpha = 0.5$ in (4.2). To obtain these spectra, we assume that P is inverted exactly. Compared with the spectrum of unpreconditioned system, the matrix P effectively reduces the condition number and also clusters the eigenvalues around the circle with a center $C = 0.5$ and a radius $R = 0.5$. This result confirms the algebraic results discussed in [19] and holds independent of the discretization and kh.

Since the preconditioner is defined based on a partial differential equation, we examine the effect of different discretizations of the Helmholtz and preconditioning operators. Figure 2 shows spectra of AP^{-1}, with P based on second order accurate differences and A based either on fourth order or sixth order accurate scheme. Again, similar clustering to Figure 1 is observed, except that some differences in the detail, especially in the half right of the spectra. Both possible combinations, however, lead to very similar spectra. Of course the spectra are not identical but the differences do not seem to impact the convergence rate as demonstrated by the computational results. In general, it is difficult to deduce which will lead to a better convergence. Hence, we find that the order of accuracy of both the scheme and the preconditioner has little effect on the spectrum and this is borne

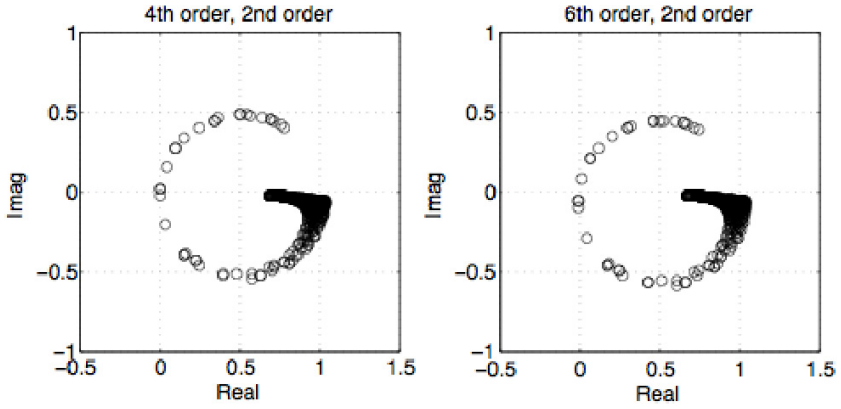

FIGURE 2. Spectrum of the preconditioned system, with a combination of second order discretization for P and fourth order (left) and sixth order (right) for A.

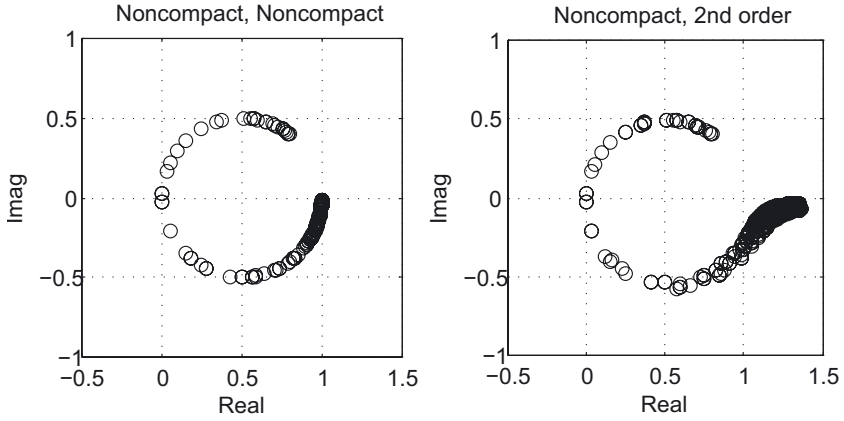

FIGURE 3. Spectrum of the preconditioned system, with a noncompact fourth order scheme preconditioned by noncompact fourth order scheme (left) and a second order scheme (right).

out by the computational results. In Figure 3 we plot the spectrum for a noncompact fourth order scheme preconditioned by the same fourth order noncompact scheme (left figure) and by the standard second order five point scheme (right figure). For the first case, the spectrum is very similar to that in Figure 1. In the latter case, however, we see a tail extending from the spectrum. It is not completely clear what effect this will have on convergence. In any case a noncompact scheme has difficulties near boundaries and has no clear advantage and so no computational results will be presented.

We note that since we solve (4.2) using only one multigrid iteration, the exact inverse of the preconditioning matrix is never available. Hence, the actual spectrum will deviate from the ideal situation.

