Journal of Computational Physics 232 (2013) 272–287

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number

Eli Turkel [a], Dan Gordon [b,*], Rachel Gordon [c], Semyon Tsynkov [d]

[a] School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel
[b] Department of Computer Science, University of Haifa, Haifa 31905, Israel
[c] Department of Aerospace Engineering, The Technion – Israel Institute of Technology, Haifa 32000, Israel
[d] Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA

ARTICLE INFO

Article history:
Received 28 March 2012
Received in revised form 12 August 2012
Accepted 13 August 2012
Available online 5 September 2012

Keywords:
Helmholtz equation
Compact high order schemes
Variable wave number
High frequency
Large wave number
Parallel computing
CARP-CG

ABSTRACT

Several studies have presented compact fourth order accurate finite difference approximation for the Helmholtz equation in two or three dimensions. Several of these formulae allow for the wave number k to be variable. Other papers have extended this further to include variable coefficients within the Laplacian which models non-homogeneous materials in electromagnetism.

Later papers considered more accurate compact sixth order methods but these were restricted to constant k. In this paper we extend these compact sixth order schemes to variable k in both two and three dimensions. Results on 2D and 3D problems with known analytic solutions verify the sixth order accuracy. We demonstrate that for large wave numbers, the second order scheme cannot produce comparable results with reasonable grid sizes.

1. Introduction

Kreiss and Oliger [18] investigated the behavior of errors of central differences in space for the time dependent convection equation. They found that the error grows in time like $t^{1/p}$ where t is the time and p is the order of accuracy of the scheme in space. They concluded that the optimal order of accuracy is between fourth and sixth. More accurate schemes lead to only a minimal reduction in the global error which does not justify the additional work involved. For the problem in Fourier space, i.e., the Helmholtz equation, Bayliss et al. [3] and later Babuška and Sauter [2] found that the number of grid points required for a given accuracy increases with the wave number, but at a slower rate as the order of accuracy of the scheme increases. More precisely, if k is the wave number and N is the mesh size (the number of subintervals in one dimension), and p is the order of a finite difference or finite element scheme, then

$$N = Ck^{\frac{p+1}{p}}, \tag{1}$$

where C is a constant that depends only on the accuracy achieved. This means that if we wish to modify k and maintain the same accuracy, then we should modify N according to (1). This relation, which is called the pollution effect (or dispersion), shows the advantage of high order accurate schemes from the standpoint of efficiency, because the higher the value of p, the closer the quantity $\frac{p+1}{p}$ is to one. Hence, similarly to Kreiss and Oliger, one expects that the optimal order of accuracy of a general scheme for the Helmholtz equation is approximately sixth order.

* Corresponding author.
E-mail addresses: turkel@post.tau.ac.il (E. Turkel), gordon@cs.haifa.ac.il (D. Gordon), rgordon@tx.technion.ac.il (R. Gordon), tsynkov@math.ncsu.edu (S. Tsynkov).

0021-9991/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcp.2012.08.016

Using straightforward central differences, higher order accuracy requires a larger stencil. This has two disadvantages. Firstly, the larger the stencil the more work may be needed to invert the matrix with a larger bandwidth and more non-zero entries. Even more serious are the difficulties near boundaries. A large stencil requires some modification near the boundaries where all the points needed in the stencil are not available. This raises questions of the efficiency and stability of such schemes. These problems are eliminated if one constructs a compact scheme (3×3 in two dimensions and $3 \times 3 \times 3$ in three dimensions). In this case, no additional boundary conditions are required at the discrete level beyond those needed for the Helmholtz equation itself.

It has been shown in [21] that sixth order accuracy is the best that can be achieved, for the Poisson (and certainly Helmholtz) equation in two dimensions using a 3×3 stencil. A lower order compact scheme uses the same stencil and so requires as much work, yet it is less accurate than the sixth order scheme. A second order scheme requires less computation for the same grid dimension because it uses 5 nodes in 2D and 7 nodes in 3D vs. 9 and 27 nodes, respectively, for a compact sixth order accurate scheme. However, one can see from our results (Section 5) that a second order scheme is not competitive due to its lower accuracy and the pollution effect; the extra computations for a compact sixth order scheme are negligible compared to the increased accuracy. Hence, for compact schemes, sixth order accuracy is optimal.

Another advantage of compact schemes pertains to parallel methods that operate as follows: the domain is partitioned into subdomains which are processed in parallel, with some inter-domain communications of the boundary values. An example of this is the CARP-CG algorithm [12] used in this work. With compact high order schemes, the amount of interprocessor communications required by such methods is the same as that of the second order scheme, so no additional communications overhead is incurred by these schemes.

The wave equation for $v(x,y,z,t)$ is given by

$$\frac{1}{c^2(x,y,z)}\frac{\partial^2 v}{\partial t^2} = \Delta v - f(x,y,z,t) \tag{2}$$

where c is the propagation speed of the waves. We assume that the solution is periodic in time (or equivalently, Fourier transform in time). Then $v(x,y,z,t) = e^{i\omega t}u(x,y,z)$ and $f(x,y,z,t) = e^{i\omega t}F(x,y,z)$. Substituting into (2) we get

$$Lu \stackrel{\text{def}}{=} \Delta u + k^2(x,y,z)u = F(x,y,z), \text{ where } k(x,y,z) = \frac{\omega}{c(x,y,z)} \tag{3}$$

$k(x,y,z)$ is called the *wave number*. If the medium is homogeneous, then c (and hence k) are constant. However, in many cases the medium is non-homogeneous, in which case c depends on the position.

Previous research developed compact fourth order accurate finite difference approximation for the Helmholtz equation in two or three dimensions, including a variable wave number [15,22]. These were later extended to polar coordinates and other generalizations of the Laplacian for inhomogeneous media [5,6]. Others considered more accurate sixth order methods but were restricted to constant k [20,24,25]. In this paper we derive a compact sixth order accurate finite difference approximation for the case of variable k in both two and three dimensions. The basic approach to developing the sixth order scheme for variable k uses equation based differencing, where derivatives of the Helmholtz equation are used to eliminate higher order derivatives in the discretization error. This is the same approach that was used before [5,6,15,22,24]. However, the extension to sixth order accuracy for variable coefficients is very nontrivial. Since the emphasis is variable wave number, we shall not present a dispersion analysis, see, e.g., [1,15].

Our aim is to discretize the equation on a compact stencil with sixth order accuracy, assuming the function u is sufficiently smooth. In particular, we assume that the solution has six bounded derivatives and that k and F have four bounded derivatives. If fewer derivatives exist, e.g., in a layered medium, then the current approach will have reduced accuracy depending on the smoothness of k and F. Nevertheless, in [19] it is shown how one can recover the formal accuracy of the scheme for generally shaped boundaries and interfaces using Calderon's projections.

If the source F has a delta-like behavior, then the solution of the differential equation, along with its derivatives, is unbounded at the position of the delta source. In that case, any finite difference scheme, compact or conventional, would lose consistency. One then needs to use an approximate representation of the delta function, and to exclude a neighborhood of the delta source in the error estimates. For example, Hicks [16] suggests using a windowed Sinc function. Interestingly, if the delta function is located at a node then the Sinc function reduces to a discrete delta function. However, since we account for derivatives of the forcing function the contributions will be non-zero at all points. Of course, the window function needs to be sufficiently smooth. This approximation will lead to additional errors. The effect of all of these approximations will be discussed in a future paper. We also assume that the functions k and F and their derivatives are known explicitly. If the derivatives cannot be calculated explicitly, then they can be approximated by finite differences to the appropriate order. This would require a larger stencil for k and F but not u. We note that the reasons given above for the use of a compact scheme apply only to compactness in the stencil for the solution u. Using a non-compact stencil for the wave number k or the forcing term F does not affect the computational effort.

2. Derivation of the compact scheme

Let δ_x and δ_{xx} (and similarly in the other directions) be central difference approximations to the first and second derivatives respectively using the immediate neighbors. So we define:

$$h\delta_x u = \frac{u_{i+1,j,k} - u_{i-1,j,k}}{2}$$
$$h^3\delta_{xyy} u = \frac{u_{i+1,j+1,k} + u_{i+1,j-1,k} - u_{i-1,j+1,k} - u_{i-1,j-1,k} - 2\left(u_{i+1,j,k} - u_{i-1,j,k}\right)}{2}$$

and similarly in the other directions. We also have

$$h^2\left(\delta_{xx} + \delta_{yy} + \delta_{zz}\right)u = u_{i+1,j,k} + u_{i-1,j,k} + u_{i,j+1,k} + u_{i,j-1,k} + u_{i,j,k+1} + u_{i,j,k-1} - 6u_{i,j,k}$$

$$\begin{aligned} h^4\left(\delta_{xxyy} + \delta_{xxzz} + \delta_{yyzz}\right)u = {} & u_{i+1,j+1,k} + u_{i-1,j+1,k} + u_{i+1,j-1,k} + u_{i-1,j-1,k} + u_{i+1,j,k+1} + u_{i-1,j,k+1} + u_{i+1,j,k-1} + u_{i-1,j,k-1} + u_{i,j+1,k+1} \\ & + u_{i,j-1,k+1} + u_{i,j+1,k-1} + u_{i,j-1,k-1} - 4\left(u_{i+1,j,k} + u_{i-1,j,k} + u_{i,j+1,k} + u_{i,j-1,k} + u_{i,j,k+1} + u_{i,j,k-1}\right) \\ & + 12u_{i,j,k} \end{aligned}$$

$$\begin{aligned} h^6\delta_{xxyyzz} u = {} & u_{i+1,j+1,k+1} + u_{i-1,j+1,k+1} + u_{i+1,j-1,k+1} + u_{i-1,j-1,k+1} + u_{i+1,j+1,k-1} + u_{i-1,j+1,k-1} + u_{i+1,j-1,k-1} + u_{i-1,j-1,k-1} \\ & - 2\left(u_{i+1,j,k+1} + u_{i-1,j,k+1} + u_{i,j+1,k+1} + u_{i,j-1,k+1} + u_{i+1,j,k-1} + u_{i-1,j,k-1} + u_{i,j+1,k-1} + u_{i,j-1,k-1}\right. \\ & \left. + u_{i+1,j+1,k}u_{i-1,j+1,k} + u_{i+1,j-1,k} + u_{i-1,j-1,k}\right) + 4\left(u_{i+1,j,k} + u_{i-1,j,k} + u_{i,j+1,k} + u_{i,j-1,k} + u_{i,j,k+1} + u_{i,j,k-1}\right) - 8u_{i,j,k} \end{aligned}$$

