

# A quasilinear complexity algorithm for the numerical simulation of scattering from a two-dimensional radially symmetric potential

James Bremer

*Department of Mathematics, University of California, Davis*

---

## Abstract

Standard solvers for the variable coefficient Helmholtz equation in two spatial dimensions have running times which grow at least quadratically with the wavenumber  $k$ . Here, we describe a solver which applies only when the scattering potential is radially symmetric but whose running time is  $\mathcal{O}(k \log(k))$  in typical cases. We also present the results of numerical experiments demonstrating the properties of our solver, the code for which is publicly available.

*Keywords:* Helmholtz equation, scattering theory, fast algorithms, numerical solution of partial differential equations

---

## 1. Introduction

The variable coefficient Helmholtz equation

$$\Delta u(x) + k^2 (1 + q(x)) u(x) = 0 \quad \text{for all } x \in \mathbb{R}^2 \quad (1)$$

can be used to model the scattering of waves in inhomogeneous media. Derivations of this equation from physical principles and discussions of its role in applications can be found in [22], [1] and [7], among other sources.

In many applications, the wavenumber  $k$  is real-valued and the scattering potential  $q$  is piecewise smooth with compact support contained inside of a disk  $\Omega$  of radius  $R$  centered at 0. Moreover,  $u$ , which is also referred to as the total field, is the sum of a known incident field  $u_i$  that is a solution of the constant coefficient Helmholtz equation

$$\Delta u_i(x) + k^2 u_i(x) = 0 \quad \text{for all } x \in \mathbb{R}^2 \quad (2)$$

and an unknown scattered field  $u_s$  which satisfies the Sommerfeld radiation condition

$$\lim_{r \rightarrow \infty} \sup_{0 \leq t \leq 2\pi} \sqrt{r} \left| \frac{\partial u_s}{\partial r}(r, t) - iku_s(r, t) \right| = 0. \quad (3)$$

In (3) and in what follows,  $r$  and  $t$  are the usual polar coordinates for the point  $x \in \mathbb{R}^2$  so that  $x = r \exp(it)$ . Together (1), (2) and (3) imply that the scattered field satisfies the boundary value problem

$$\begin{cases} \Delta u_s(x) + k^2 (1 + q(x)) u_s(x) = -k^2 q(x) u_i(x) & \text{for all } x \in \mathbb{R}^2 \\ \sup_{0 \leq t \leq 2\pi} \sqrt{r} \left| \frac{\partial u_s}{\partial r}(r, t) - iku_s(r, t) \right| = 0, \end{cases} \quad (4)$$

---

*Email address:* `bremer@math.ucdavis.edu` (James Bremer)

and it is well known that (4) suffices to uniquely determine  $u_s$  (see, for instance, Theorem 8.7 in Chapter 8 of [7]).

The scattered field  $u_s$  becomes increasingly oscillatory as  $k$  grows, and  $\mathcal{O}(k^2)$  unknowns are required to discretize it accurately using standard methods, such as collocation, orthogonal polynomial bases and finite element bases. This is true even in the case of the constant coefficient Helmholtz equation

$$\Delta u(x) + k^2 u(x) = 0 \quad \text{for all } x \in \mathbb{R}^2. \quad (5)$$

Indeed,

$$\cos(kx_1) + \sin(kx_2), \quad (6)$$

where  $x_1$  and  $x_2$  denote the usual rectangular coordinates in  $\mathbb{R}^2$ , is a typical solution of (5). According to the Nyquist sampling theorem (see, for instance, Chapter 3 of [21]),  $\mathcal{O}(k^2)$  discretization nodes are required to accurately represent (6).

Consequently, all numerical methods for (4) which rely on standard methods for representing its solutions have running times which grow at least quadratically with the wavenumber  $k$ . In fact, most existing solvers for (4) use  $n = \mathcal{O}(k^2)$  discretization nodes to represent the desired solution and have running times which grow superlinearly with  $n$ . See, for instance, [8], [7], [11], [2], and their references for examples of such methods and discussions of the asymptotic complexity of state-of-the-art solvers for the variable coefficient Helmholtz equation.

Numerical-asymptotic methods, such as [18, 16, 17, 18, 24, 5, 6, 20, 28], combine asymptotic models with conventional numerical schemes (finite element methods, discontinuous Galerkin methods, integral equation methods, etc.) to represent solutions more efficiently. Typically, these schemes are able to achieve high-accuracy at a wide range of wavenumbers. However, although they generally require far fewer degrees of freedom per wavelength than conventional numerical methods, their running times still typically grow quadratically with  $k$ , at least until such large frequencies are reached that the underlying asymptotic models become highly accurate.

Here, we describe a high accuracy method for solving the variable coefficient Helmholtz equation whose running time is  $\mathcal{O}(k \log(k))$  in typical cases but which only applies when the scattering potential  $q$  is radially symmetric. Our approach exploits this symmetry to construct an expansion of the total field through separation of variables. More explicitly, we represent  $u$  in the interior of the disk  $\Omega$  via a sum of the form

$$u(r, t) = \sum_{n=-m}^m a_n \psi_{|n|}(r) \exp(int), \quad (7)$$

where, for each nonnegative integer  $n$ ,  $\psi_n$  satisfies a second order differential equation whose coefficients depend on the wavenumber  $k$  and the index  $n$  and which we refer to as the “perturbed Bessel equation.” The value of  $m$  in (7) is equal to the number of Fourier modes needed to accurately represent the restriction of the incident wave  $u_i$  to the boundary of the scatterer  $\partial\Omega$ . Since  $u_i$  satisfies the constant coefficient Helmholtz equation at wavenumber  $k$ ,  $m$  is  $\mathcal{O}(k)$  in typical applications. Moreover, we solve the perturbed Bessel equation using a method whose worst case running time appears to be  $\mathcal{O}(\log(k))$ .

We say “appears to be” because our evidence for this claim is, in part, experimental. To construct the  $\psi_n$ , we use an approach based on the method of [3] for the numerical solution of second order differential equations of the form

$$y''(t) + k^2 \eta(t) y(t) = 0 \quad \text{for all } a < t < b. \quad (8)$$

When  $\eta$  is positive, the solutions of (8) are oscillatory and  $\mathcal{O}(k)$  unknowns are required to discretize them using standard approaches (e.g., by sampling their values). The solver of [3] instead represents them using a nonoscillatory phase function which can be calculated and stored efficiently. Indeed, according to the estimates of [4], the method of [3] runs in time independent of  $k$  under mild assumptions on  $\eta$  including that it is smooth and *strictly positive* on the interval  $[a, b]$ . The equation which defines  $\psi_n$  can have turning points, however, and in the neighborhood of such a point, the estimates of [4] do not apply. Here, we present the results of numerical experiments which strongly indicate that at worst our method requires  $\mathcal{O}(\log(k))$  to construct each  $\psi_n$ . Assuming this is correct,  $\mathcal{O}(k \log(k))$  time is required to construct all of the functions  $\psi_n$  appearing in (7). A further  $\mathcal{O}(k \log(k))$  operations are needed to calculate the coefficients in this expansion, so that the time required to solve (4) is  $\mathcal{O}(k \log(k))$ .

It is well known that the condition number of evaluation of oscillatory functions increases with the frequency of their oscillations and that this generally limits the accuracy with which they can be evaluated numerically (see, for instance, Chapter 1 of [13]). As a consequence, numerical schemes for solving (4), including the algorithm of this paper, lose accuracy with increasing  $k$ . Section 5 describes numerical experiments which were conducted to assess the speed and accuracy of our scheme.

Separation of variables is hardly a new idea; but, it appears to have been rarely used as a numerical tool for solving (4). This is most likely because it offers no advantage in asymptotic running time when standard numerical methods are used to solve the differential equations which arise. Moreover, when the parameter  $n$  is large,  $\psi_n$  behaves like a rapidly increasing exponential function on some part of its domain, and the accurate numerical solution of the corresponding differential equation becomes difficult. One of the few robust numerical algorithms for solving (4) via separation of variables is described in [15]. It proceeds by reformulating the differential equations that arise as integral equations which are then inverted using a “fast direct solver.” This results in an extremely accurate and robust method for solving (4); however, the running time of this method still grows quadratically with  $k$ . Our scheme also bears some similarities to that of [11]. There, spectral methods are used to construct expansions of the total field in the interior of  $\Omega$  and of the scattered field in the exterior of  $\Omega$ , and the coefficients in these expansion are found using a mechanism very similar to that used by our algorithm to compute the coefficients in the expansion (7). However, the scheme of [11] is not limited to radially symmetric potentials and its running time grows faster than  $\mathcal{O}(k^2)$ .