4.2. Multigrid as the preconditioner solver

Multigrid is typically very efficient for Poisson-like equations. For this class of problems, a smoothing method (typically based on a basic iterative method) effectively reduces high frequency components of error, while a coarse-grid correction takes care the low frequency components. This is however not the case for the Helmholtz equation, unless the wavenumber is sufficiently small, *i.e.*, k^2 less than the smallest eigenvalue of the Laplacian. For high wavenumbers, some eigenvalues will have a negative sign (indefinite). For indefinite systems, Jacobi

TABLE 1. Fourier analysis asymptotic convergence for the Helmholtz equation. μ is the smoothing factor, ρ_2 and ρ_3 are the two- and three-grid convergence, respectively. Bilinear interpolation and full weighting are used, with coarse-grid approximation based on the Galerkin operator and one pre- and post-smoothing.

k	Grid	μ	ρ_2	ρ_3
5	32^2	0.56	0.56	0.56
20	32^2	0.60	1.82	2.04
50	128^2	0.58	1.05	4.04

iteration, for instance, does not converge. In multigrid, the smoothing factor of a basic iterative method is more important than its convergence, and for the Helmholtz equation, Jacobi iteration still gives satisfactory smoothing properties. Due to the increase of indefiniteness, its smoothing properties will deteriorate. (A convergent Jacobi iteration can be based on two-step Jacobi iteration with complex relaxation, as discussed in [11]. But experiments show no advantage of using this method as a smoother for multigrid as compared to a one-step Jacobi iteration). Another source of difficulty comes from the coarse-grid correction. Due to indefiniteness, a sign change may occur to the eigenvalues close to zero after coarse-grid discretization. This will eventually lead to a convergence degradation.

A powerful tool to evaluate properties of a multigrid method is based on Fourier analysis [4,20**?**]. An example of results from Fourier analysis is shown in Table 1 for a W-cycle multigrid method with one pre- and post point-wise Jacobi smoothing, with underrelaxation $\omega_R = 0.5$. (For three-grid methods, F-cycle is identical to W-cycle). The coarse-grid correction is based on a Galerkin approximation, with bilinear interpolation and full weighting as the prolongation and the restriction operator, respectively. For $k = 5$ (small wavenumber), a convergent multigrid method can still be obtained, indicated by a satisfactory smoothing factor μ, and two- and three-grid convergence, denoted by ρ_2 and ρ_3, respectively. For higher wavenumbers, even though the smoothing properties are still satisfactory ($\mu < 1$), results from Fourier analysis indicate problematic coarse-grid correction, especially with three grids.

Having a complex wavenumber in the Helmholtz equation makes the situation quite different. We first analyze the use of damped Jacobi for the complex equation (4.2). To simplify the analysis we consider the discretization of the one dimensional equation with central differences. So we consider

$$v_{j+1} - 2v_j + v_{j-1} + (1 + i\alpha)(kh)^2 v_j = 0. \tag{4.3}$$

The damped Jacobi iteration for this is given by

$$\begin{aligned} \hat{v}_j &= \frac{v_{j+1} + v_{j-1}}{2 - (1 + i\alpha)(kh)^2} \\ v_j &= \omega_R \hat{v}_j + (1 - \omega_R) v_j. \end{aligned} \tag{4.4}$$

The amplification matrix is given by

$$\begin{aligned} G(\theta) &= 1 - \omega_R + \omega_R \frac{\cos(\theta)}{1 - \frac{1+i\alpha}{2} k^2 h^2} = 1 - \omega_R + \omega_R \cos(\theta) \frac{1 + \frac{1+i\alpha}{2} k^2 h^2}{(1 + \frac{k^2 h^2}{2})^2 + \frac{\alpha^2 k^4 h^4}{4}} \\ &= 1 - \omega_R + \frac{\omega_R \cos(\theta) \left(1 + \frac{k^2 h^2}{2}\right)}{1 + k^2 h^2 + \frac{(\alpha^2+1) k^4 h^4}{4}} + i\alpha \frac{\omega_R \cos(\theta)}{1 + k^2 h^2 + \frac{(\alpha^2+1) k^4 h^4}{4}}. \end{aligned}$$