By a Taylor series expansion we get

$$\delta_x u = u_x + \frac{h^2}{6}u_{xxx} + \frac{h^4}{120}u_{xxxxx} + O(h^6)$$
$$\delta_{xx} u = u_{xx} + \frac{h^2}{12}u_{xxxx} + \frac{h^4}{360}u_{xxxxxx} + O(h^6) \tag{4}$$

Define the (second order) discrete approximation to the left hand side of (3):

$$\Delta_h u = \delta_{xx} u + \delta_{yy} u + \delta_{zz} u \tag{5a}$$

$$L_h u = \Delta_h u + k^2 u \tag{5b}$$

Then, by a Taylor Series expansion we have

$$L_h u = Lu + \frac{h^2}{12}\left(u_{xxxx} + u_{yyyy} + u_{zzzz}\right) + \frac{h^4}{360}\left(u_{xxxxxx} + u_{yyyyyy} + u_{zzzzzz}\right) + O(h^6) \tag{6}$$

Define:

$$I_1 = u_{xxxx} + u_{yyyy} + u_{zzzz}$$
$$I_2 = u_{xxxxxx} + u_{yyyyyy} + u_{zzzzzz} \tag{7}$$

and

$$Q = \delta_{xx}\delta_{yy}(k^2 u) + \delta_{xx}\delta_{zz}(k^2 u) + \delta_{yy}\delta_{zz}(k^2 u) - (F_{xxyy} + F_{xxzz} + F_{yyzz})$$
$$R = (k^2 u)_{xx} + (k^2 u)_{yy} + (k^2 u)_{zz} = \Delta(k^2 u)$$
$$S = (k^2 u)_{xxxx} + (k^2 u)_{yyyy} + (k^2 u)_{zzzz}$$

We need to approximate I_1 to fourth order accuracy and I_2 to second order accuracy. By a Taylor series expansion we get

$$Lu = L_h u - \frac{h^2}{12}\left(I_1 + O(h^4)\right) - \frac{h^4}{360}\left(I_2 + O(h^2)\right) + O(h^6) = f \tag{8}$$

Lemma 1

$$u_x = \delta_x u + \frac{h^2}{6}\left[\delta_{xyy} u + \delta_{xzz} u + \delta_x(k^2 u) - F_x\right] + O(h^4) \tag{9}$$

Proof

$$\delta_x u = u_x + \frac{h^2}{6}u_{xxx} + O(h^4)$$
$$u_x = \delta_x u - \frac{h^2}{6}u_{xxx} + O(h^4) \tag{10}$$

Differentiating the Helmholtz equation with respect to x we get

$$u_{xxx} = -\left(u_{xyy} + u_{xzz} + (k^2 u)_x\right) + F_x$$

So substituting into (10)

$$u_x = \delta_x u + \frac{h^2}{6}\left(u_{xyy} + u_{xzz} + (k^2u)_x - F_x\right) + O(h^4) \tag{11}$$

We replace the derivatives by central differences with an error $O(h^2)$ and substitute into (11) yielding (9). □
u_y and u_z are derived similarly.

Lemma 2

$$\Delta(k^2u) = k^2F + \left(\Delta(k^2) - k^4\right)u + 2\left((k^2)_x u_x + (k^2)_y u_y + (k^2)_z u_z\right) \tag{12}$$

Proof

$$(k^2u)_{xx} = (k^2)_{xx}u + 2(k^2)_x u_x + k^2 u_{xx}$$

So adding each direction

$$\begin{aligned}\Delta(k^2u) &= \Delta(k^2)u + 2\left(\left(k^2\right)_x u_x + \left(k^2\right)_y u_y + \left(k^2\right)_z u_z\right) + k^2\Delta(u)\\ &= k^2F + \left(\Delta(k^2) - k^4\right)u + 2\left(\left(k^2\right)_x u_x + \left(k^2\right)_y u_y + \left(k^2\right)_z u_z\right) \qquad \square\end{aligned}$$

Combining this with Lemma 1 we have found $\Delta(k^2u)$ to fourth order accuracy.

$$\begin{aligned}\Delta(k^2u) = k^2F &+ \left(\Delta(k^2) - k^4\right)u + 2(k^2)_x\left(\delta_x u + \frac{h^2}{6}\left[\delta_{xyy}u + \delta_{xzz}u + (k^2)_x u - F_x\right]\right)\\ &+ 2(k^2)_y\left(\delta_y u + \frac{h^2}{6}\left[\delta_{xxy}u + \delta_{yzz}u + (k^2)_y u - F_y\right]\right) + 2(k^2)_z\left(\delta_z u + \frac{h^2}{6}\left[\delta_{xxz}u + \delta_{yyz}u + (k^2)_z u - F_z\right]\right)\\ &+ O(h^4)\end{aligned} \tag{13}$$

We wish to approximate $(k^2u)_{xxxx} + (k^2u)_{yyyy} + (k^2u)_{zzzz}$ with second order accuracy. A straightforward expansion requires fourth derivatives of k^2. If k^2 is a complicated formula then its fourth derivatives become exceedingly complex. Instead, we derive another formula that requires only second derivatives of k. Thus, an explicit formula for these derivatives is simpler.

Lemma 3

$$\frac{h^2}{12}S \equiv \frac{h^2}{12}\left[(k^2u)_{xxxx} + (k^2u)_{yyyy} + (k^2u)_{zzzz}\right] = \Delta_h(k^2u) - \Delta(k^2u) + O(h^4) \tag{14}$$

where $\Delta(k^2u)$ is given to fourth order accuracy by (13) and $\Delta_h(k^2u) = (\delta_x^2 + \delta_y^2 + \delta_z^2)u$.

Proof

$$\delta_{xx}(k^2u) = (k^2u)_{xx} + \frac{h^2}{12}(k^2u)_{xxxx} + O(h^4)$$

Adding from each direction we get

$$\frac{h^2}{12}\left[(k^2u)_{xxxx} + (k^2u)_{yyyy} + (k^2u)_{zzzz}\right] = \Delta_h(k^2u) - \Delta(k^2u) + O(h^4) \qquad \square$$

We begin with approximating I_2 to second order accuracy. Differentiating (3) four times, with respect to $xxxx, yyyy, zzzz$ we get

$$\begin{aligned}u_{xxxxxx} + u_{xxxxyy} + u_{xxxxzz} &= -(k^2u)_{xxxx} + F_{xxxx}\\ u_{xxyyyy} + u_{yyyyyy} + u_{yyyyzz} &= -(k^2u)_{yyyy} + F_{yyyy}\\ u_{xxzzzz} + u_{yyzzzz} + u_{zzzzzz} &= -(k^2u)_{zzzz} + F_{zzzz}\end{aligned} \tag{15}$$

Adding we get

$$\begin{aligned} I_2 &\equiv u_{xxxxxx} + u_{yyyyyy} + u_{zzzzzz} \\ &= -\left(u_{xxxxyy} + u_{xxxxzz} + u_{xxyyyy} + u_{yyyyzz} + u_{xxzzzz} + u_{yyzzzz}\right) - \left[(k^2u)_{xxxx} + (k^2u)_{yyyy} + (k^2u)_{zzzz}\right] \\ &\quad + \left[F_{xxxx} + F_{yyyy} + F_{zzzz}\right] \end{aligned} \tag{16}$$

Differentiating the Helmholtz equation with respect to $xxyy, xxzz, yyzz$, we have

$$\begin{aligned} u_{xxxxyy} + u_{xxyyyy} + u_{xxyyzz} &= -(k^2u)_{xxyy} + F_{xxyy} \\ u_{xxxxzz} + u_{xxyyzz} + u_{xxzzzz} &= -(k^2u)_{xxzz} + F_{xxzz} \\ u_{xxyyzz} + u_{yyyyzz} + u_{yyzzzz} &= -(k^2u)_{yyzz} + F_{yyzz} \end{aligned}$$

Adding these 3 equations we get

$$\begin{aligned} u_{xxxxyy} + u_{xxxxzz} + u_{xxyyyy} + u_{yyyyzz} + u_{xxzzzz} + u_{yyzzzz} &= -3u_{xxyyzz} - \left[(k^2u)_{xxyy} + (k^2u)_{xxzz} + (k^2u)_{yyzz}\right] + \left[F_{xxyy} + F_{xxzz} + F_{yyzz}\right] \\ &= -3u_{xxyyzz} - Q + O(h^2) \end{aligned} \tag{17}$$