The remainder of this document is structured as follows. Section 2 discusses the solution of the boundary value problem (4) via the method separation of variables. Our algorithm for the numerical solution of the boundary value problem (4) is discussed in Sections 3 and 4. In Section 5, we describe several numerical experiments which demonstrate the properties of our solver. We close in Section 6 with a few brief remarks regarding this work and a discussion of possible directions for future research.

## 2. Determination of the Scattered Field Through Separation of Variables

In this section, we describe how the method of separation of variables can be used to solve the boundary value problem (4).

The total field  $u$  satisfies

$$\Delta u(r, t) + k^2(1 + q(r))u(r, t) = 0 \quad (9)$$

in  $\Omega$  and is nonsingular at 0. Separating variables in (9) gives us the representation formula

$$u(r, t) = \sum_{n=-\infty}^{\infty} a_n \psi_{|n|}(r) \exp(int), \quad (10)$$

where, for each nonnegative integer  $n$ ,  $\psi_n$  is a nonzero solution of

$$r^2 y''(r) + ry'(r) + (k^2(1 + q(r))r^2 - n^2) y(r) = 0, \quad 0 < r < R, \quad (11)$$

which is regular at 0. When  $q = 0$ , (11) becomes Bessel's differential equation and we refer to it as the perturbed Bessel equation. Since this equation has a regular singular point at 0 and the roots of the indicial equation are  $n$  and  $-n$ , there exists a basis  $\{u_n, v_n\}$  in the space of its solutions such that

$$u_n(r) = \mathcal{O}(r^n) \quad \text{as } r \rightarrow 0 \quad (12)$$

and

$$v_n(r) = \begin{cases} \mathcal{O}(\log(r)) & \text{as } r \rightarrow 0 \quad \text{if } n = 0 \\ \mathcal{O}(r^{-n}) & \text{as } r \rightarrow 0 \quad \text{if } n > 0 \end{cases} \quad (13)$$

(see, for instance, Chapter 5 of [14]). Since  $\psi_n$  is nonsingular at 0, it must be a multiple of the solution  $u_n$ . Because we do not impose a second boundary condition on  $\psi_n$ , it is only determined up to a (nonzero) multiplicative constant. It is perhaps tempting to impose a second condition on  $\psi_n$  in the hopes of uniquely determining it. However, doing so generally leads to a boundary value problem which is not solvable for all possible  $k$ . Moreover, this ambiguity has no impact on the form of the expansion (10) used to represent the total field in the interior of  $\Omega$ , although the particular value of the coefficient  $a_n$  depends on the choice of  $\psi_n$ .

Because  $q$  is supported inside of the disk  $\Omega$ , the scattered field  $u_s$  solves

$$\begin{cases} \Delta u_s(r, t) + k^2 u_s(r, t) = 0 & \text{in } \Omega^c \\ \lim_{r \rightarrow \infty} \sup_{0 \leq t \leq 2\pi} \sqrt{r} \left| \frac{\partial u_s}{\partial r}(r, t) - iku_s(r, t) \right| = 0. \end{cases} \quad (14)$$

Separating variables in (14) shows that  $u_s$  can be represented in the exterior of  $\Omega$  via a sum of the form

$$u_s(r, t) = \sum_{n=-\infty}^{\infty} b_n \sigma_{|n|}(kr) \exp(int), \quad (15)$$

where, for each nonnegative  $n$ ,  $\sigma_n$  is a solution of Bessel's differential

$$r^2 y''(r) + ry'(r) + (k^2 r^2 - n^2) y(r) = 0, \quad R < r < \infty, \quad (16)$$

consistent with the Sommerfeld radiation condition. It follows from the integral representation formula

$$H_n(z) = \exp(iz) \left( \frac{-2^{n+1}i}{\sqrt{\pi} \Gamma(n + \frac{1}{2})} z^n \int_0^\infty (x^2 - ix)^{n-\frac{1}{2}} \exp(-2xz) dx \right) \quad (17)$$

for the Hankel function of the first kind of order  $n$  (which can be found as Formula 3.388(4) in [12]) and standard results regarding the decay properties of Laplace transforms (see, for instance, [27]) that  $H_n(kr)$  is a solution of (16) which has the appropriate behavior at infinity. In particular,  $u_s$  admits the representation

$$u_s(r, t) = \sum_{n=-\infty}^{\infty} b_n H_n(kr) \exp(int) \quad (18)$$

in the exterior of  $\Omega$ .

Standard elliptic regularity results (see, for instance, Section 8.4 of [10]) imply that the total field  $u$  and its derivative  $\partial u / \partial r$  with respect to the radial variable are continuous across the boundary  $\partial\Omega$  of the disk  $\Omega$ . This suffices to determine the coefficients  $a_n$  in (10) and  $b_n$  in (18). To see this, we first let

$$u_i(R, t) = \sum_{n=-\infty}^{\infty} c_n \exp(int) \quad (19)$$

and

$$\frac{\partial u_i}{\partial r}(R, t) = \sum_{n=-\infty}^{\infty} d_n \exp(int) \quad (20)$$

be the Fourier expansions of the restrictions of the incident field  $u_i$  and its radial derivative to  $\partial\Omega$ . Since  $u - u_s = u_i$ , the continuity of  $u$  and its radial derivative imply that

$$\sum_{n=-\infty}^{\infty} a_n \psi_{|n|}(R) \exp(int) - \sum_{n=-\infty}^{\infty} b_n H_n(kR) \exp(int) = \sum_{n=-\infty}^{\infty} c_n \exp(int) \quad (21)$$

and

$$\sum_{n=-\infty}^{\infty} a_n \psi'_{|n|}(R) \exp(int) - \sum_{n=-\infty}^{\infty} b_n k H'_n(kR) \exp(int) = \sum_{n=-\infty}^{\infty} d_n \exp(int) \quad (22)$$

for all  $0 \leq t \leq 2\pi$ . Owing to the orthogonality of the set  $\{\exp(int)\}$ , (21) and (22) hold if and only if for each integer  $n$ ,  $a_n$  and  $b_n$  satisfy the linear system of equations

$$\begin{cases} a_n \psi_{|n|}(R) - b_n H_n(kR) = c_n \\ a_n \psi'_{|n|}(R) - b_n k H'_n(kR) = d_n. \end{cases} \quad (23)$$

Since the scattered field is uniquely determined by (4), each of the systems (23) must be uniquely solvable. In particular, for each integer  $n$ , the determinant

$$H_n(kR) \psi'_{|n|}(R) - k \psi_{|n|}(R) H'_n(kR) \quad (24)$$

of the coefficient matrix in (23) is necessarily nonzero and the coefficients  $a_n$  and  $b_n$  are given by the formulas

$$a_n = \frac{-k H'_n(kR) c_n + H_n(kR) d_n}{H_n(kR) \psi'_{|n|}(R) - k \psi_{|n|}(R) H'_n(kR)} \quad (25)$$

and

$$b_n = \frac{-\psi'_{|n|}(R) c_n + \psi_{|n|}(R) d_n}{H_n(kR) \psi'_{|n|}(R) - k \psi_{|n|}(R) H'_n(kR)}. \quad (26)$$

This procedure determines the scattered field in the sense that once the coefficients  $a_n$  and  $b_n$  are determined through (25) and (26), the scattered field can be evaluated at any point outside of  $\Omega$  via (18) and it can be evaluated at any point inside of  $\Omega$  or on  $\partial\Omega$  by first evaluating the total field via (10) and then subtracting the value of the (known) incident field  $u_i$  from the result.

### 3. A Numerical Algorithm for the Determination of the Scattered Field

This is the first of two sections detailing our method for the solution of the boundary value problem (4). In it, we describe all elements of the algorithm except for the procedure for solving Equation (11). The following section addresses that topic.

Our algorithm operates in two phases: a precomputation phase in which the perturbed Bessel

equation is repeatedly solved in order to construct the functions  $\psi_n$  used in the representation of the total field, and a solution phase in which the coefficients in the expansions of the total and scattered fields are calculated. We refer to the first procedure as the “precomputation phase” because in many applications it is necessary to solve (4) for multiple incident fields  $u_i$  while the wavenumber  $k$  and scattering potential  $q$  are fixed. In this event, the precomputation phase is only executed once and the solution phase is executed once for each incident field.