So

$$|G(\theta)|^2 = \left(1 - \omega_R + \frac{\omega_R \cos(\theta) \left(1 + \frac{k^2 h^2}{2}\right)}{1 + k^2 h^2 + \frac{(\alpha^2+1) k^4 h^4}{4}}\right)^2 + \left(\frac{\alpha k^2 h^2 \omega_R \cos(\theta)}{1 + k^2 h^2 + \frac{(\alpha^2+1) k^4 h^4}{4}}\right)^2.$$

TABLE 2. Fourier analysis asymptotic convergence for the Helmholtz equation with complex wavenumber. μ is the smoothing factor, ρ_2 and ρ_3 are the two- and three-grid convergence, respectively. Bilinear interpolation and full weighting are used, with coarse-grid approximation based on the Galerkin operator and one pre- and post-smoothing with point Jacobi and $\omega_R = 0.5$.

	$k = 20$, 32^2 grid			$k = 50$, 128^2 grid			$k = 100$, 512^2 grid		
	μ	ρ_2	ρ_3	μ	ρ_2	ρ_3	μ	ρ_2	ρ_3
2th order	0.60	0.60	0.59	0.58	0.58	0.58	0.57	0.57	0.57
4th order	0.53	0.53	0.73	0.50	0.50	0.50	0.49	0.49	0.49
6th order	0.52	0.52	0.52	0.50	0.50	0.50	0.49	0.49	0.49
Noncompact	0.62	0.62	0.61	0.60	0.60	0.60	0.59	0.59	0.59

We thus, see that α only multiplies terms k^4h^4. Hence, if kh is small enough the effect of α on the convergence is minimal. In other words changing k^2 to a complex coefficient in the Helmholtz equation has a minimal effect on the smoother on fine grids. Only on coarser grids where kh is not small does it have an effect. It is straightforward to check that changes to the discretization of the Helmholtz equation has no effect on these conclusions.

To quantify multigrid convergence properties, we perform Fourier analysis on (4.2), with $\alpha = 0.5$. The associated matrix is obtained via second, fourth, or sixth order finite differences. Results from this analysis are shown in Table 2. In addition to the satisfactory smoothing factor, these results now indicate no problematic coarse-grid corrections. Fourier analysis, including for based on noncompact scheme, also shows that these properties hold almost independently of discretizations used and the wave number k.

5. RESULTS

We consider the following problem on a uniform Cartesian grid with mesh width h. We have Dirichlet boundary conditions on three sides and a simplified absorbing boundary condition (ABC) on the fourth side. The results are fairly independent of the details of the discretization of the original problem. Hence, replacing this poor absorbing boundary condition by an improved one should have little effect on the convergence rates. To verify the high order accuracy of the scheme we consider a system with an explicit solution:

$$\nabla^2 u + k^2 u = 0 \quad \text{in} \quad \Omega = [0,1] \times \left[-\frac{1}{2}, \frac{1}{2}\right], \tag{5.1}$$

$$u\left(x, -\frac{1}{2}\right) = u\left(x, \frac{1}{2}\right) = 0,$$

$$u(0, y) = \cos(\pi y),$$

$$\frac{\partial u}{\partial x} + i\beta u\Big|_{x=1} = 0.$$

The exact solution is $u(x, y) = \cos(\pi y)\mathrm{e}^{-i\beta x}$ with $\pi^2 + \beta^2 = k^2$.

The preconditioned system is similar with k replaced by k_{prec} in the Helmholtz equation but not the ABC. Using this replacement also in the Robin boundary condition has a minor effect on the convergence.

We consider several discretizations of both the original and preconditioned Helmholtz equation. Using a second (2.3) or fourth (2.4) or sixth order (2.5) (with $\delta = 0$) accurate approximation for both equations will be denoted as scheme 2, HO4, HO6 respectively. If we use a fourth order or sixth high order accurate approximation for the original equation but second order for the preconditioned equation it will be denoted as scheme HO42, HO62 respectively. Scheme HO64 is similarly defined.

In Table 3 we check the accuracy of the code. We reduced the residual by 10 orders of magnitude. This was necessary to get the full accuracy on the finest mesh 512×512. On coarser meshes we give the number of

TABLE 3. $k=30$, BICGSTAB with F cycle; damped Jacobi $\alpha = .5$, $\omega_R = 0.4$.