Inserting (16) and (17) into (7) we get

$$I_2 \equiv u_{xxxxxx} + u_{yyyyyy} + u_{zzzzzz} = 3u_{xxyyzz} + Q - S + \left[F_{xxxx} + F_{yyyy} + F_{zzzz}\right]$$

and so

$$\begin{aligned} \frac{h^4}{360} I_2 &= \frac{h^4}{120}\delta_{xx}\delta_{yy}\delta_{zz}u + \frac{h^4}{360}\left(Q + \left[F_{xxxx} + F_{yyyy} + F_{zzzz}\right]\right) - \frac{h^2}{30}\frac{h^2}{12}S + O(h^6) \\ &= \frac{h^4}{120}\delta_{xx}\delta_{yy}\delta_{zz}u + \frac{h^4}{360}\left(\delta_{xx}\delta_{yy} + \delta_{xx}\delta_{zz} + \delta_{yy}\delta_{zz}\right)(k^2u) - \frac{h^2}{30}\frac{h^2}{12}S \\ &\quad + \frac{h^4}{360}\left[F_{xxxx} + F_{yyyy} + F_{zzzz} - F_{xxyy} - F_{xxzz} - F_{yyzz}\right] + O(h^6) \end{aligned} \tag{18}$$

In Lemma 3 we have found $\frac{h^2}{12}S = (\delta_{xx} + \delta_{yy} + \delta_{zz})(k^2u) - \Delta(k^2u) + O(h^4)$. So combining we get

$$\begin{aligned} \frac{h^4}{360} I_2 &= \frac{h^4}{120}\delta_{xx}\delta_{yy}\delta_{zz}u + \frac{h^4}{360}\left(\delta_{xx}\delta_{yy} + \delta_{xx}\delta_{zz} + \delta_{yy}\delta_{zz}\right)(k^2u) - \frac{h^2}{30}(\delta_{xx} + \delta_{yy} + \delta_{zz})(k^2u) + \frac{h^2}{30}\Delta(k^2u) \\ &\quad + \frac{h^4}{360}\left[F_{xxxx} + F_{yyyy} + F_{zzzz} - F_{xxyy} - F_{xxzz} - F_{yyzz}\right] + O(h^6) \end{aligned} \tag{19}$$

We next consider

$$I_1 = u_{xxxx} + u_{yyyy} + u_{zzzz}$$

which we need to fourth order accuracy. Differentiating the Helmholtz equation twice with respect to xx, yy, zz respectively, we get

$$\begin{aligned} u_{xxxx} + u_{xxyy} + u_{xxzz} &= -(k^2u)_{xx} + F_{xx} \\ u_{xxyy} + u_{yyyy} + u_{yyzz} &= -(k^2u)_{yy} + F_{yy} \\ u_{xxzz} + u_{yyzz} + u_{zzzz} &= -(k^2u)_{zz} + F_{zz} \end{aligned}$$

Adding we get

$$I_1 = u_{xxxx} + u_{yyyy} + u_{zzzz} = -2\left(u_{xxyy} + u_{xxzz} + u_{yyzz}\right) - \Delta(k^2u) + \Delta F$$

But

$$\delta_{xx}\delta_{yy}u = u_{xxyy} + \frac{h^2}{12}(u_{xxxxyy} + u_{xxyyyy}) + O(h^4)$$

and so adding each direction

$$\left(\delta_{xx}\delta_{yy} + \delta_{xx}\delta_{zz} + \delta_{yy}\delta_{zz}\right)u = u_{xxyy} + u_{xxzz} + u_{yyzz} + \frac{h^2}{12}\left(u_{xxxxyy} + u_{xxyyyy} + u_{xxxxzz} + u_{xxzzzz} + u_{yyyyzz} + u_{yyzzzz}\right) + O(h^4)$$

But from Eq. (17) we have

$$\begin{aligned} u_{xxxxyy} + u_{xxxxzz} + u_{xxyyyy} + u_{yyyyzz} + u_{xxzzzz} + u_{yyzzzz} &= -3u_{xxyyzz} - \left[(k^2u)_{xxyy} + (k^2u)_{xxzz} + (k^2u)_{yyzz}\right] + \left[F_{xxyy} + F_{xxzz} + F_{yyzz}\right] \\ &= -3\delta_{xx}\delta_{yy}\delta_{zz}u - Q + O(h^2) \end{aligned}$$

So

$$u_{xxyy}+u_{xxzz}+u_{yyzz}=\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)u+\frac{h^2}{4}\delta_{xx}\delta_{yy}\delta_{zz}u+\frac{h^2}{12}Q+O(h^4)$$

Thus,

$$I_1=-2\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)u-\frac{h^2}{2}\delta_{xx}\delta_{yy}\delta_{zz}u-\frac{h^2}{6}Q-\Delta(k^2u)+\Delta F+O(h^4) \tag{20}$$

We have found already (13) $\Delta(k^2u)$ to fourth order accuracy. So we have found I_1 and I_2 to their necessary accuracy. We also have I_1 to second order

$$I_1=-2\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)u-\Delta_h(k^2u)-k^2F+\Delta F+O(h^2) \tag{21}$$

So the Helmholtz equation approximated to fourth order accuracy is given by

$$\Delta_h u+k^2u+\frac{h^2}{12}\left[\Delta_h(k^2u)+2\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)u\right]=F+\frac{h^2}{12}\Delta F \tag{22}$$

Adding (20) and (19) we have

$$\begin{aligned}\frac{h^2}{12}I_1+\frac{h^4}{360}I_2=&-\frac{h^2}{6}\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)u-\frac{h^2}{30}\left(\delta_{xx}+\delta_{yy}+\delta_{zz}\right)(k^2u)-\frac{h^2}{20}\Delta(k^2u)-\frac{h^4}{30}\delta_{xx}\delta_{yy}\delta_{zz}u\\&-\frac{h^4}{90}\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)(k^2u)+\frac{h^2}{12}\Delta F+\frac{h^4}{360}\left[F_{xxxx}+F_{yyyy}+F_{zzzz}\right]+\frac{h^4}{90}\left[F_{xxyy}+F_{xxzz}+F_{yyzz}\right]\\&+O(h^6)\end{aligned}$$

Thus, the Helmholtz equation $Lu=f$ is approximated by $L_hu-\left(\frac{h^2}{12}I1+\frac{h^4}{360}I_2\right)+O(h^6)=F$. Multiplying by h^2 yields

$$\begin{aligned}&h^2\left(\delta_{xx}+\delta_{yy}+\delta_{zz}\right)\left(1+\frac{k^2h^2}{30}\right)u+k^2h^2u+\frac{h^4}{6}\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)\left(1+\frac{k^2h^2}{15}\right)u+\frac{h^4}{20}\Delta(k^2u)+\frac{h^6}{30}\delta_{xx}\delta_{yy}\delta_{zz}u\\&=h^2\left[F+\frac{h^2}{12}\Delta F+\frac{h^4}{360}\left(F_{xxxx}+F_{yyyy}+F_{zzzz}\right)+\frac{h^4}{90}\left(F_{xxyy}+F_{xxzz}+F_{yyzz}\right)\right]+O(h^8)\end{aligned} \tag{23}$$

where

$$\begin{aligned}h^4\Delta(k^2u)=&h^4\left[k^2F+\left(\Delta(k^2)-k^4\right)u\right]+2(k^2)_xh^3\left(h\delta_xu+\frac{h^3}{6}\left[\delta_{xyy}u+\delta_{xzz}u+\delta_x(k^2u)-F_x\right]\right)\\&+2(k^2)_yh^3\left(h\delta_yu+\frac{h^3}{6}\left[\delta_{xxy}u+\delta_{yzz}u+\delta_y(k^2u)-F_y\right]\right)\\&+2(k^2)_zh^3\left(h\delta_zu+\frac{h^3}{6}\left[\delta_{xxz}u+\delta_{yyz}u+\delta_z(k^2u)-F_z\right]\right)+O(h^8)\end{aligned} \tag{24}$$

This contains terms that depend on F that need to be transferred to the right hand side of the equation. Thus, we require the first derivatives of $k$and $\Delta(k^2)$ either analytically or computationally to fourth order accuracy. We also require fourth derivatives of F either analytically or with fourth order accuracy.

We now rewrite (23) in terms of the points (9 in 2D and 27 in 3D) in the stencil. To simplify the notation we note that all the terms except for parts of $\Delta(k^2u)$ are in self-adjoint form even on the discrete level to sixth order accuracy. However, parts of $\Delta(k^2u)$ are discretely self-adjoint only to fourth order accuracy but lose that property for the sixth order accuracy which will complicate the formulae. Hence, we split our discretization into two parts, A represents the self-adjoint portion while B represents the non-self-adjoint part. Define:

$$\begin{aligned}A=&h^2\left(\delta_{xx}+\delta_{yy}+\delta_{zz}\right)\left(1+\frac{k^2h^2}{30}\right)u+k^2h^2u+\frac{h^4}{6}\left(\delta_{xx}\delta_{yy}+\delta_{xx}\delta_{zz}+\delta_{yy}\delta_{zz}\right)\left(1+\frac{k^2h^2}{15}\right)u+\frac{h^6}{30}\delta_{xx}\delta_{yy}\delta_{zz}u\\&+\frac{h^4}{20}\left[\left(\Delta(k^2)-k^4\right)u\right]\end{aligned} \tag{25}$$

$$\begin{aligned}B=&\frac{2h^3}{20}\left[(k^2)_x\left(h\delta_xu+\frac{h^3}{6}\left[\delta_{xyy}u+\delta_{xzz}u+\delta_x(k^2u)\right]\right)+(k^2)_y\left(h\delta_yu+\frac{h^3}{6}\left[\delta_{xxy}u+\delta_{yzz}u+\delta_y(k^2u)\right]\right)\right.\\&\left.+\left(k^2\right)_z\left(h\delta_zu+\frac{h^3}{6}\left[\delta_{xxz}u+\delta_{yyz}u+\delta_z(k^2u)\right]\right)\right]\end{aligned} \tag{26}$$

$$RHS = h^2\left\{\left(1-\frac{k^2h^2}{20}\right)F+\frac{h^2}{12}\Delta F+\frac{h^4}{360}\left(F_{xxxx}+F_{yyyy}+F_{zzzz}\right)+\frac{h^4}{90}\left(F_{xxyy}+F_{xxzz}+F_{yyzz}\right)+\frac{h^4}{60}\left[(k^2)_xF_x+(k^2)_yF_y+(k^2)_zF_z\right]\right\} \tag{27}$$

Our final equation is then

$$A+B=RHS$$

We first rewrite A in stencil notation.