The precomputation phase takes as input the wavenumber  $k$  for the problem, a subroutine for evaluating the scattering potential  $q$  at any specified point, a list

$$\chi_1 < \chi_2 < \dots < \chi_s \quad (27)$$

of all of the points on the interval  $(0, R)$  at which the scattering potential  $q(r)$  is nonsmooth (recall that  $q$  is assumed to be piecewise smooth), and a positive integer  $m$ . It consists of calculating functions

$$\psi_0, \psi_1, \dots, \psi_m \quad (28)$$

such that for each integer  $0 \leq n \leq m$ ,  $\psi_n$  is a solution of the perturbed Bessel equation (11) which is regular at 0. We detail our algorithm for solving the perturbed Bessel equation in Section 4, which follows this one. Based on strong experimental evidence, we believe that the running time of the precomputation phase is  $\mathcal{O}(m \log(m))$ .

The solution phase takes as input a routine for evaluating the incident field  $u_i$  and its derivative  $\partial u_i / \partial r$  with respect to the radial variable  $r$ . It proceeds by first forming the approximations

$$u_i(R, t) \approx \sum_{n=-m}^m c_n \exp(int) \quad (29)$$

and

$$\frac{\partial u_i}{\partial r}(R, t) \approx \sum_{n=-m}^m d_n \exp(int) \quad (30)$$

of the restrictions of  $u_i$  and its normal derivative to the boundary  $\partial\Omega$  in the usual way — that is, using the fast Fourier transform. The integer  $m$  must be sufficiently large for the approximations (29) and (30) to be highly accurate. In the next step of the solution phase, the coefficients in the truncated expansion

$$u(r, t) = \sum_{n=-m}^m a_n \psi_{|n|}(r) \exp(int) \quad (31)$$

which represents the total wave in the interior of  $\Omega$  are computed using Formula (25). Finally, the coefficients in the truncated expansion

$$u_s(r, t) = \sum_{n=-m}^m b_n H_n(kr) \exp(int) \quad (32)$$

used to represent the scattered field in the exterior of  $\Omega$  are computed using (26). The fast Fourier transforms take  $\mathcal{O}(m \log(m))$  operations, and they dominate the cost of this phase of the algorithm. Only  $\mathcal{O}(m)$  operations are required to construct the coefficients in the expansions (31) and (32).

The coefficients in the expansions (31) and (32) are the principal outputs of the algorithm of this paper. Once they have been determined, the scattered field can be evaluated at any point in the exterior of  $\Omega$  in  $\mathcal{O}(m)$  operation by evaluating the sum (32), and it can be evaluated at any point in the interior of  $\Omega$  in  $\mathcal{O}(m)$  operations by evaluating (31) and subtracting the value of the incident field  $u_i$ .

Since  $u_i$  is a solution of the constant coefficient Helmholtz equation at wavenumber  $k$ , it is expected that  $m$  will be on the order of  $k$  so that the running time of our algorithm is  $\mathcal{O}(k \log(k))$ . We found  $m = \pi/2Rk$  to be sufficient in all of the numerical experiments discussed in this paper. If  $m$  is not known *a priori* it can be determined through an adaptive procedure. For instance, starting from an initial guess,  $m$  could be gradually increased until the coefficients in the expansions (29) and (30) decay sufficiently fast.

#### 4. Numerical Solution of the Perturbed Bessel Equation

This is the second of two sections detailing our method for the solution of the boundary value problem (4). In it, we discuss the method used to solve Equation (11) numerically.

It can be easily seen that if  $\psi_n$  solves (11), then  $\varphi_n(r) = \sqrt{r} \psi_n(r)$  is a solution of

$$y''(r) + Q(r)y(r) = 0, \quad 0 < r < R, \quad (33)$$

where

$$Q(r) = \lambda^2 (1 + q(r)) + \frac{\frac{1}{4} - n^2}{r^2}. \quad (34)$$

We refer to (33) as the normal form of the perturbed Bessel equation and we find it more convenient to work with than (11). Among other things, a great deal of information about the behavior of the solutions of (33) can be easily discerned from the coefficient (34). Indeed, according to standard asymptotic results (see, for instance, [25] or [9]), in intervals on which  $Q$  is positive the solutions of (33) behave roughly as oscillatory exponential functions, while they resemble increasing or decreasing exponential functions in intervals on which  $Q$  is negative. The zeros of  $Q$  which separate these regions are known as turning points for (33). We do not, in fact, produce a solution of the perturbed Bessel equation over the entire interval  $[0, R]$  since  $Q$  is singular at 0 and, as a consequence, all but one of the solutions of (33) are also singular at 0. Instead, we produce a solution over  $[10^{-15}, R]$ , which generally suffices for the purposes of numerical computation.

The first step of our algorithm for solving the perturbed Bessel equation consists of forming a partition

$$10^{-15} = \xi_1 < \xi_2 < \dots < \xi_t = R \quad (35)$$

of the solution interval  $[10^{-15}, R]$  such that  $Q$  is smooth and does not change sign on each interval  $(\xi_j, \xi_{j+1})$ . We do so by finding the set of all zeros of  $Q$  on  $[0, R]$  and merging it with the list (27) of the singularities of  $Q$  provided by the user as well as the additional points  $10^{-15}$  and  $R$ .

Next, for each  $j = 1, \dots, t$ , our solver forms a basis

$$\mathcal{B}_j = \{u_j, v_j\} \quad (36)$$

in space of the restrictions of solutions of (33) to  $(\xi_j, \xi_{j+1})$ . The mechanisms used to construct and represent these basis functions differ depending on whether  $Q$  is positive in the interval (the oscillatory regime) or negative in the interval (the nonoscillatory regime). We discuss the details in each case below.

Finally, for each  $j = 1, \dots, t$ , it calculates coefficients  $\gamma_j$  and  $\zeta_j$  such that the restriction of the desired solution  $\varphi_n$  of (33) to the interval  $(\xi_j, \xi_{j+1})$  is

$$\gamma_j u_j(r) + \eta_j v_j(r). \quad (37)$$

In the case of the first interval  $(\xi_1, \xi_2)$ , the values of  $\varphi_n$  and its derivative at the point  $\xi_1 = 10^{-15}$

are first estimated via the asymptotic approximation

$$\varphi_n(r) \sim \sqrt{r} J_n \left( \sqrt{q(0) + k^2 r} \right) \quad \text{as } r \rightarrow 0, \quad (38)$$

which can be easily derived using standard methods (see, for instance, [25] or [9]). Then, the linear system of equations

$$\begin{cases} \gamma_1 u_1(\xi_1) + \zeta_{j-1} u_1(\xi_1) = \varphi_n(\xi_1) \\ \gamma_1 u'_1(\xi_1) + \zeta_{j-1} u'_1(\xi_1) = \varphi'_n(\xi_1) \end{cases} \quad (39)$$

is solved for  $\gamma_1$  and  $\zeta_1$ . For each  $j > 1$ , the coefficients  $\gamma_j$  and  $\zeta_j$  are determined by enforcing the continuity of  $\varphi_n$  and its derivative at the point  $\xi_j$ . More explicitly, we solve the system of linear equations

$$\begin{cases} \gamma_{j-1} u_{j-1}(\xi_j) + \eta_{j-1} u_{j-1}(\xi_j) = \gamma_j u_j(\xi_j) + \eta_j u_j(\xi_j) \\ \gamma_{j-1} u'_{j-1}(\xi_j) + \eta_{j-1} u'_{j-1}(\xi_j) = \gamma_j u'_j(\xi_j) + \eta_j u'_j(\xi_j) \end{cases} \quad (40)$$

for  $\gamma_j$  and  $\eta_j$ .