Scheme	Mesh	Tol	# iter	Sec	Tol	# iter	L_2 error	Conv rate
HO42	65×65	10	78	0.39	6	53	1.08×10^{-3}	
HO42	129×129	10	61	1.27	7	47	7.10×10^{-5}	3.93
HO42	257×257	10	54	5.03	8	45	4.57×10^{-6}	3.96
HO42	513×513	10	53	22.22	10	53	2.83×10^{-7}	3.98
HO62	65×65	10	78	0.38	6	53	4.21×10^{-5}	
HO62	129×129	10	64	1.33	7	49	1.39×10^{-6}	4.92
HO62	257×257	10	54	5.16	8	45	4.54×10^{-8}	4.93
HO62	513×513	12	53	21.56	12	53	5.09×10^{-9}	4.94

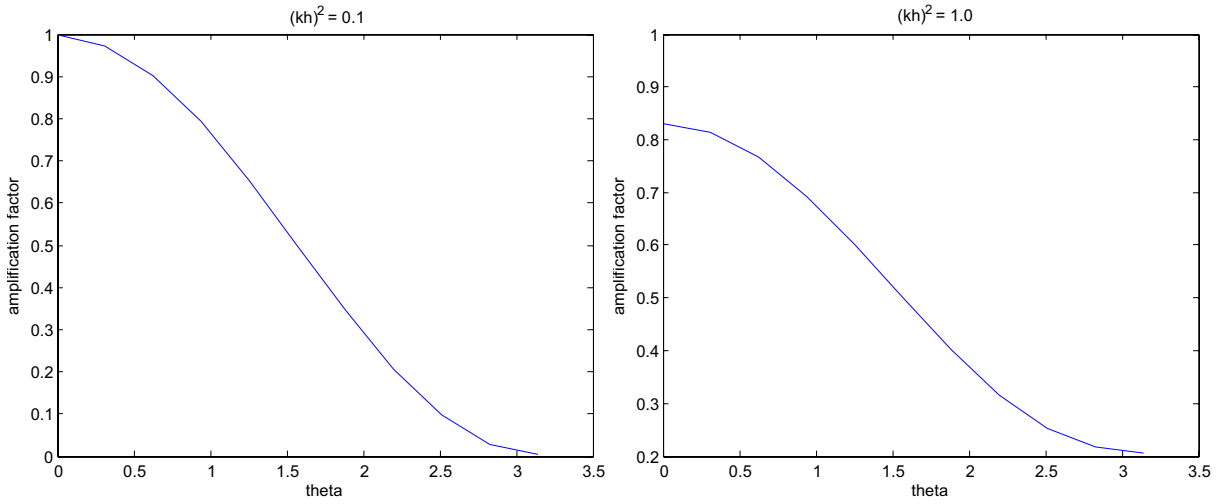

FIGURE 4. Amplification factor of damped Jacobi.

iterations required to reduce the residual by 10 orders of magnitude and also to reduce the residual by enough not to affect the final accuracy. The number of orders of magnitude of reduction is given by "tol". We clearly see the global fourth order accuracy of HO42 including the simplified absorbing boundary condition while HO62 seems to give only fifth order accuracy. In this test the Helmholtz equation was approximated by a fourth or sixth order approximation while the preconditioner was approximated by a standard five point second order finite difference (HO42/HO62). BICGSTAB was used for the solver with point Jacobi used as the smoother for the MG of the preconditioner. To graphically see the amplification factor as a function of θ we plot in Figure 4 the amplification factor for the 512×512 grid.

For the Laplacian the best smoother with SOR is $\omega_r = 1$, *i.e.* Gauss-Seidel. However, for the modified Helmholtz equation we have found that under-relaxation is required. In addition using a red-black ordering increases the convergence rate and allows for parallelization. Hence, in Table 4 we repeat the computations but using RB-SOR with $\omega_R = 0.5$ as the smoother. We see that this converges slower than with the damped Jacobi. Lexicographic GS with the same damping factor was not as robust.

In Table 5 we consider the two smoothers for a range of wavenumbers k. All the numbers in the table are still based on a reduction of 10 orders. This uses a fourth (or sixth) order nine point discretization of the original equation but a five point second order discretization of the modified (preconditioner) equation. The preconditioned equation is solved by a F-cycle multigrid down to a 2×2 grid with one iteration of a damped point Jacobi preconditioner with a relaxation factor, $\omega_R = .5$ or $\omega_R = \frac{2}{3}$. The dependence on ω_R is very slight.