A_{cc} is the coefficient of the 8 corner points
$(i+1,j+1,k+1),(i-1,j+1,k+1),(i+1,j-1,k+1),(i-1,j-1,k+1),$
$(i+1,j+1,k-1),(i-1,j+1,k-1),(i+1,j-1,k-1),(i-1,j-1,k-1)$

A_{sc} is the coefficient of the 12 corner-side points
$(i+1,j,k+1),(i-1,j,k+1),(i,j+1,k+1),(i,j-1,k+1),$
$(i+1,j,k-1),(i-1,j,k-1),(i,j+1,k-1),(i,j-1,k-1),$
$(i+1,j+1,k),(i-1,j+1,k),(i+1,j-1,k),(i-1,j-1,k)$

A_{ss} is the coefficient of the 6 immediate neighbors
$(i+1,j,k),(i-1,j,k),(i,j+1,k),(i,j-1,k),(i,j,k+1),(i,j,k-1)$

A_0 the coefficient of the center point (i,j,k). Then we have

$$A_0=-\frac{64}{15}+\frac{14k^2h^2}{15}-\frac{k^4h^4}{20}+\frac{h^4}{20}\Delta(k^2) \tag{28}$$
$$A_{ss}=\frac{7}{15}-\frac{k^2h^2}{90}\quad A_{sc}=\frac{1}{10}+\frac{k^2h^2}{90}\quad A_{cc}=\frac{1}{30}$$

We strongly stress that the various points in the corner and side do not have the same coefficients since the k^2 that appears is evaluated at the appropriate stencil point.

For B we do not have a self-adjoint form and so we need to give explicitly all the coefficients. So

$$\begin{aligned}B=&\frac{(k^2)_xh^3}{20}\left[\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i+1,j,k}-\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i-1,j,k}+\frac{1}{6}\left(u_{i+1,j+1,k}+u_{i+1,j-1,k}+u_{i+1,j,k+1}+u_{i+1,j,k-1}-u_{i-1,j+1,k}-u_{i-1,j-1,k}-u_{i-1,j,k+1}-u_{i-1,j,k-1}\right)\right]\\&+\frac{(k^2)_yh^3}{20}\left[\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j+1,k}-\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j-1,k}+\frac{1}{6}\left(u_{i+1,j+1,k}+u_{i-1,j+1,k}+u_{i,j+1,k+1}+u_{i,j+1,k-1}-u_{i+1,j-1,k}-u_{i-1,j-1,k}-u_{i,j-1,k+1}-u_{i,j-1,k-1}\right)\right]\\&+\frac{(k^2)_zh^3}{20}\left[\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j,k+1}-\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j,k-1}+\frac{1}{6}\left(u_{i+1,j,k+1}+u_{i-1,j,k+1}+u_{i,j+1,k+1}+u_{i,j-1,k+1}-u_{i+1,j,k-1}-u_{i-1,j,k-1}-u_{i,j+1,k-1}-u_{i,j-1,k-1}\right)\right]\end{aligned} \tag{29}$$

Note, that for k constant $B=0$ and in (28) $\Delta(k^2)=0$. Reorganizing into the additions to A_0,A_{ss},A_{sc},A_{cc} we get

$$B_0=0 \tag{30}$$

$$\begin{aligned}B_{ss}=&\frac{(k^2)_xh^3}{20}\left[\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i+1,j,k}-\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i-1,j,k}\right]+\frac{(k^2)_yh^3}{20}\left[\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j+1,k}-\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j-1,k}\right]\\&+\frac{(k^2)_zh^3}{20}\left[\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j,k+1}-\left(\frac{1}{3}+\frac{k^2h^2}{6}\right)u_{i,j,k-1}\right]\end{aligned}$$

$$\begin{aligned}B_{sc}=&\frac{h^3}{120}\left[\left((k^2)_x+(k^2)_y\right)u_{i+1,j+1,k}+\left((k^2)_x+(k^2)_z\right)u_{i+1,j,k+1}+\left((k^2)_y+(k^2)_z\right)u_{i,j+1,k+1}+\left((k^2)_x-(k^2)_y\right)u_{i+1,j-1,k}\right.\\&+\left((k^2)_x-(k^2)_z\right)u_{i+1,j,k-1}+\left((k^2)_y-(k^2)_z\right)u_{i,j+1,k-1}-\left((k^2)_x-(k^2)_y\right)u_{i-1,j+1,k}-\left((k^2)_x-(k^2)_z\right)u_{i-1,j,k+1}\\&\left.-\left((k^2)_y-(k^2)_z\right)u_{i,j-1,k+1}-\left((k^2)_x+(k^2)_y\right)u_{i-1,j-1,k}-\left((k^2)_x+(k^2)_z\right)u_{i-1,j,k-1}-\left((k^2)_y+(k^2)_z\right)u_{i,j-1,k-1}\right]\end{aligned}$$

$$B_{cc}=0$$

If we eliminate the $O(h^6)$ we get a fourth order accurate scheme. In this case $A_{cc}=0$ and the stencil contains only 21 points instead of 27. Similarly B greatly simplifies and in particular $B_{sc}=0$.

In two dimensions we have only the corner points A_c and the side points A_s and this reduces to

$$A_c = \frac{1}{6} + \frac{k^2h^2}{90} \quad A_s = \frac{2}{3} + \frac{k^2h^2}{90} \quad A_0 = -\frac{10}{3} + \frac{41k^2h^2}{45} - \frac{k^4h^4}{20} + \frac{h^4}{20}\Delta(k^2) \tag{31}$$

and

$$B_0 = 0 \tag{32}$$

$$B_c = \frac{h^3}{120}\left[\left((k^2)_x + (k^2)_y\right)u_{i+1,j+1} + \left((k^2)_x - (k^2)_y\right)u_{i+1,j-1} - \left((k^2)_x - (k^2)_y\right)u_{i-1,j+1} - \left((k^2)_x + (k^2)_y\right)u_{i-1,j-1}\right]$$

$$B_s = \frac{h^3}{20}\left[(k^2)_x\left(\frac{k^2h^2}{6} + \frac{2}{3}\right)u_{i+1,j} - (k^2)_x\left(\frac{k^2h^2}{6} + \frac{2}{3}\right)u_{i-1,j} + (k^2)_y\left(\frac{k^2h^2}{6} + \frac{2}{3}\right)u_{i,j+1} - (k^2)_y\left(\frac{k^2h^2}{6} + \frac{2}{3}\right)u_{i,j-1}\right]$$

As before k^2 is evaluated at the appropriate stencil point. However, the derivatives $(k^2)_x, (k^2)_y, (k^2)_z, \Delta(k^2)$ are evaluated at the center point of the stencil (i,j,k).

Note,

- For constant k we do not need the formulae for the first derivative (9).
- For constant k this formula is different from those in [20,24,25] but the difference is of higher order.
- Because of the term $\delta_{xx}\delta_{yy}\delta_{zz}u$ there are terms in three dimensions that have no two dimensional counterpart.

3. Boundary conditions

When a Dirichlet boundary condition is imposed then the previous formulae can be used at all interior points. For Neumann boundary conditions, when k is constant, a fourth order accurate technique was developed in [22]. We now develop a sixth order accurate method for a Neumann condition $u_x = g(y,z)$ when g has at least 3 continuous derivatives in y and z. Similar formulae hold in the other directions. To be specific, we consider the coordinate line $i = 0$ and introduce a ghost point $i = -1$ (for ease of notation we omit the j,k indexes). At the boundary $i = 0$ we specify both the Helmholtz equation and the Neumann boundary condition. Furthermore, we take tangential derivatives of the Neumann boundary condition. This offers a considerably more straightforward venue (compared to [20]) toward removing all higher order derivatives from the truncation error expansion in the approximation of the boundary condition. We have:

$$u_{xx} + u_{yy} + u_{zz} + k^2u = F$$
$$\frac{u_{i+1} - u_{i-1}}{2h} = u_x + \frac{h^2}{6}u_{xxx} + \frac{h^4}{120}u_{xxxxx}$$
$$u_{xxx} = -u_{xyy} - u_{xzz} - k^2u_x - (k^2)_xu + F_x$$

$$\begin{aligned}
u_{xxxxx} &= -\left(u_{xxxyy} + u_{xxxzz}\right) - k^2u_{xxx} - 3(k^2)_xu_{xx} - 3(k^2)_{xx}u_x - (k^2)_{xxx}u + F_{xxx} = u_{xyyyy} + 2u_{xyyzz} + u_{xzzzz} + \left(k^2u\right)_{xyy} \\
&+ \left(k^2u\right)_{xzz} - F_{xyy} - F_{xzz} + k^2\left(u_{xyy} + u_{xzz} + k^2u_x + (k^2)_xu - F_x\right) - 3(k^2)_xu_{xx} - 3(k^2)_{xx}u_x - (k^2)_{xxx}u + F_{xxx} = u_{xyyyy} \\
&+ 2u_{xyyzz} + u_{xzzzz} + \left(k^2u_{xyy} + 2\left(k^2\right)_yu_{xy} + \left(k^2\right)_{yy}u_x + \left(k^2\right)_xu_{yy} + 2\left(k^2\right)_{xy}u_y + \left(k^2\right)_{xyy}u\right) + \left(k^2u_{xzz} + 2\left(k^2\right)_zu_{xz}\right. \\
&\left.+\left(k^2\right)_{zz}u_x + \left(k^2\right)_xu_{zz} + 2\left(k^2\right)_{xz}u_z + \left(k^2\right)_{xzz}u\right) + k^2\left(u_{xyy} + u_{xzz} + k^2u_x + (k^2)_xu\right) - 3(k^2)_xu_{xx} - 3(k^2)_{xx}u_x \\
&-(k^2)_{xxx}u + F_{xxx} - F_{xyy} - F_{xzz} - k^2F_x
\end{aligned}$$