#### 4.1. The oscillatory regime

For intervals  $(\xi_j, \xi_{j+1})$  on which  $Q$  is positive, we use a basis generated by a nonoscillatory phase function. A function  $\alpha$  is a phase function for the differential equation

$$y''(t) + \eta(t)y(t) = 0 \quad \text{for all } a < t < b \quad (41)$$

provided  $\alpha'(r) > 0$  for all  $a < r < b$  and

$$\frac{\sin(\alpha(r))}{\sqrt{\alpha'(r)}} \quad \text{and} \quad \frac{\cos(\alpha(r))}{\sqrt{\alpha'(r)}} \quad (42)$$

is a basis in its space of solutions. An extensive discussion of phase functions for second order differential equations can be found in [23]. Among other things, it is shown there that  $\alpha$  is a phase function for (41) if and only if its derivative satisfies the nonlinear second order differential equation

$$(\alpha'(r))^2 = \eta(r) - \frac{1}{2} \frac{\alpha'''(r)}{\alpha'(r)} + \frac{3}{4} \left( \frac{\alpha''(r)}{\alpha'(r)} \right)^2, \quad a < r < b, \quad (43)$$

which we refer to as Kummer's equation after E. E. Kummer who studied it in [19]. We note that (43) only determines  $\alpha$  up to a constant. For our purposes, the constant is largely irrelevant (our only requirement is that it not be too large in magnitude), and we always determine it by taking  $\alpha$  to be zero at left-hand endpoint of the interval on which it is defined.

In [4], it is shown that, under mild assumptions on the coefficient  $\eta$  (including the condition that it be positive on the interval  $[a, b]$ ), there exists a phase function for (41) which is roughly as oscillatory as the coefficient  $\eta$ . Moreover, in [3], a fast algorithm for the numerical calculation of this nonoscillatory phase functions is presented. It operates by first introducing a "windowed version"  $\tilde{\eta}$  of  $\eta$  such that

$$\tilde{\eta}(r) = \begin{cases} \lambda^2 & \text{for all } a < r < \frac{3a+b}{4} \\ \eta(r) & \text{for all } \frac{a+3b}{4} < r < b \end{cases} \quad (44)$$

with  $\lambda$  a constant chosen to be roughly on the order of  $\sqrt{\eta(a)}$ . Since  $\tilde{\eta} = \lambda^2$  near  $a$ , the nonoscillatory phase function for the equation

$$y''(r) + \tilde{\eta}(r)y(r) = 0 \quad (45)$$

is equal to  $\lambda r$  near  $a$ . By solving the initial value problem

$$\begin{cases} (\tilde{\alpha}'(r))^2 = \tilde{\eta}(r) - \frac{1}{2} \frac{\tilde{\alpha}'''(r)}{\tilde{\alpha}'(r)} + \frac{3}{4} \left( \frac{\tilde{\alpha}''(r)}{\tilde{\alpha}'(r)} \right)^2 \\ \tilde{\alpha}'(r) = \lambda \text{ and } \tilde{\alpha}''(r) = 0, \end{cases} \quad (46)$$

the values of  $\tilde{\alpha}'(b)$  and  $\tilde{\alpha}''(b)$  are determined. Since  $\tilde{\eta}$  is equal to  $\eta$  near  $b$ , these values closely approximate  $\alpha'(b)$  and  $\alpha''(b)$ . The function  $\alpha$  is then determined over the interval  $(a, b)$  by solving the terminal value problem

$$\begin{cases} (\alpha'(r))^2 = \eta(r) - \frac{1}{2} \frac{\alpha'''(r)}{\alpha'(r)} + \frac{3}{4} \left( \frac{\alpha''(r)}{\alpha'(r)} \right)^2 \\ \alpha'(b) = \tilde{\alpha}'(b) \\ \alpha''(b) = \tilde{\alpha}''(b). \end{cases} \quad (47)$$

To solve (46) and (47), we use a spectral method which represents the phase function  $\alpha$  and its derivatives using piecewise Chebyshev expansions over a collection of subintervals of  $(a, b)$ . The subintervals are chosen adaptively. Other mechanisms for the solution of these differential equations could be used, so long as they are well-suited for stiff problems. When the coefficient  $\eta$  is strictly positive on  $[a, b]$ , the nonoscillatory phase function for (41) produced by this algorithm can be constructed and evaluated in time independent of the magnitude of  $\eta$  (which is a measure of the frequency of oscillation of the solutions of (41)).

Since (33) can have turning points at the endpoints  $\xi_j$  and  $\xi_{j+1}$ , the estimates of [4] do not apply. However, we have found experimentally (see Sections 5.1 and 5.2) that the worst case running time of the algorithm of [3] is  $\mathcal{O}(\log(k))$  in this case.

#### 4.2. The nonoscillatory regime

For intervals  $(\xi_j, \xi_{j+1})$  on which  $Q$  is negative, we use a basis  $\{u_j, v_j\}$  of solutions of (33) such that  $u_j$  resembles an increasing exponential function and  $v_j$  resembles a decreasing exponential function. Because the cost of representing these functions using standard methods, such as through expansions in orthogonal polynomials or via sampling their values, increases rapidly with  $n$  and  $k$ , we instead construct their logarithms and use these to evaluate  $u_j$  and  $v_j$  as needed.

If  $y(r) = \exp(\sigma(t))$  satisfies the second order differential equation (33), then it can be easily verified that  $\sigma'$  satisfies the Riccati equation

$$\sigma''(r) + (\sigma'(r))^2 + Q(r) = 0 \quad (48)$$

(see, for instance, Chapter 4 of [14] for a discussion of the Riccati equation). To construct the logarithm of  $u_j$ , we first calculate a solution  $\sigma'$  of (48) which satisfies the initial condition  $\sigma'(\xi_j) = 0$ . Again, we use an adaptive spectral solver whose output is a piecewise Chebyshev expansion representing the solution. We then use spectral integration to form the antiderivative  $\sigma$  of  $\sigma'$  such that  $\sigma(\xi_{j+1}) = 0$ . Since the dominant solution of (48) when solving in the forward direction is increasing, the function  $u_j$  constructed in this fashion resembles an increasing exponential function.

To construct the logarithm of  $v_j$ , we solve (48), imposing the terminal condition  $\sigma'(\xi_{j+1}) = 0$ . Next we use spectral integration to form the antiderivative  $\sigma$  of  $\sigma'$  such that  $\sigma(\xi_j) = 0$ . Since the dominant solution of (48) in the backward direction is decreasing,  $v_j$  resembles a decreasing exponential function.

Based on the extensive numerical experiments of Section 5, we believe that the worst case running time of this procedure for constructing  $u_j$  and  $v_j$  is  $\mathcal{O}(\log(k))$ .

## 5. Numerical Experiments

In this section, we describe numerical experiments which were conducted to evaluate the performance of the algorithm of this paper. Our code was written in Fortran with OpenMP extensions and was compiled with the GNU Fortran compiler version 7.4.0. All calculations were performed on a workstation computer equipped with 28 Intel Xeon E5-2697 processor cores running at 2.6 GHz. We used P. Swarztrauber's FFTPACK library [26] to apply the fast Fourier transform. We used a code provided by V. Rokhlin to evaluate the Hankel functions. Our implementation of the algorithm of this paper and our code for conducting the numerical experiments described here are available on GitHub at the following address:

<https://github.com/JamesCBremerJr/HelmRad>

We used the following procedure to measure the accuracy of solutions produced by our solver for (4). We first executed it using extended precision (Fortran REAL\*16) arithmetic, which gives about 33 decimal digits of accuracy. When possible, we directly verified that the obtained scattered field satisfies the partial differential equation

$$\Delta u_s(x) + k^2 (1 + q(x)) u_s(x) + k^2 q(x) u_i(x) = 0 \quad (49)$$

to at least 15 decimal digits of accuracy using a Chebyshev spectral method. We then executed our algorithm a second time using double precision (Fortran REAL\*8) arithmetic and measured the error in the obtained solution by comparison with the reference solution produced using extended precision arithmetic. The condition number of the spectral discretization of (49) increases rapidly with the number of discretization nodes needed and hence with  $k$ , which is why extended precision arithmetic was necessary to verify the reference solutions. Even so, ill-conditioning limited the use of this technique to problems in which  $k$  was less than or equal to 1024. When reporting errors, we use parentheses to indicate experiments in which we could not verify the accuracy of our extended precision solution through a spectral method. To summarize, we first used our solver to generate solutions using extended precision arithmetic. For all  $k \leq 1024$ , these were verified to be accurate to at least 15 decimal digits of precision using a spectral method. We then generated solutions in double precision arithmetic using our solver and measured the error of these solutions by comparison with those generated in extended precision.

In each experiment, we used an excessive number of Fourier modes to represent the restriction of the incoming field to  $\partial\Omega$  in order to avoid any loss of accuracy from the truncation of its Fourier expansion. The number of Fourier modes used is reported in each experiment. In many cases, we expect that the Fourier expansion of the incoming wave will be constructed using an adaptive algorithm (and an implementation of such an algorithm is included with our publicly available code).