TABLE 4. $k = 30$, HO42, BICGSTAB with F cycle; SOR; $\alpha = .7$, $\omega_R = 0.5$.

Mesh	Tol	# iter	Sec
65×65	10	81	0.43
129×129	10	64	1.62
257×257	10	61	4.56
513×513	10	57	29.20

TABLE 5. BICGSTAB F cycle MG and damped Jacobi, 513×513 mesh, HO42.

k	α	Jacobi $\omega_R = .4$	α	GS $\omega_R = .5$	L_2 error
5	.2	12	.5	12	1.73×10^{-11}
10	.4	25	.5	19	1.28×10^{-9}
20	.5	33	.5	45	3.03×10^{-8}
30	.5	57	.6	70	2.89×10^{-7}
40	.5	63	.7	96	1.33×10^{-6}
50	.5	98	.7	118	3.96×10^{-6}
60	.5	153	.7	208	9.17×10^{-6}

For the highest k the number of iterations is large but this is mainly due to a slowdown after the residual has been reduced 7 orders of magnitude. We also present the optimal α for $k^2_{\text{prec}} = (1 + i\alpha)k^2$. We again see that the damped Jacobi is more efficient and more robust compared with RB-SOR. In Table 6 we solve (5.1) with BICGSTAB and the preconditioned system (4.2) with multigrid. A F cycle is used with 1 iteration of damped Jacobi with $\omega_R = 0.4$ as the smoother, before and after each cycle. However, for higher values of k we found that lower values of ω_R is necessary. We reduce the residual by 10 orders of magnitude for all k.

We see from Table 6 that the convergence for fixed k is independent of the mesh. Similarly, when we use a standard second order five point stencil to approximate the preconditioned equation we get the same convergence independent of whether a second order or high order discretization is used for the original equation. We stress that the accuracy of the total scheme is independent of the treatment of the preconditioner. As k increases, α (the imaginary part of k_{prec}, (4.2)), the convergence slows down. For $k \leq 20$ the number of iterations grows linear in k. However, for larger k it grows faster than linear. If instead of damped Jacobi we use GS underrelaxation then the results are very dependent on the ω_R chosen. For large k we need to choose $\omega_R = 0.5$ and the number of iterations for convergence is larger than with the damped Jacobi. In Figure 5 we display the convergence history for the case $k = 50$, HO42 and a 512×512 grid. We see that the convergence slows down for smaller residual levels. Hence, if we reduce the residual by less than 10 orders of magnitude the time required is much less than indicated in the tables. This is typical of all the cases described here.

A main point is that the convergence rate is insensitive to the details of the discretization both of the equation and the preconditioner. Higher and lower order accuracy schemes for both the Helmholtz equation and the preconditioner converge at similar rates. The damping factor ω_R has a greater influence on the convergence than α as long as α is sufficiently large which depends on k.

We also note that the convergence in Tables 6 and 5 depends on k. Similar behavior is also observed for lower order scheme, as reported in [9]. This convergence degradation when k increases, is relatively mild, compared to the unpreconditioned Bi-CGSTAB. The latter typically shows extremely slow convergence for even small k [8, 10] and an exponential increase of the number of iterations as k increases.

In Table 7 we compare, on a 512×512 mesh, Red-Black SOR, SOR or damped Jacobi as the smoother for the multigrid for the preconditioner. We again reduce the residual by 10 orders of magnitude. We see that RB-SOR requires fewer iterations but more computer time. The extra CPU is caused by needing 2 loops to go through the mesh and moving data to and from a temporary storage. However, this is very dependent on the computer architecture.

TABLE 6. BICGSTAB with F cycle, damped Jacobi.