Using the Neumann boundary condition $u_x = g$ and its tangential derivatives we obtain:

$$u_{xxx} = -g_{yy} - g_{zz} - k^2g - (k^2)_xu + F_x$$

$$\begin{aligned}
u_{xxxxx} &= g_{yyyy} + 2g_{yyzz} + g_{zzzz} + \left(k^2g_{yy} + 2\left(k^2\right)_yg_y + \left(k^2\right)_{yy}g + \left(k^2\right)_xu_{yy} + 2\left(k^2\right)_{xy}u_y + \left(k^2\right)_{xyy}u\right) + \left(k^2g_{zz} + 2\left(k^2\right)_zg_z\right. \\
&\left.+\left(k^2\right)_{zz}g + \left(k^2\right)_xu_{zz} + 2\left(k^2\right)_{xz}u_z + \left(k^2\right)_{xzz}u\right) + k^2\left(g_{yy} + g_{zz} + k^2g + (k^2)_xu\right) - 3(k^2)_xu_{xx} - 3(k^2)_{xx}g - (k^2)_{xxx}u \\
&+F_{xxx} - F_{xyy} - F_{xzz} - k^2F_x = g_{yyyy} + 2g_{yyzz} + g_{zzzz} + 2k^2\left(g_{yy} + g_{zz}\right) + 2\left(k^2\right)_yg_y + 2\left(k^2\right)_zg_z + \left(\left(k^2\right)_{yy} + \left(k^2\right)_{zz}\right. \\
&\left.-3(k^2)_{xx} + k^4\right)g + (k^2)_x\left(-3u_{xx} + u_{yy} + u_{zz} + k^2u\right) + 2\left(k^2\right)_{xy}u_y + 2\left(k^2\right)_{xz}u_z + \left(\left(k^2\right)_{xyy} + \left(k^2\right)_{xzz} - (k^2)_{xxx}\right)u \\
&+F_{xxx} - F_{xyy} - F_{xzz} - k^2F_x
\end{aligned}$$

Discretizing, centered at $(0,j,k)$, we subsequently get

$$\delta_x u+\frac{h^2}{6}(k^2)_x u-\frac{h^4}{120}\Big(-3(k^2)_x\delta_{xx}u+(k^2)_x(\delta_{yy}u+\delta_{zz}u)+2(k^2)_{xy}\delta_y u+2(k^2)_{xz}\delta_z u+\Big((k^2)_{xyy}+(k^2)_{xzz}+k^2(k^2)_x-(k^2)_{xxx}\Big)u\Big)$$
$$=g-\frac{h^2}{6}\Big(g_{yy}+g_{zz}+k^2g-F_x\Big)+\frac{h^4}{120}\Big[g_{yyyy}+2g_{yyzz}+g_{zzzz}+2k^2(g_{yy}+g_{zz})+2\Big(k^2\Big)_y g_y+2\Big(k^2\Big)_z g_z+\Big(\Big(k^2\Big)_{yy}$$
$$+\Big(k^2\Big)_{zz}-3(k^2)_{xx}+k^4\Big)g+F_{xxx}-F_{xyy}-F_{xzz}-k^2F_x\Big]$$

This yields a stencil of 5 points in 2D and 7 points in 3D for the Neumann boundary points. The matrix is now inverted including the extra artificial line $i=-1$, and similarly for other boundaries with a Neumann condition.

We next consider the case that along the right boundary we have a simplified radiation condition of the form

$$\frac{\partial u}{\partial x}+i\beta u=0 \quad \text{at } x=x_0 \tag{33}$$

Radiation boundary conditions are imposed in the far-field where we assume that the medium, and hence k, is constant. Therefore, we choose a constant β in the neighborhood of the right boundary. Furthermore, since the position of the artificial surface is arbitrary we assume that (33) is valid in a layer about x_0. Hence, we can differentiate (33) in the x direction. We note that having Eq. (33) hold in a layer of finite thickness in the x direction is essentially equivalent to having the solution transition from its domain $x<x_0$ to a duct with one-way wave propagation for $x>x_0$. This, however, does not present a limitation for our subsequent discussion, because we need boundary condition (33) only for methodological purposes to eliminate real eigenvalues, see Section 5.4. We do not necessarily assume that it represents a physical setting.

We approximate the x derivative in (33) by (dropping the index in the y direction for simplicity)

$$\frac{u_{N+1}-u_{N-1}}{2h}=\frac{\partial u}{\partial x}+\frac{h^2}{6}\frac{\partial^3 u}{\partial x^3}+\frac{h^4}{120}\frac{\partial^5 u}{\partial x^5}+O(h^6) \tag{34}$$

where the $N+1$th grid point lies outside the computational domain, and hence is a ghost point. Therefore, we need to eliminate u_{N+1}. Differentiating (33) several times, we obtain a relation for the third-order and fifth-order derivatives in (34):

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

Substitution of the above relation into (34) yields

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

or

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

and hence from (33) we derive

$$0=\frac{\partial u}{\partial x}+i\beta u=\frac{u_{N+1}+2i\beta h\left(1-\frac{\beta^2h^2}{6}+\frac{\beta^4h^4}{120}\right)u_N-u_{N-1}}{2h}+O(h^6) \tag{35}$$

This is combined with the discretization of the Helmholtz equation at the N-th grid point (i.e., the right boundary point) to eliminate the u_{N+1} term (see also [10]).

If we wish to use an accurate absorption condition one can introduce a PML. This requires the approximation of an equation with variable coefficients within a generalized self-adjoint form of the Laplacian in 2D: $\partial_x(a(x,y)\partial_x u)+\partial_y(b(x,y)\partial_y u)$. This equation can be discretized with fourth order accuracy, see [6]. Since the accuracy inside the PML is irrelevant, it does not affect the overall global accuracy inside the physical domain (see [23]).

4. Solver

For three dimensional problems solving directly using Gaussian elimination or some other direct method is not feasible. Hence, we shall use an iterative solver. For the solver, we used the block-parallel CARP-CG algorithm [12]. It is simple to implement on structured and unstructured grids, and it is particularly useful for linear systems with large off-diagonal elements, including cases with discontinuous coefficients [13]. On one processor, it is identical to the CGMN algorithm [4]. CARP-CG has also been successfully used for the high frequency Helmholtz equations [14] with high order schemes from [15,22,25], but only for constant k in 3D. Note that this algorithm is suitable for both symmetric and non-symmetric systems. The following is a very brief description of CARP-CG; for more details, see [12].

CARP-CG is a CG acceleration of CARP (component-averaged row projections) [11], which is a block-parallel extension of KACZ – the Kaczmarz algorithm [17]. KACZ is inherently sequential: starting from an arbitrary initial point, it sweeps through the equations by successively projecting the current iterate towards the hyperplane defined by the next equation. The extent of the projection is determined by the value of a relaxation parameter. It is well known that KACZ is SOR (successive over-relaxation) on the normal equations system $AA^Ty = b, x = A^T$.

CARP divides the equations into blocks (which may overlap) and assigns each block to a processor. Every variable shared by several blocks is copied to all processors whose assigned block of equations contains that variable. The following two steps are then repeated until convergence:

1. Operating in parallel, each processor executes a KACZ sweep on the equations of its assigned block. For each shared variable, the processor uses its copy of that variable.
2. The new value of every shared variable is obtained by averaging all its copies, and it is then distributed to the processors which share it.

For PDE problems defined over some domain, CARP can be used as a form of domain decomposition, with blocks corresponding to subdomains. This way, shared variables are limited to grid points at subdomain boundaries.

The interesting point about CARP is that in some superspace, it is equivalent to KACZ with cyclic relaxation parameters (i.e., each equation has its own fixed relaxation parameter). This provides a convenient theoretical proof of convergence and enables the CG acceleration of CARP as follows: by running CARP in a double (forward and backward) sweep of the equations, one obtains a symmetric positive semi-definite iteration matrix B (even if the original matrix is nonsymmetric). Thus, CG can be applied to B to obtain CARP-CG. B is never formed explicitly: matvec operations with B are obtained by running CARP in a double sweep. This technique extends the one used in [4] to KACZ with cyclic relaxation parameters.

5. Results

We note that in three dimensions a second order accurate scheme requires only a stencil with 7 points. For a fourth order scheme we need 21 points while for the sixth order scheme we require the full compact stencil with 27 points. A Krylov iterative method requires an inner product and so the work increases when the number of points in the stencil increases, even though the stencil remains compact. Thus, a sixth order accurate method should require 2–4 times the number of operations of a second order scheme for the same grid. Alternately, one can use a second order scheme with a mesh about 50% finer in all directions, and then the two algorithms would require about the same storage and computational time per iteration. As will be seen in the results, even when taking this into account, the sixth order method is much more efficient than the second order method. In fact for many three dimensional problems a second order scheme would not give any reasonable results due to memory restrictions on the grid size. This is especially true for high frequencies where, because of the pollution effect, the grid requirement for a second order scheme behaves as $k^{3/2}$, while for a sixth order accurate scheme the required grid increases only as $k^{7/6}$.