The code for the precomputation phase of our algorithm is multithreaded in order to take advantage of the embarrassingly parallel nature of the calculation (each  $\psi_n$  can be constructed entirely independent of the others). The FFTPACK library, on the other hand, is single-threaded and although the solution phase of our algorithm could no doubt be accelerated by switching to a multithreaded FFT library, we opted not to do so because the source code for the FFTPACK library was readily available and easy to modify to use extended precision arithmetic.

In the course of conducting these experiments, we found that there is a large jump in the cost of applying the FFT using the FFTPACK library when the dimension of the transform is increased from  $2^{15}$  to  $2^{16}$ , and this is reflected in the timings for the solution phase of our algorithm. Since algorithms for applying the fast Fourier transform are not our principal concern here, we did not extensively investigate this issue. However, we suspect that it is a cache effect.

### 5.1. The numerical solution of the perturbed Bessel equation, part I

We now describe a set of experiments conducted to measure performance of our method for solving the normal form of the perturbed Bessel equation (33). In each of them,  $q$  was taken to be

$$q(r) = r^2 - 1, \quad (50)$$

and the equation was solved over the interval  $[0, 2]$ . The functions

$$\left\{ J_{\frac{n}{2}} \left( \frac{\lambda}{2} r^2 \right) \sqrt{r}, Y_{\frac{n}{2}} \left( \frac{\lambda}{2} r^2 \right) \sqrt{r} \right\} \quad (51)$$

form a basis in the space of solutions of (33) in this case, so any solution which is regular at the origin is necessarily multiple of  $J_{\frac{n}{2}} \left( \frac{\lambda}{2} r^2 \right) \sqrt{r}$ . This made assessing the accuracy of obtained solutions possible — we did so by measuring their absolute error at 100 points on the interval  $[0, 2]$ .

In the first of these experiments, the results of which are shown in the first row of Figure 1, we held  $k$  fixed at  $2^{17} = 131,072$  and increased  $n$  from 0 to  $2^{17}$ . We report the time taken by our solver and the largest observed absolute error as functions of  $n$ .

In the second, the results of which are shown in the second row of Figure 1, we let  $n = 0$  and increased  $k$  from 0 to  $2^{17}$ . We once again report the time taken by our solver and the largest observed absolute error as functions of  $k$ .

In the third and fourth experiments, the results of which are reported in the third and fourth rows of Figure 1, respectively, we fixed  $n$  to be a constant multiple of  $k$  and increased  $k$  from 0 to  $2^{17}$ . We again measured the running time and largest absolute error. In the third experiment, we set  $n = k/2$  and in the fourth we set  $n = k$ .

We observe that in the case in which the solutions are purely oscillatory (i.e., when  $n = 0$ ), the running time of the procedure is essentially independent of  $k$ . This is consistent with the estimates of [4]. When  $k$  is fixed and  $n$  is increased, there is modest growth in the runtime of the procedure. In the cases in which the equation has a turning point and  $n$  is increased in proportion to  $k$ , the running time of the procedure appears to grow logarithmically with  $k$ . This is consistent with our conjecture that the running time of this procedure grows logarithmically with  $k$  in the worst case.

### 5.2. The numerical solution of the perturbed Bessel equation, part II

In most applications, it is necessary to calculate the solutions of the perturbed Bessel equation for a fixed  $k$  and a range of value of  $n$ . For instance, it is often necessary to construct the set

$$S_k = \{\varphi_n : n = 0, 1, \dots, k\}, \quad (52)$$

where, for each  $n$ ,  $\varphi_n$  is a solution of (33). We conducted several experiments to measure the time required by our solver to construct the set  $S_k$  as a function of  $k$ . In each of them,  $R$  was taken to be 2 and, for each  $k = 2^8, 2^9, \dots, 2^{16}, 2^{17}$ , we measured the time required to construct  $S_k$ . In the first experiment, the results of which are shown in Figure 2,  $q(r)$  was taken to be  $r^2 - 1$ . In the second experiment, the results of which appear in Figure 3,  $q(r) = 14r^2 \exp(-5r^2)$ . In the third experiment,  $q(r) = 3\chi_{[1,2]}(r)$ , where  $\chi_{[1,2]}(r)$  denotes the characteristic function of the interval  $[1, 2]$ . The results of this third experiment appear in Figure 4. In the case of the first and third experiment, the solution of (33) is known, and we were able to measure the absolute errors in the obtained solutions.

We observe first that the time required to construct  $S_k$  appears to be  $\mathcal{O}(k \log(k))$ , which is consistent with our conjecture. We also observe that the accuracy of the obtained solutions deteriorates as  $k$  increases. This is expected as the condition number of both (4) and (33) increase with  $k$ .

### 5.3. Scattering of a plane wave from a Gaussian potential

In this experiment, the incident field was the plane wave

$$u_i(r, t) = \exp\left(ikr \cos\left(t - \frac{\pi}{4}\right)\right) \quad (53)$$

and the scattering potential was defined by

$$q(r) = \exp(-5r^2). \quad (54)$$

We solved (4) for each  $k = 2^4, 2^5, \dots, 2^{17}$  and measured the time required by each phase of our algorithm. We also measured the absolute error in the obtained solution in the fashion described at the beginning of this section. The results are shown in Figure 5 and Table 1. Figure 5 also displays images of the incoming field, scattered field and total field when  $k = 16$ , as well as a plot of the function  $q(r)$ . A plot of the scattering potential as a function of the two spatial variables  $x$  and  $y$  appears in Figure 6.

We observe that the running time of our solver appears to grow in line with our conjecture — that is, as  $\mathcal{O}(k \log(k))$  — and that while the accuracy of our solver deteriorates with  $k$ , at least 6-7 digit accuracy is achieved all cases, even for a problem of more than 100,000 wavelengths in size.

### 5.4. Scattering of a circular wave from a potential resembling a volcano

In this experiment, the incident field was the circular wave

$$u_i(z) = H_0(k|z - z_0|), \quad (55)$$

where  $z_0 = 6i$ , and the scattering potential was defined by

$$q(r) = 14r^2 \exp(-5r^2). \quad (56)$$

Again, we solved (4) for each  $k = 2^4, 2^5, \dots, 2^{17}$  and measured the time required by each phase of our algorithm. We also measured the absolute error in the obtained solution in the fashion described at the beginning of this section. The results are shown in Figure 7 and Table 2. Figure 7 also displays images of the incoming field, scattered field and total field when  $k = 16$ , as well as a plot of the function  $q(r)$ . A plot of the scattering potential as a function of the two spatial variables  $x$  and  $y$  appears in Figure 8.

Again, we see that the running time of our solver appears to grow as  $\mathcal{O}(k \log(k))$ , and that, in the worst case for a problem of more than one hundred thousand wavelengths in size, roughly 6-7 digit accuracy is obtained.

### 5.5. Scattering of a plane wave from a discontinuous potential

In this experiment, the incident field was the plane wave

$$u_i(r, t) = \exp\left(ikr \cos\left(t - \frac{\pi}{4}\right)\right) \quad (57)$$

and the scattering potential was defined by

$$q(r) = \begin{cases} 1 & 0 < r < 1 \\ 2 & 2 < r < 3 \\ 0 & \text{otherwise.} \end{cases} \quad (58)$$

We once again solved (4) for each  $k = 2^4, 2^5, \dots, 2^{17}$  and measured the time required by each phase of our algorithm. We also measured the absolute error in the obtained solution in the fashion described at the beginning of this section. The results are shown in Figure 9 and Table 3. Figure 9 also displays images of the incoming field, scattered field and total field when  $k = 16$ , as well as a

plot of the function  $q(r)$ . A plot of the scattering potential as a function of the two spatial variables  $x$  and  $y$  appears in Figure 10.

Again, we see that the running time of our solver appears to grow as  $\mathcal{O}(k \log(k))$ , and that, in the worst case for a problem of more than one hundred thousand wavelengths in size, roughly 6-7 digit accuracy is obtained.