Mesh	k	ω_R	α	Scheme	γ	# iterations
129×129	5	.4	0	2		18
129×129	5	.3	0	HO42	0	11
513×513	5	.4	0	2		14
513×513	5	.3	0	2		11
513×513	5	.3	0	HO4	0	11
513×513	5	.3	0	HO6		11
129×129	10	.2	.2	2		32
129×129	10	.2	.2	HO4	0	27
129×129	10	.4	.2	HO4	0	37
129×129	10	.4	.2	HO42	0	37
129×129	10	.4	.2	HO6		35
513×513	10	.4	.2	2		58
513×513	10	.2	.2	2		26
513×513	10	.4	.2	HO4	0	64
513×513	10	.2	.4	HO4	0	25
513×513	10	.2	.2	HO4	0	24
513×513	10	.2	.4	HO42	0	26
513×513	10	.4	.2	HO42	0	54
513×513	10	.2	.2	HO42	0	27
513×513	10	.2	.4	HO6		26
513×513	30	.4	.5	2		54
513×513	30	.4	.5	HO4	0	51
513×513	30	.4	.5	HO42	0	56
513×513	30	.4	.5	HO6		50
513×513	30	.4	.5	HO62		53
513×513	30	.4	.7	HO64		56
513×513	50	.4	.5	2		101
513×513	50	.4	.5	HO4	0	95
513×513	50	.4	.5	HO42	0	98
513×513	50	.4	.5	HO6		101
513×513	50	.4	.5	HO62		96
513×513	50	.4	.5	HO64		101

In Table 8 we compare the use of V-, F- or W-cycles within the multigrid used to solve the preconditioned equation. We now only consider the fourth order discretization for the original equation and the five point second order stencil for the preconditioned equation (HO42) on a 513×513 mesh. As usual we reduce the residual by 10 orders of magnitude. The V cycle is less efficient than either F and W cycles especially for larger k.

In Table 9 we examine the effect of k and the mesh on the optimal α. To keep all other parameters the same we choose the HO42 scheme with a F multigrid strategy and a damped Jacobi smoother, $\omega = 0.5$. To account for the pollution effect whenever we double k we increase the number of mesh points in each direction, N, by approximately $k^{\frac{5}{4}} \times 2$ which should keep the error approximately constant. In this table we present various k and the appropriate mesh and the experimentally found best α (*i.e.* the one that minimized the number of iterations). In many cases changing α slightly had little effect on the number of iterations and so the optimal α is actually an interval around that listed in the table. The optimal α seems to grow as approximately $k^{1/2}$. We also present the L_2 and maximum errors, to indeed verify the fourth order accuracy and the pollution effect. In Figure 5 we plot a typical convergence history for the above case with a 512×512 mesh. We see that there is

TABLE 7. BICGSTAB with F cycle and SOR or damped Jacobi smoother, mesh 513×513.

k	α	Smoother	ω_R	Scheme	# iterations	Time (s)
5	.2	RB-SOR	0.4	2	12	6.70
5	.2	SOR	0.4	2	12	5.35
5	.2	JAC	0.4	2	12	5.41
5	.2	RB-SOR	0.4	HO42	12	6.74
5	.2	SOR	0.4	HO42	12	5.36
5	.2	JAC	0.4	HO42	12	5.37
30	.5	RB-SOR	0.4	2	55	28.45
30	.5	SOR	0.8	2	106	41.75
30	.5	JAC	0.4	2	54	21.78
30	.5	RB-SOR	0.4	HO42	57	29.68
30	.5	SOR	0.8	HO42	113	44.97
30	.5	JAC	0.4	HO42	57	20.94
50	.5	RB-SOR	0.5	2	84	42.52
50	.5	SOR	0.5	2	122	47.86
50	.5	JAC	0.5	2	100	39.20
50	.5	RB-SOR	0.5	HO42	84	42.64
50	.5	SOR	0.5	HO42	120	46.58
50	.5	JAC	0.5	HO42	88	34.73

TABLE 8. HO42, BICGSTAB with V,F,W MG cycles; RB-SOR or Jacobi, mesh 513×513.

k	α	Smoother	ω_R	Cycle	# iterations	Time (s)
5	0	RB-SOR	0.4	V	11	5.84
5	.2	RB-SOR	0.4	F	12	7.23
5	.2	RB-SOR	0.4	W	12	7.53
5	0	JAC	0.4	V	12	4.95
5	.2	JAC	0.4	F	14	5.60
5	0	JAC	0.4	W	16	5.03
30	.5	RB-SOR	0.4	V	123	56.31
30	.5	RB-SOR	0.4	F	57	30.94
30	.5	RB-SOR	0.4	W	62	35.35
30	.5	JAC	0.4	V	108	37.56
30	.5	JAC	0.4	F	57	23.51
30	.5	JAC	0.4	W	65	28.11
50	.5	RB-SOR	0.6	V	151	99.37
50	.5	RB-SOR	0.5	F	84	45.47
50	.5	RB-SOR	0.5	W	79	63.34
50	.5	JAC	0.5	V	160	77.40
50	.5	JAC	0.5	F	88	50.14
50	.5	JAC	0.5	W	81	48.29

initially a slow convergence. After settling down we have a rapid convergence until the residual is reduced by about 6 orders of magnitude. Then the convergence slows down to achieve a greater reduction in the residual.