5.1. Setup of the numerical experiments

Tests were run on a Supermicro cluster consisting of 12 nodes connected by an Infiniband network. Each node has two Intel Xeon E5520 quad CPUs running at 2.27 GHz, so the cluster can provide a total of 96 cores. The two CPUs share 8 GB of memory and each has its own 8 MB cache. The cluster runs under Debian Linux, and message passing used the MPICH2 public domain MPI package. For the timing experiments, CARP-CG was run in parallel mode using 12 cores, one from each node. All the experiments were run with $u^0 = 0$ as the initial estimate, and the relaxation parameter was taken as 1.6.

We used three standard measures of convergence, defined as follows. Let u^0, u, u^* denote the initial estimate, the current iterate and the analytic solution, respectively. The measures are

- **rel-res** (the relative residual) is $\|Au - b\|_2/\|Au_0 - b\|_2$.
- L_2**-err** (the relative L_2 error) is $\|u^* - u\|_2/\|u^*\|_2$.
- **max-err** denotes the maximal component-wise error: $\max_i|u_i^* - u_i|$.

We consider two and three dimensional problems with nonzero right hand side and known analytic solutions. The examples use the same variable k, with three parameters a, b, c:

$$k(x) = a - b\sin(cx) \quad \text{with } a > b \geqslant 0 \tag{36}$$

Thus, a and b control the range of values of k, and $|\,c\,|$ controls the number of oscillations of k in the domain. Fig. 1 shows a plot of $k(x)$.

5.2. Problem 1 (2D)

We consider the Helmholtz equation with the following analytical solution.

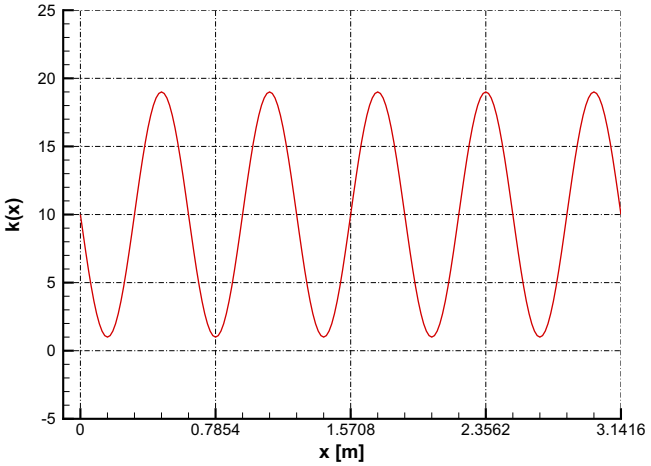

Fig. 1. Plot of $k(x)$ for $a = 10,\ b = 9,\ c = 10$.

$$u(x,y) = e^{-\frac{k(x)}{c}}\sin(\beta y), \quad \text{where } \beta = \sqrt{a^2 + b^2}$$

defined over the domain $D = \pi \times \pi$. A 3D plot of $u(x,y)$ with $a = 10,\ b = 9,\ c = 10$ is shown in Fig. 2.

The right hand side is

$$Lu = \Delta u + k^2 u = F = -b(2a + c)\sin(cx)e^{-\frac{k(x)}{c}}\sin(\beta y)$$

Choosing $c = -2a$ gives us a homogeneous case. For the computations, the derivatives of k and F are evaluated analytically. The number of grid points per wavelength is given by $N_g = \frac{2\pi}{kh} = \frac{2N}{k}$. Note that k is the dimensional wave number. Since the length of the domain is π, the non-dimensional wave number is $\bar{k} = \pi k$.

We first consider two cases, both with $c = 10$, so k has 5 periods as in Fig. 1.

Case 1a: $a = 10,\ b = 4 (6 \leqslant k \leqslant 14, 18.85 \leqslant \bar{k} \leqslant 43.98)$.

Case 1b: $a = 10,\ b = 9 (1 \leqslant k \leqslant 19, 3.14 \leqslant \bar{k} \leqslant 59.69)$.

Table 1 shows the results of the second and sixth order accurate schemes for cases 1a and 1b of Problem 1, for various grid sizes. Also shown are the minimal values of N_g in each case. The second order accurate scheme with a grid of 403×403 has an error which is about 5–8 times larger than that of the sixth order accurate scheme with a mesh of 52×52 (i.e. $\frac{1}{60}$th the number of grid points of the second order scheme).

We now present some plots of the relative residual and error for a third case.

Case 1c: $a = 5,\ b = 2,\ c = 10 (3 \leqslant k \leqslant 7, 9.425 \leqslant \bar{k} \leqslant 22)$.

Fig. 3 shows the convergence of the relative residual for Case 1c with a grid of 112×112 (the 27×27 grid will be explained below). Note that both schemes show a good convergence, with the sixth order scheme being somewhat better.

Fig. 4 shows the L_2-error for Case 1c, with the second and sixth order schemes. We can see that on the 112×112 grid, the sixth order scheme produces much better results, even though the relative residual results were not so different. We can only conclude from this that the second order scheme does not model the problem well on this grid, and that relative residual results alone can be misleading. Furthermore, if we are satisfied with an error goal of 10^{-2} (which is the error obtained with the second order scheme), then we can achieve it with the sixth order scheme on a much smaller grid of 27×27, in only 0.5% of the time. The fast convergence of the relative residual on this smaller grid is shown in Fig. 3. These results clearly demonstrate the computational advantage of the sixth order scheme over the second order scheme.

5.3. Problem 2 (3D)

Our 3D example uses the same $k(x)$ as before, but u is extended to three dimensions as follows.

$$u(x,y) = e^{-\frac{k(x)}{c}}\sin(\beta y)\sin(\gamma z), \ \text{where } \beta^2 + \gamma^2 = a^2 + b^2$$

defined over the domain $D = \pi \times \pi \times \pi$. The right hand side is

$$Lu \equiv \Delta u + k^2 u = F = -b(2a + c)\sin(cx)e^{-\frac{k(x)}{c}}\sin(\beta y)\sin(\gamma z)$$

The three dimensional example consists of 3 cases.

Case 2a: $a = 10,\ b = 9,\ c = 10,\ \gamma = 9 (1 \leqslant k \leqslant 19, 3.14 \leqslant \bar{k} \leqslant 59.69)$.

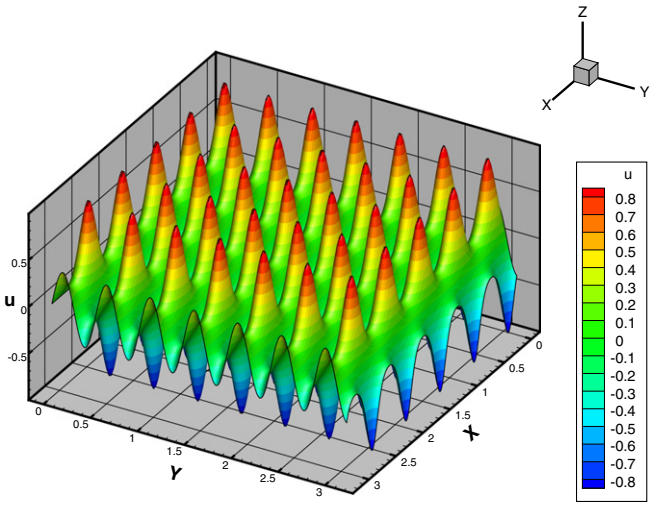

Fig. 2. Plot of $u(x,y)$ for $a = 10,\ b = 9,\ c = 10$.

Table 1
Problem 1 (2D), Cases 1a and 1b: second and sixth order error results for various grid sizes, and the corresponding minimal values of N_g.

a	b	N	min N_g	2nd order		6th order	
				L_2-err	max-err	L_2-err	max-err
10	4	52	7.29	2.62E−1	2.97E−1	5.19E−4	5.18E−4
		103	14.57	6.78E−2	9.16E−2	7.94E−6	8.19E−6
		203	28.86	1.59E−2	1.95E−2	1.31E−7	1.36E−7
		403	57.43	3.96E−3	4.73E−3	2.11E−9	2.19E−9
10	9	52	5.37	8.29E−2	1.08E−1	3.40E−4	5.29E−4
		103	10.74	2.08E−2	2.75E−2	4.73E−6	7.59E−6
		203	21.26	6.95E−3	1.02E−2	7.64E−8	1.25E−7
		403	42.32	1.68E−3	2.64E−3	1.22E−9	2.00E−9

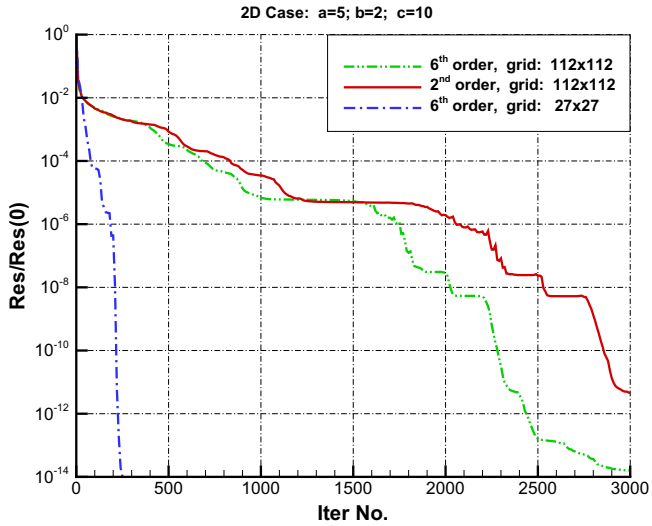

Fig. 3. Problem 1 (2D), Case 1c: relative residual for the second and sixth order schemes.