### 5.6. Scattering of an incident wave generated by a linear array from a discontinuous potential

In this experiment, the incident field was generated by a linear array of charges placed along the line segment connecting the points  $(-5, 10)$  and  $(5, 10)$ . More explicitly,  $u_i$  was given by the formula

$$u_i(z) = \sum_{j=1}^{10} H_0 \left( k \left| z - \left( -5 + 10 \frac{j-1}{9} + 10i \right) \right| \right). \quad (59)$$

The scattering potential was defined by

$$q(r) = \begin{cases} 1 & 0 < r < 1 \\ 2 - 4 \left( r^2 - \frac{5}{2} \right)^2 & 2 < r < 3 \\ 0 & \text{otherwise.} \end{cases} \quad (60)$$

We once again solved (4) for each  $k = 2^4, 2^5, \dots, 2^{17}$  and measured the time required by each phase of our algorithm. We also measured the absolute error in the obtained solution in the fashion described at the beginning of this section. The results are shown in Figure 11 and Table 4. Figure 11 also displays images of the incoming field, scattered field and total field when  $k = 16$ , as well as a plot of the function  $q(r)$ . A plot of the scattering potential as a function of the two spatial variables  $x$  and  $y$  appears in Figure 12.

Again, we see that the running time of our solver appears to grow as  $\mathcal{O}(k \log(k))$ , and that, in the worst case for a problem of more than one hundred thousand wavelengths in size, roughly 6-7 digit accuracy is obtained.

## 6. Conclusions and Future Work

We have developed a fast method for the numerical solution of the two-dimensional variable coefficient Helmholtz equation in the radially symmetric case. It is based on separation of variables and exploits the fact that a large class of second order differential equations admit nonoscillatory phase functions. Using our scheme, problems of up to several hundred thousands wavelengths in size can be solved with relatively high accuracy (at least 6-7 digits) in a matter of minutes on a typical workstation computer.

Even so, there are many inefficiencies in our approach. For instance, a reduction in the asymptotic running time of our method for solving the perturbed Bessel differential equation could be achieved through the use of standard asymptotic methods for equations of the form (8). We could represent the functions  $\psi_n$  near turning points via expansions in Airy functions (see, for instance, Chapter 11 of [25] or Chapter 4 of [9]). There are some numerical difficulties involved in computing the coefficients in such expansions; however, assuming that these can be overcome, the asymptotic complexity of the precomputation phase could be reduced to  $\mathcal{O}(\lambda)$  and a dramatic speedup in the running time of our algorithm effected.

A more challenging problem is to extend the results of this paper to the case of nonsymmetric scattering potentials. The notion of phase function extends easily to this case. Indeed, if  $u(x, y) =$

$\exp(\sigma(x, y))$  solves

$$\Delta u(x, y) + q(x, y)u(x, y) = 0, \quad (61)$$

then  $\sigma$  must satisfy the two-dimensional Riccati equation

$$\Delta\sigma(x, y) + |\nabla\sigma(x, y)|^2 + q(x, y) = 0. \quad (62)$$

Moreover, preliminary numerical experiments suggest that, just like its one-dimensional counterpart, (62) admits solutions which are nonoscillatory even when  $q$  is of large magnitude. It seems likely, then, that the rapid numerical solution of various boundary value problems for (61) can be carried out by constructing a collection of nonoscillatory solutions of (62).

Suppose, for example, that  $\Omega \subset \mathbb{R}^2$  is the disk of radius  $R > 0$  centered at 0. If

$$\sigma_{-m}, \sigma_{-m+1}, \dots, \sigma_{-1}, \sigma_0, \sigma_1, \dots, \sigma_{m-1}, \sigma_m \quad (63)$$

are nonoscillatory solutions of (62) such that for each  $n$  the restriction of  $\exp(\sigma_n(x, y))$  to the boundary of  $\Omega$  is equal to the exponential function  $\exp(int)$ . If the restriction of  $f$  to  $\partial\Omega$  admits the expansion

$$\sum_{n=-m}^m a_n \exp(int), \quad (64)$$

then the solution of the Dirichlet boundary value problem

$$\begin{cases} \Delta u(x, y) + q(x, y)u(x, y) = 0 & \text{in } \Omega \\ u(x, y) = f(x, y) & \text{on } \partial\Omega \end{cases} \quad (65)$$

is

$$\sum_{n=-m}^m a_n \exp(\sigma_n(x, y)). \quad (66)$$

Assuming that (62) can be solved in  $\mathcal{O}(\log(k))$  time, this would provide a method for solving (65) in the general case whose running time behaves as  $\mathcal{O}(\log(k))$ . This line of inquiry is being vigorously pursued by the author and will be reported on at a later date.

## 7. Acknowledgments

We thank the anonymous reviewers for their many helpful comments. We thank Vladimir Rokhlin for providing us with his code for evaluating the Hankel functions, and for several useful discussions. This work was supported in part by National Science Foundation grant DMS-1418723, and by a UC Davis Chancellor's Fellowship.

## References

- [1] ACHENBACH, J. D. *Wave propagation in elastic media*. North Holland, New York, 1980.
- [2] AMBIKASARAN, S., BORGES, C., IMBERT-GERARD, L., AND GREENGARD, L. Fast, adaptive, high-order accurate discretization of the Lippmann–Schwinger equation in two dimensions. *SIAM Journal on Scientific Computing* 38, 3 (2016), A1770–A1787.
- [3] BREMER, J. On the numerical solution of second order differential equations in the high-frequency regime. *Applied and Computational Harmonic Analysis* 44 (2018), 312–349.
- [4] BREMER, J., AND ROKHLIN, V. Improved estimates for nonoscillatory phase functions. *Discrete and Continuous Dynamical Systems, Series A* 36 (2016), 4101–4131.

[5] BRUNO, O. P., AND REITICH, F. High-order methods for high-frequency scattering applications. In *Modeling and Computations in Electromagnetics* (Berlin, Heidelberg, 2008), H. Ammari, Ed., Springer Berlin Heidelberg, pp. 129–163.

[6] CHANDLER-WILDE, S., GRAHAM, I., LANGDON, S., AND SPENCE, E. Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering. *Acta Numerica* 21 (04 2012), 89 – 305.

[7] COLTON, D., AND KRESS, R. *Inverse Acoustic and Electromagnetic Scattering Theory (Third Edition)*. Springer-Verlag, New York, 2013.

[8] ENGQUIST, B., AND YING, L. Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation. *Communications on Pure and Applied Mathematics* 64, 5 (2011), 697–735.

[9] FEDORYUK, M. V. *Asymptotic Analysis*. Springer-Verlag, 1993.

[10] GILBARG, D., AND TRUDINGER, N. *Elliptic Partial Differential Equations of Second Order (Third Edition)*. Springer, 1997.

[11] GILLMAN, A., BARNETT, A., AND MARTINSSON, P. A spectrally accurate direct solution technique for frequency-domain scattering problems with variable media. *BIT Numerical Mathematics* 55 (2015), 141–170.

[12] GRADSHTEYN, I. S., AND RYZHIK, I. M. *Table of Integrals, Series and Products*, Eighth ed. Elsevier, 2015.

[13] HIGHAM, N. *Accuracy and Stability of Numerical Algorithms*, second edition ed. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2002.

[14] HILLE, E. *Ordinary differential equations in the complex domain*. Wiley, New York, 1976.

[15] HOSKINS, J., AND ROKHLIN, V. A fast adaptive algorithm for scattering from a two-dimensional radially-symmetric potential. *Yale Technical Report TR-1538* (2019).

[16] IMBERT-GÉRARD, L.-M. Interpolation properties of generalized plane waves. *Numerische Mathematik* 131, 4 (Dec 2015), 683–711.

[17] IMBERT-GÉRARD, LISE-MARIE, AND MONK, PETER. Numerical simulation of wave propagation in inhomogeneous media using generalized plane waves. *ESAIM: M2AN* 51, 4 (2017), 1387–1406.

[18] IMBERT-GÉRARD, L.-M., AND DESPRÉS, B. A generalized plane-wave numerical method for smooth nonconstant coefficients. *IMA Journal of Numerical Analysis* 34, 3 (10 2013), 1072–1103.

[19] KUMMER, E. De generali quadam aequatione differentiali tertii ordinis. *Progr. Evang. Köngil. Stadtgymnasium Liegnitz* (1834).

[20] LANGDON, S., AND CHANDLER-WILDE, S. N. A wavenumber independent boundary element method for an acoustic scattering problem. *SIAM Journal on Numerical Analysis* 43, 6 (2006), 2450–2477.