We finally did the computations with a variable k. Thus, we now solve (5.1) with $k(x, y) = k_0 g(x, y)$ and

$$g(x, y) = \begin{cases} 1 & \text{near boundary} \\ 1 + f \sin(px) \sin(qy) & \text{in the interior } f = 0.05 \end{cases}$$

TABLE 9. HO42, BICGSTAB with F,cycle and damped Jacobi. Optimal α and error for various k with $N = k^{5/4}$. mesh 513×513.

k	Mesh $N \times N$	α	# iterations	L_2 error	Max error
4	8×8	0.0	8	4.0610^{-4}	6.73×10^{-4}
8	16×16	0.15	13	2.5210^{-4}	9.00×10^{-4}
16	38×38	0.35	32	4.2310^{-4}	1.04×10^{-3}
32	88×88	0.35	70	4.0910^{-4}	1.06×10^{-3}
64	210×210	0.55	119	4.3010^{-4}	1.09×10^{-3}
128	512×512	0.85	104	4.3110^{-4}	1.06×10^{-3}
256	1210×1210	1.08	177	4.3310^{-4}	1.07×10^{-3}

TABLE 10. Variable k same parameters as Table 9.

k	Mesh $N \times N$	Optimal α	# iterations
4	8×8	0.00	7
8	16×16	0.08	13
16	38×38	0.30	37
32	88×88	0.45	77
64	210×210	0.55	177
128	512×512	0.70	1590
256	1210×1210	1.10	263

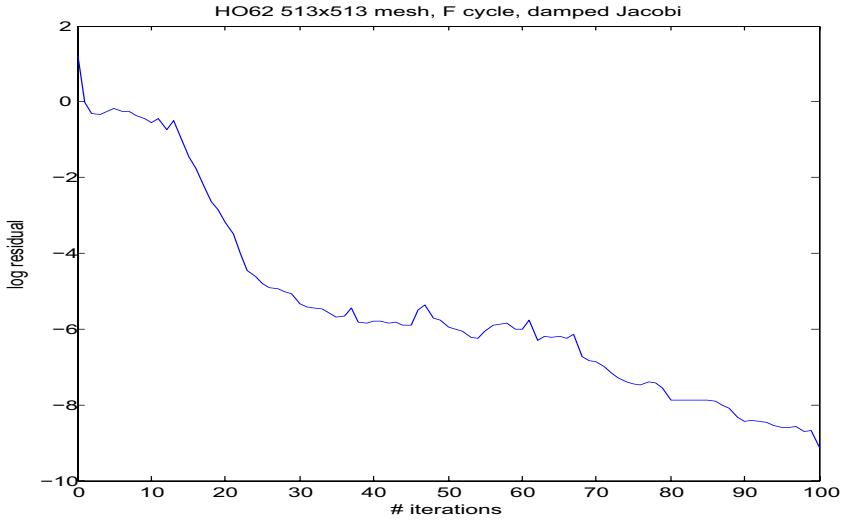

FIGURE 5. typical convergence, HO42, F cycle, damped Jacobi, 512×512 mesh.

p and q are chosen so that there is a complete sine wave perturbation in the interior in a centered subdomain with half the width and length of the entire domain. We ran the exact parameters used in Table 9. The results are presented in Table 10. We see that approximately the same optimal α is needed but that the number of iterations required increases. With greater perturbations than $f = 0.05$ the residuals only converged about 5 orders of magnitude, for large k, rather than the previous 7 orders of magnitude reduction. For k up to about 16 we could considerably increase the amount of variability (f) without any adverse effects. Note, that the perturbation is multiplicative and so smaller in absolute numbers for smaller k. Since only kh appears in the discretization a nonuniform mesh has a similar effect to a variable wavenumber.