For this case, we compare the efficiency of the second and sixth order schemes in 3D by finding the number of iterations and CPU time required for two predetermined L_2 error goals of 0.01 and 0.001. The results in Table 2 show the obvious supe-

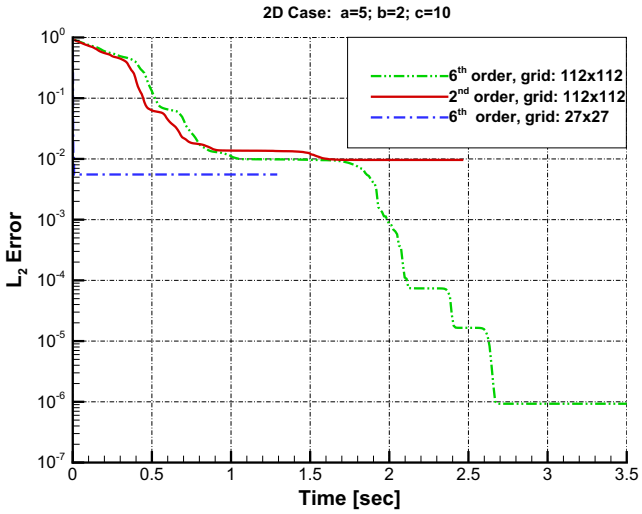

Fig. 4. Problem 1 (2D), Case 1c: L_2-error for the second and sixth order schemes.

Table 2
Problem 2 (3D), Case 2a: comparison of second and sixth order schemes for a fixed error goal. $a = 10,\ b = 9,\ c = 10,\ \gamma = 9$.

L_2-err goal	Scheme	N	min N_g	# iter.	Time (s)	Time ratio
0.01	2nd ord.	103	10.74	540	5.52	17
	6th ord.	31	3.16	251	0.33	
0.001	2nd ord.	333	34.95	1970	703	696
	6th ord.	45	4.63	350	1.01	

Table 3
Problem 2 (3D), Case 2b: grid sizes, number of iterations and time required to reach two error goals, for $a = 80,\ b = 40,\ \gamma = 40$, and various values of the oscillation parameter c. The minimal values of N_g are also shown.

L_2-err goal	c	N	min N_g	# iter.	Time (sec)
0.01	10	147	2.43	10	0.85
	50	187	3.10	26	4.31
	70	216	3.58	137	33.4
	80	242	4.02	425	147
0.001	10	229	3.80	200	122
	50	266	4.42	280	289
	70	312	5.18	893	642
	80	> 402	(See text for explanation)		

riority of the sixth order scheme with CPU gain factors of about 17 (1% error) and 696 (0.1% error). So the gains are especially impressive for the higher accuracy goal.

Case 2b: $a = 80,\ b = 40,\ \gamma = 40 (40 \leqslant k \leqslant 120, 125.66 \leqslant \bar{k} \leqslant 376.99)$.

In Table 3 we find the mesh needed to achieve 1% and 0.1% accuracy, for various values of the oscillation parameter c. Also shown are the number of iterations and runtimes, as well as the minimal values of N_g for each grid size. Only the results for the sixth order scheme are shown because the second order scheme could not achieve the error goals with grids of manageable sizes. Note that for $c = 10$, we get 1% accuracy with a grid size of $N = 147$, which is less than 2.5 points per wavelength! An important point to observe here is that the number of oscillations of k (determined by c) has a very significant effect on the required grid size and hence on the time to achieve the required goal. There is a very obvious reason for this: the grid should not only sample the waves, but for variable k, it should also be fine enough to provide a good sampling of k. For $c = 80$, our resources did not suffice to reach the error goal of 0.001; the best we could do, with a grid of 402^3, was to reach an error of 0.001736 in 10,000 iterations, in 267 min.

Case 2c: This case consists of one example of variable k (with $a = 80,\ b = 40,\ \gamma = 60$) and three examples of constant valued k (with $a = 40, 80, 120$ and $b = 0$). $c = 10$ in all cases, and γ is chosen so that $\gamma \approx \beta$. A fixed grid of 147^3 was used for all four examples. Table 4 shows the L_2 error values of these four examples. These results show that the error with

Table 4
Problem 2 (3D), Case 2c: comparison of variable k (first row) with constant values of k, on a fixed grid of 147^3 with $c = 10$, for the sixth order scheme.

k	N_g	a	b	γ	L_2-err
40–120	2.43–7.30	80	40	60	9.43E−3
40	7.30	40	0	30	9.65E−3
80	3.65	80	0	60	6.72E−1
120	2.43	120	0	85	9.55E−1

the oscillatory k is approximately equal to the error obtained when k is fixed at the minimum value of the oscillatory k. Note that these results were obtained with a moderate value of the oscillation parameter c (compare with Table 3).

Note further from line 2 of Table 4 that for an L_2 error of 0.01, we only need $N_g = 7.30$ for $k = 40$ ($\bar{k} \approx 125$). For other wave numbers the required value of N_g is affected by pollution. However, for a sixth order accurate scheme this effect is quite mild. To see this, consider the pollution formula (1) (according to which $N \approx k^{(p+1)/p}$), and the definition of N_g as $\frac{2N}{k}$, from which we get $N_g \approx k^{1/p}$. The constant of proportionality, as calculated from the line with $k = 40$, is $7.3/40^{1/6} = 3.95$. So, for example, if we wish to double k to 80 and maintain the same accuracy of 0.01, the required number of grid points per wave length should be $N_g = 3.95 \times 80^{1/6} = 8.2$. A similar calculation with the pollution formula (1) shows that the mesh size should be increased from 147 to 330.

Fig. 5 shows the L_2-error for the second and sixth order schemes for Case 2c corresponding to the oscillatory example in Table 4 ($a = 80,\ b = 40,\ c = 10,\ \gamma = 60$). On the 402^3 grid, the sixth order scheme provides better accuracy. Also, we can get the accuracy of the second order scheme (3×10^{-2}) with the sixth order scheme in 1% of the time by using a grid of 127^3.

5.4. Verification of the pollution formula

As discussed in the introduction, the number of grid points needed to maintain a given error grows as k increases. Hence, in order to achieve good accuracy for problems with high frequency, a low order finite difference or finite element method requires a very fine grid, i.e., a large grid dimension. However, this "pollution" effect is mitigated by using high order accurate schemes. For these schemes (non-spectral), a point of diminishing returns occurs, where the extra work associated with the higher order of approximation no longer justifies the small decrease in the error. As previously shown, for compact schemes this point is achieved at sixth order accuracy. For fourth order accurate schemes, the connection between $k$and h has been verified in [5,10].

When a Dirichlet boundary condition is specified along the entire boundary, we found that the accuracy we obtained deviated from that predicted by the pollution formula (1). To explain this, we recall that if the Helmholtz equation $\Delta u + k^2 u = 0$ has a non-trivial solution u on a given domain subject to zero Dirichlet boundary conditions, then $-k^2$ is called an eigenvalue of the Laplace operator, see, e.g., [8, Chapter V]. In this case, the eigenfunction can be added to a solution of any non-homogeneous Dirichlet problem for the Helmholtz equation on the same domain. In other words, the solution to this problem is not unique, which is called a resonance. For the domain $[0, \pi] \times [0, \pi]$, $-k^2$ is an eigenvalue if $k^2 = m^2 + n^2$, where

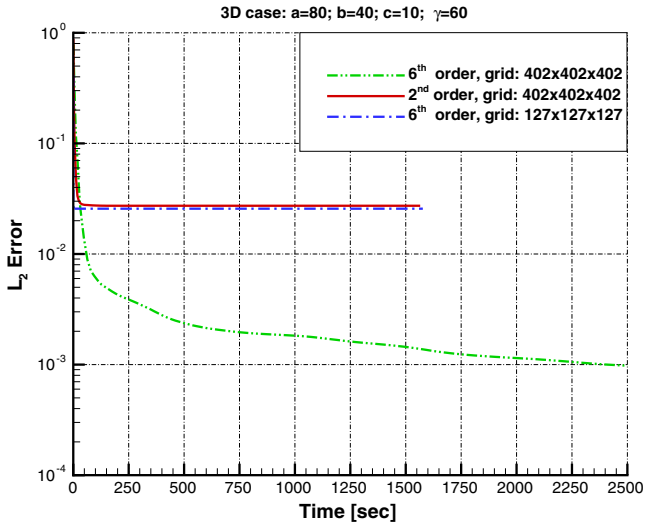

Fig. 5. Problem 2 (3D), Case 2c: L_2 error for the second and sixth order schemes.

Table 5
Problem 3: L_2 error obtained with the standard second order scheme and the fourth order scheme of [22] for different values of N and k.

N	2nd order		4th order [22]	
	k	L_2-err	k	L_2-err
40	6.2996	1.33×10^{-1}	5.7435	8.46×10^{-4}
80	10	1.21×10^{-1}	10	8.87×10^{-4}
120	13.1037	1.23×10^{-1}	13.8316	9.07×10^{-4}
160	15.8740	1.16×10^{-1}	17.4110	8.90×10^{-4}
200	18.4202	1.19×10^{-1}	20.8138	9.09×10^{-4}
240	20.8008	1.14×10^{-1}	24.0832	8.97×10^{-4}
280	23.0522	1.20×10^{-1}	27.2430	8.94×10^{-4}
320	25.1984	1.21×10^{-1}	30.3143	8.95×10^{-4}
	$N = 2.5298 \times k^{3/2}$		$N = 4.4987 \times k^{5/4}$	

Table 6
Problem 3: L_2 error obtained with the sixth order schemes of [10] and the present scheme for different values of N and k.