[21] MALLAT, S. *A Wavelet Tour of Signal Processing (Third Edition)*. Academic Press, Boston, 2009.

- [22] MORSE, P. M., AND FESHBACH, H. *Methods of Mathematical Physics, Part II*. Feshbach Publishing, New York, 1998.
- [23] NEUMAN, F. *Global Properties of Linear Ordinary Differential Equations*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.
- [24] NGUYEN, N., PERAIRE, J., REITICH, F., AND COCKBURN, B. A phase-based hybridizable discontinuous galerkin method for the numerical solution of the helmholtz equation. *Journal of Computational Physics 290* (2015), 318 – 335.
- [25] OLVER, F. W. *Asymptotics and Special Functions*. A.K. Peters, Natick, MA, 1997.
- [26] SWARZTRAUBER, P. FFTPACK. <https://www.netlib.org/fftpack/> (1985).
- [27] WIDDER, D. *The Laplace Transform*. Dover, 2010.
- [28] WU, Y. M., AND CHEW, W. C. The modern high frequency methods for solving electromagnetic scattering problems (invited paper). *Progress In Electromagnetics Research 156* (2016), 63–82.

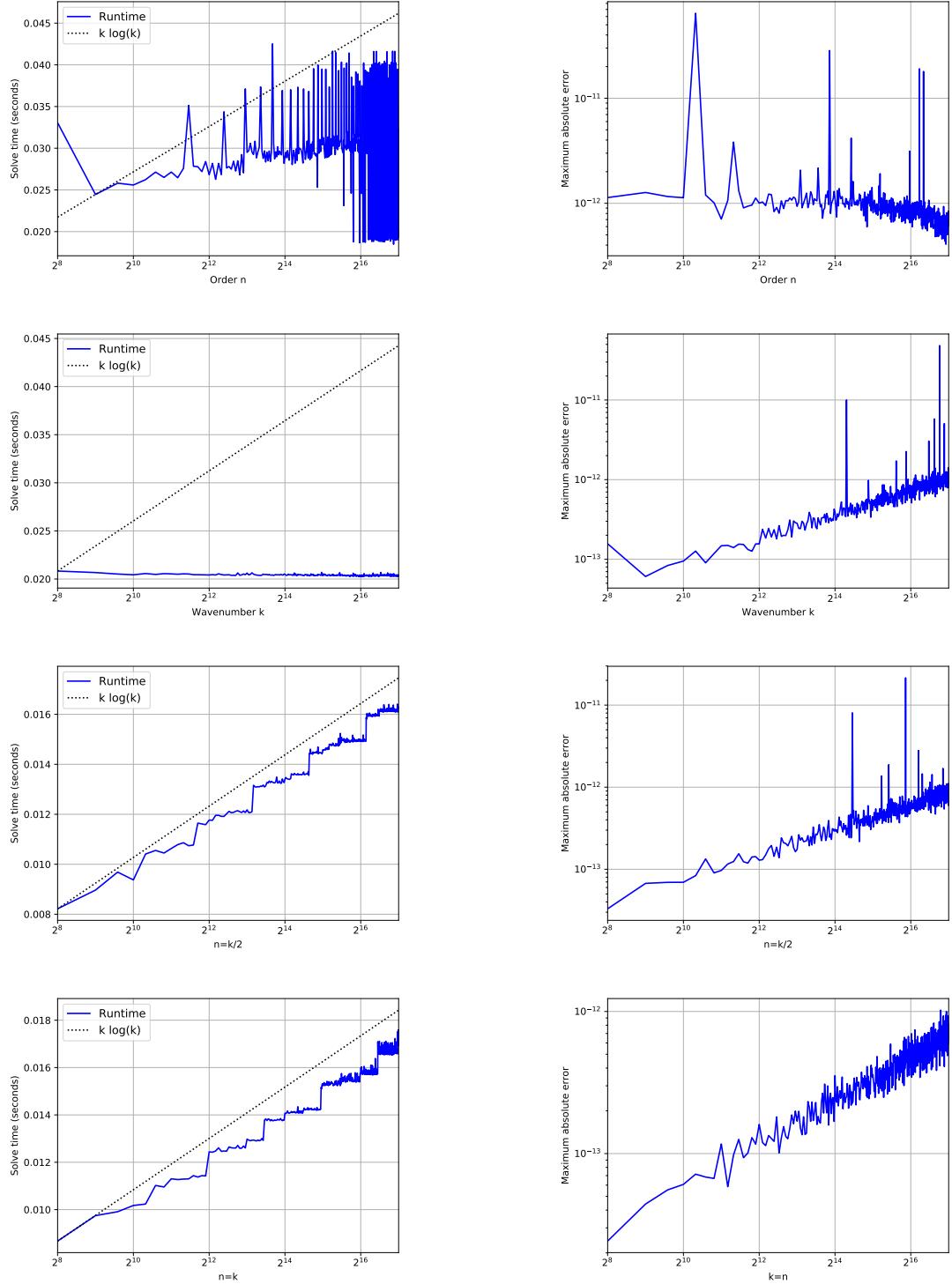


Figure 1: The results of the experiments described in Section 5.1 in which the performance of our solver for the perturbed Bessel equation was tested. In each of these experiments  $q(r) = r^2 - 1$  and  $R = 2$ . The first row gives the running time and largest observed absolute error as functions of  $n$  when  $k$  is held fixed at  $2^{17}$ . The second row gives them as functions of  $k$  when  $n = 0$  and  $k$  is increased from 0 to  $2^{17}$ . The third and fourth row give the running time and largest observed absolute error as functions of  $n$  when  $n$  is equal to a multiple of  $k$ ;  $n = k/2$  in third row and  $n = k$  in the fourth.

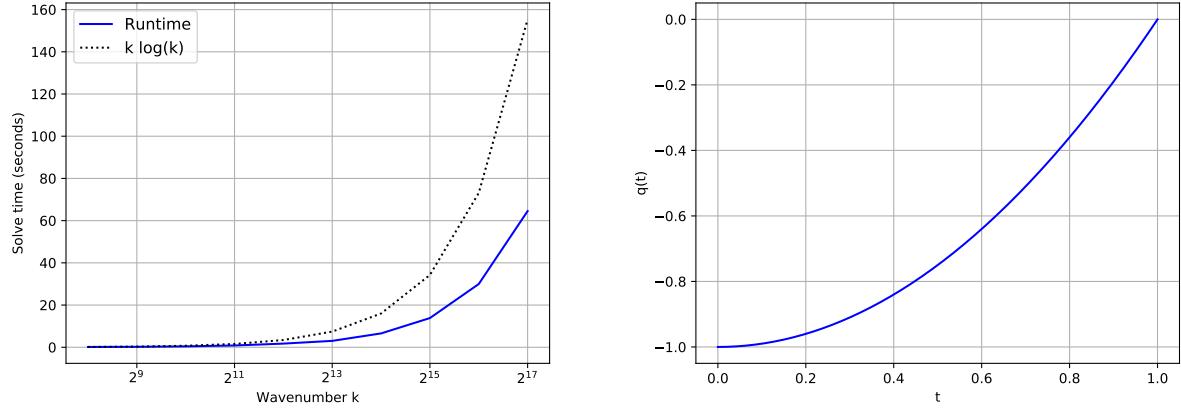


Figure 2: The results of the first experiment of Section 5.2. The plot in the upper left gives the time required to construct the set of solutions  $S_k = \{\varphi_n : n = 0, \dots, k\}$  of the perturbed Bessel equation when  $q(r) = 3\chi_{1,2}(r)$  as function of  $k$ . In the upper right is a graph of the function  $q(r)$ . Each row of the table corresponds to one value of  $k$  and reports the time  $t_k$  required to compute the set of solutions  $S_k$ , the maximum observed absolute error and, when applicable, the ratio of  $t_k$  to  $t_{k/2}$ .