6. Conclusion

We have considered second, fourth and sixth order compact discretizations of the Helmholtz equation coupled with an absorbing boundary condition. The resultant linear system has complex coefficients, is not self-adjoint and not positive definite. The system is solved using BICGSTAB. To accelerate the convergence we employ a right preconditioner using a modified Helmholtz equation with a complex wavenumber. This modified equation is then solved by multigrid going down to a 2×2 coarse grid.

The convergence rate of the resultant scheme is relatively independent of the grid. Furthermore, if the preconditioned is discretized by a 5-point second-order stencil then the convergence rate is independent of the discretization of the original equation which determines the final accuracy. The multigrid scheme uses only 1 iteration of a pre and post smoother. Using a SOR scheme for the preconditioner uses fewer iterations but more CPU time than a point damped Jacobi scheme. Both the SOR and damped Jacobi work best with a relaxation parameter $\omega_R = .5$. Because of the complex wavenumber this relaxation factor is different than that appropriate for the Laplace equation. A Fourier analysis is presented to justify the computational results.

Acknowledgements. E. Turkel was supported by a grant from the Ministry of Science, Culture & Sport, Israel & the Ministry of Research, France.

References

[1] I.M. Babuška and S.A. Sauter, Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? *SIAM Rev.* **42** (2000) 451–484.

[2] A. Bayliss, C.I. Goldstein and E. Turkel, An iterative method for the Helmholtz equation. *J. Comput. Phys.* **49** (1983) 443–457.

[3] A. Bayliss, C.I. Goldstein and E. Turkel, On accuracy conditions for the numerical computation of waves. *J. Comput. Phys.* **59** (1985) 396–404.

[4] A. Brandt, Multi-level adaptive solution to the boundary- value problems. *Math. Comp.* **31** (1977) 333-390.

[5] A. Brandt and I. Livshits, *Remarks on the wave-ray Multigrid Solvers for Helmholtz Equations*, *Computational Fluid and Solid Mechanics*, edited by K.J. Bathe. Elsevier (2003) 1871–1871.

[6] H.C. Elman and D.P. O'Leary, Efficient iterative solution of the three dimensional Helmholtz equation. *J. Comput. Phys.* **142** (1998) 163–181.

[7] Y.A. Erlangga, Advances in iterative methods and preconditioners for the Helmholtz equation. *Arch. Comput. Methods Eng.* **15** (2008) 37–66.

[8] Y.A. Erlangga, C. Vuik and C.W. Oosterlee, On a class of preconditioners for the Helmholtz equation. *Appl. Numer. Math.* **50** (2004) 409–425.

[9] Y.A. Erlangga, C.W. Oosterlee and C. Vuik, A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM J. Sci. Comput.* **27** (2006) 1471–1492.

[10] Y.A. Erlangga, C. Vuik and C.W. Oosterlee, Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation. *Appl. Numer. Math.* **56** (2006) 648–666.

[11] G.R. Hadley, A complex Jacobi iterative method for the indefinite Helmholtz equation. *J. Comput. Phys.* **203** (2005) 358–370.

[12] I. Harari and E. Turkel, Accurate finite difference methods for time-harmonic wave propagation. *J. Comput. Phys.* **119** (1995) 252–270.

[13] I. Singer and E. Turkel, High order finite difference methods for the Helmholtz equation. *Comput. Meth. Appl. Mech. Eng.* **163** (1998) 343–358.

[14] I. Singer and E. Turkel, Sixth order accurate finite difference schemes for the Helmholtz equation. *J. Comp. Acous.* **14** (2006) 339–351.

[15] H. Tal-Ezer and E. Turkel, Iterative Solver for the Exterior Helmholtz Problem. *SIAM J. Sci. Comput.* **32** (2010) 463–475.

[16] E. Turkel, Numerical methods and nature. *J. Sci. Comput.* **28** (2006) 549–570.

[17] E. Turkel, *Boundary Conditions and Iterative Schemes for the Helmholtz Equation in Unbounded Regions, Computational Methods for Acoustics Problems*, edited by F. Magoules. Saxe-Coburg Publ. UK (2008).

[18] H.A. van der Vorst, Bi-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.* **13** (1992) 631–644.

[19] M.B. van Gijzen, Y.A. Erlangga and C. Vuik, Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplace precondtioner. *SIAM J. Sci. Comput.* **29** (2006) 1942–1958.

[20] R. Wienands, C.W. Oosterlee, On three-grid Fourier analysis for multigrid. *SIAM J. Sci. Comput.* **22** (2001) 651–671.