N	k	L_2-err [10]	L_2-err [new scheme]
40	5.5204	4.49×10^{-6}	1.95×10^{-6}
80	10	5.13×10^{-6}	1.81×10^{-6}
120	14.1558	5.47×10^{-6}	1.42×10^{-6}
160	18.1145	4.49×10^{-6}	1.44×10^{-6}
200	21.9327	5.42×10^{-6}	1.59×10^{-6}
240	25.6425	5.43×10^{-6}	1.59×10^{-6}
280	29.2647	5.56×10^{-6}	1.36×10^{-6}
320	32.8134	5.45×10^{-6}	1.57×10^{-6}
$N = 5.4503 \times k^{7/6}$			

$m, n \geqslant 0$ are integers, and the corresponding eigenfunction is $u(x, y) = \sin(nx)\sin(my)$. For the discretization, the eigenvalues may appear nearby the exact value depending on the details of the scheme and the grid. Therefore, when $k^2 \approx m^2 + n^2$, the resulting system matrix is close to singular. The near singular behavior of the system matrix presents a serious problem for direct methods. However, the iterative projection methods, such as Kaczmarz (and hence CARP-CG), converge to a solution which depends on the initial guess [4]. This explains the irregular behavior of the error that we observed. The chances of $-k^2$ being close to an eigenvalue cannot be ignored: for example, 27.5% of all integers between 1 and 10,000 are sums of two squares, and this affects the values of k between 1 and 100. We will therefore use the following problem, which has a unique solution, to verify the pollution formula.

Problem 3 (2D): Consider a homogeneous Helmholtz equation on the domain $[0, \pi] \times [0, \pi]$. The boundary condition at the side $y = \pi$ is taken as $u_y - i\beta u = 0$ [cf. formula (33)], and is implemented in the sense of (34), (35). Dirichlet boundary conditions are given on the remaining three sides. The solution (with constant k) is given by $u = \sin(nx)e^{i\beta y}$, with $n^2 + \beta^2 = k^2$, and we choose $n = 1$. The reason for using a Sommerfeld-type boundary condition $u_y - i\beta u = 0$ at $y = \pi$ is purely methodological: to move the eigenvalues to the complex plane and hence eliminate any possibility of resonances at real eigenvalues; see [26, Chapter VII], as well as [9, Chapter IV, Section 5]. From this standpoint, any boundary condition that would help us do that serves the purpose. In doing so, it need not be physically relevant; we rather want to use it simply as a means of removing the real eigenvalues. At the same time, it does need to admit a sixth order accurate approximation by compact finite differences. For the Sommerfeld-type condition $u_y - i\beta u = 0$, this can be achieved under the assumptions made in Section 3.

We use Problem 3 to verify the pollution formula (1). We chose $N = 80$ and $k = 10$ as the base from which to calculate the relation $N = Ck^{(p+1)/p}$, where C is a constant calculated as $C = 80/10^{(p+1)/p}$. Table 5 shows the relative L_2 errors for the standard second order scheme and the fourth order scheme of Singer and Turkel [22]. Table 6 shows the results for two sixth order schemes: the scheme from [10] and the present scheme with k constant. We find that the results with the new scheme are somewhat better than those of [10]. The values of C are also given in the tables. These results provide a good numerical validation of the pollution formula (1).

6. Conclusions

We have developed a compact sixth order accurate finite difference scheme for the Helmholtz equation with variable wave number for 2D and 3D. We have also derived a sixth order accurate approximation to a Neumann boundary condition and a simplified farfield absorbing boundary condition. The new scheme was tested in both two and three dimensions. The

scheme allows, in particular, the accurate solution of acoustic wave propagation problems in the frequency space for heterogeneous media. This setting frequently occurs in geophysical applications.

Computational experiments verify that the new compact scheme is much more efficient than a second order scheme. Since the scheme is compact, no extra numerical boundary conditions are needed. Coupled with the CARP-CG iterative algorithm, the scheme has enabled solution of high frequency problems in three dimensions that were beyond the computer resources of a multi-processor parallel computer. Our results also strengthen the finding from [14] that relative residual results with the second order scheme can be misleading: even if they indicate a good convergence of the relative residual, the actual error from the true solution can be very significant.

The pollution effect states that as the wave number k increases, the required number of subintervals (per domain side), N, to obtain a given accuracy increases according to the relation $N \approx k^{(p+1)/p}$, where p is the order of accuracy of the scheme. Thus, N increases faster than linear with k, but the higher the accuracy, the slower the growth in N. Hence, higher order accuracy is even more important for high frequency problems. We have also verified numerically this relation between k and N. In addition, the iterative method CARP-CG also works even better for high frequency problems. Hence, the combination of the sixth order accurate discretization with CARP-CG leads to a very efficient overall algorithm for the high frequency Helmholtz equation in non-homogeneous media.

We have also studied the accuracy of the scheme when the variable wave number k itself is highly oscillatory. It was found that the accuracy and convergence were equally good even when there was a large ratio between $k_{\max}$ and $k_{\min}$. For a moderate number of oscillations within the domain (e.g., 10), the accuracy obtained was close to that obtained for constant $k = k_{\min}$. We also found that as the number of oscillations within k increased, the grid had to be refined in order to maintain a given accuracy. Thus, it was not sufficient to choose a mesh based only on $k_{\min}$ but one also needed to account for resolving the oscillations within k.

Acknowledgments

The authors thank the anonymous reviewers for their useful comments. The research of the first and fourth authors was partially supported by the US-Israel Binational Science Foundation (BSF) under Grant No. 2008094. The work of the fourth author was also partially supported by the US ARO under Grant No. W911NF-11-1-0384.

References

[1] M. Ainsworth, Discrete dispersion relation for hp-version finite element approximation at high wave number, SIAM J. Numer. Anal. 42 (2) (2004) 553–575.
[2] I.M. Babuška, 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. Reprint of SIAM J. Numer. Anal. 34 (1997), 2392–2423..
[3] A. Bayliss, C.I. Goldstein, E. Turkel, On accuracy conditions for the numerical computation of waves, J. Comput. Phys. 59 (3) (1985) 396–404.
[4] A. Björck, T. Elfving, Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations, BIT 19 (1979) 145–163.
[5] S. Britt, S. Tsynkov, E. Turkel, A compact fourth order scheme for the Helmholtz equation in polar coordinates, J. Sci. Comput. 45 (2010) 26–47.
[6] S. Britt, S. Tsynkov, E. Turkel, Numerical simulation of time-harmonic waves in inhomogeneous media using compact high order schemes, Commun. Comput. Phys. 9 (3) (2011) 520–541.
[7] R. Courant, D. Hilbert, Methods of Mathematical Physics, vol. I, Wiley Interscience Publishers, New York, 1953.
[8] R. Courant, D. Hilbert, Methods of Mathematical Physics, vol. II, Wiley Interscience Publishers, New York, 1962.
[9] Y. Erlangga, E. Turkel, Iterative schemes for high order discretizations to the exterior Helmholtz equation, ESAIM: Math. Model. Numer. Anal. 46 (2012) 647–660.
[10] D. Gordon, R. Gordon, Component-averaged row projections: a robust, block-parallel scheme for sparse linear systems, SIAM J. Sci. Comput. 27 (2005) 1092–1117.
[11] D. Gordon, R. Gordon, CARP-CG: a robust and efficient parallel solver for linear systems. Applied to strongly convection dominated PDEs, Parallel Comput. 36 (2010) 495–515.
[12] D. Gordon, R. Gordon, Solution methods for linear systems with large off-diagonal elements and discontinuous coefficients, Comput. Model. Eng. Sci. 53 (2009) 23–45.
[13] D. Gordon, R. Gordon, Parallel solution of high frequency Helmholtz equations using high order finite difference schemes, Appl. Math. Comput. 218 (21) (2012) 10737–10754.
[14] I. Harari, E. Turkel, Accurate finite difference methods for time-harmonic wave propagation, J. Comput. Phys. 119 (2) (1995) 252–270.
[15] G.J. Hicks, Arbitrary source and receiver positioning in finite difference schemes using Kaiser windowed sinc functions, Geophysics 67 (2002) 156–166.
[16] S. Kaczmarz, Angenäherte Auflösung von Systemen linearer Gleichungen, Bull. Acad. Polonaise Sci. Lett. A35 (1937) 355–357.
[17] H.-O. Kreiss, J. Oliger, Comparison of accurate methods for the integration of hyperbolic equations, Tellus 24 (1972) 199–215.
[18] M. Medvinsky, S. Tsynkov, E. Turkel, The method of difference potentials for the Helmholtz equation using compact high order schemes, J Sci Comput (2012) 53:150–193, and Erratum: J Sci Comput, in press, doi: 10.1007/s10915-012-9638-z.
[19] M. Nabavi, K. Siddiqui, J. Dargahi, A new 9-point sixth-order accurate compact finite-difference method for the Helmholtz equation, J. Sound Vib. 307 (2007) 972–982.
[20] S.O. Settle, C.C. Douglas, I. Kim, D. Sheen, On the derivation of highest-order compact finite difference schemes for the one- and two-dimensional Poisson equation with Dirichlet boundary conditions, SIAM J. Numer. Anal., submitted for publication. http://iamcs.tamu.edu/file_dl.php?type=preprint&preprint_id=385.
[21] I. Singer, E. Turkel, High-order finite difference methods for the Helmholtz equation, Comput. Methods Appl. Mech. Engrg. 163 (1–4) (1998) 343–358.
[22] I. Singer, E. Turkel, A perfectly matched layer for the Helmholtz equation in a semi-infinite strip, J. Comput. Phys. 201 (2004) 439–465.
[23] I. Singer, E. Turkel, Sixth order accurate finite difference schemes for the Helmholtz equation, J. Comput. Acoust. 14 (2006) 339–351.
[24] G. Sutmann, Compact finite difference schemes of sixth order for the Helmholtz equation, J. Comput. Appl. Math. 203 (1) (2007) 15–31.
[25] A.N. Tikhonov, A.A. Samarskii, Equations of Mathematical Physics, Pergamon Press, Oxford, 1963.