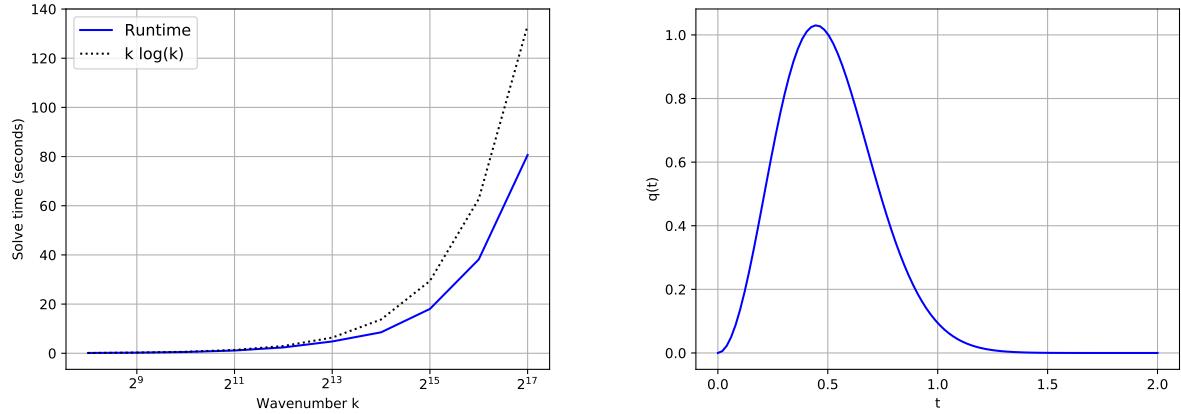


Figure 3: The results of the second experiment of Section 5.2. The plot in the upper left gives the time required to construct the set of solutions  $S_k = \{\varphi_n : n = 0, \dots, k\}$  of the perturbed Bessel equation when  $q(r) = 14r^2 \exp(-5r^2)$  as function of  $k$ . In the upper right is a graph of the function  $q(r)$ . Each row of the table corresponds to one value of  $k$  and reports the time  $t_k$  required to compute the set of solutions  $S_k$  as well as the ratio of  $t_k$  to  $t_{k/2}$  (when applicable).

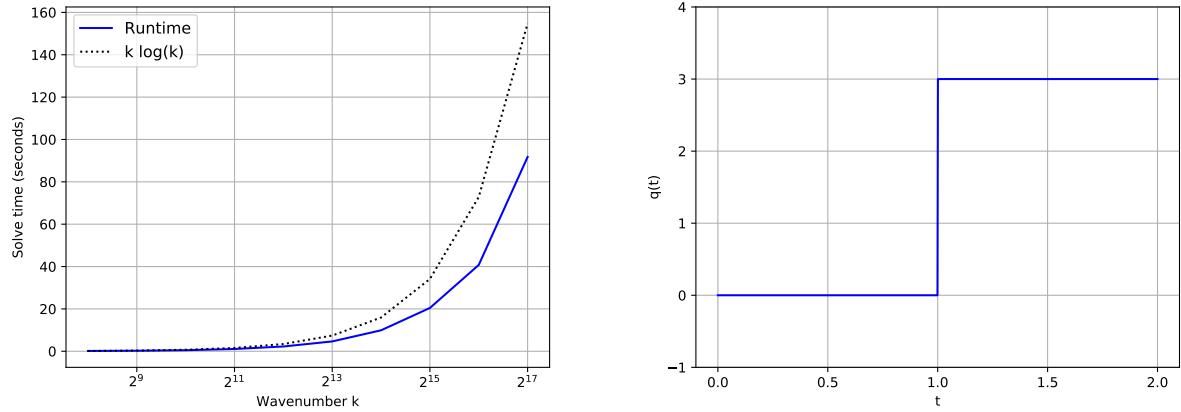


Figure 4: The results of the third experiment of Section 5.2. The plot in the upper left gives the time required to construct the set of solutions  $S_k = \{\varphi_n : n = 0, \dots, k\}$  of the perturbed Bessel equation when  $q(r) = 3\chi_{1,2}(r)$  as function of  $k$ . In the upper right is a graph of the function  $q(r)$ . Each row of the table corresponds to one value of  $k$  and reports the time  $t_k$  required to compute the set of solutions  $S_k$ , the maximum observed absolute error and, when applicable, the ratio of  $t_k$  to  $t_{k/2}$ .

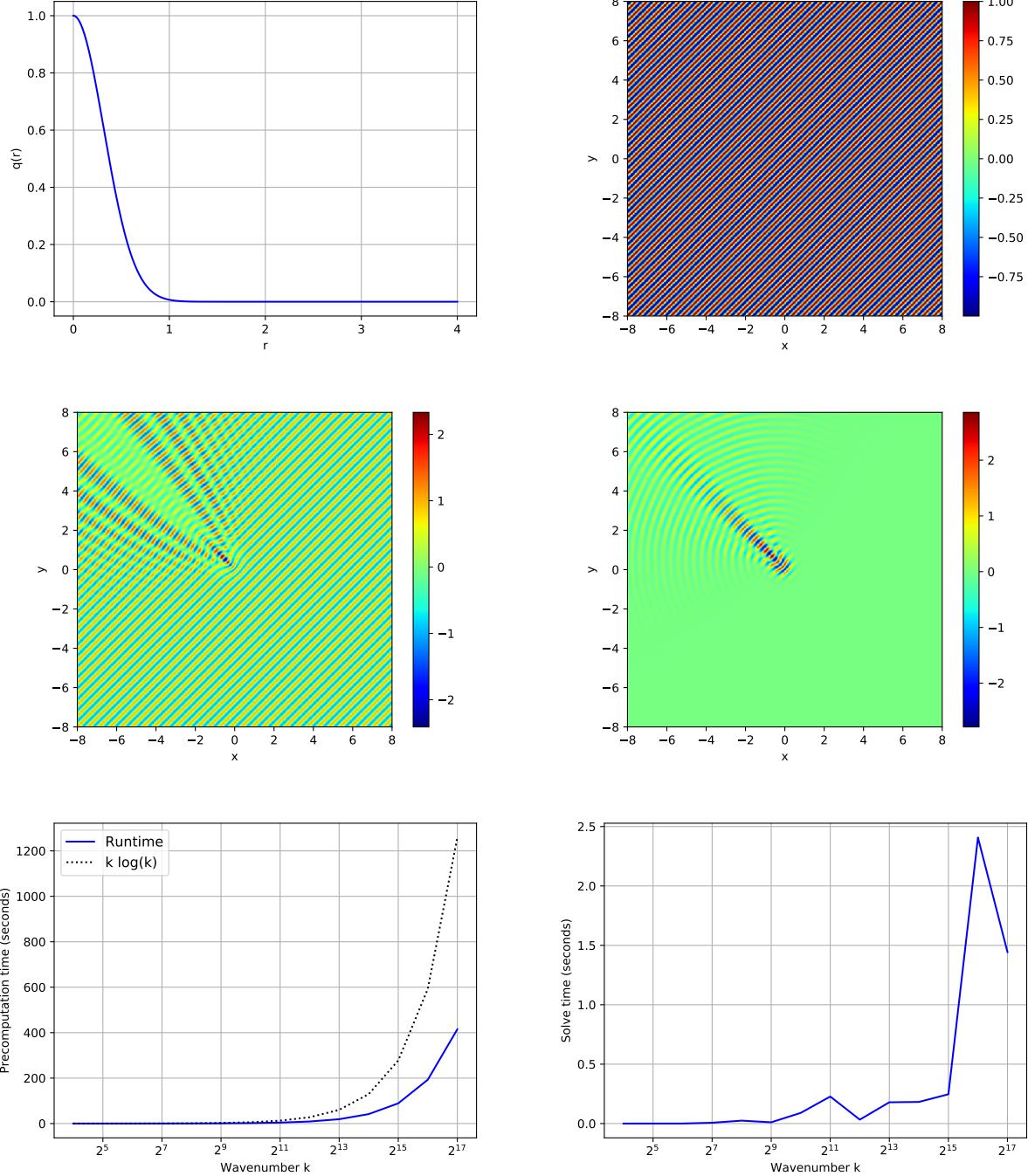


Figure 5: Figures related to the experiment of Section 5.3. In the upper left is a plot of the function  $q(r)$  and in the upper right is an image of the incident wave when  $k = 16$ . The image at middle left is of the real part of the total field  $k = 16$ . At middle right is an image of the real part of the scattered field when  $k = 16$ . At bottom left, the running time of the precomputation phase is plotted as a function of  $k$  and at bottom right, the running time of the solution phase is plotted as a function of  $k$ .

$k$	$m$	Maximum absolute error	Precomp time (in seconds)	Solve time (in seconds)
$2^4$	100	$8.34 \times 10^{-14}$	$3.61 \times 10^{-02}$	$1.71 \times 10^{-04}$
$2^5$	201	$1.55 \times 10^{-12}$	$5.38 \times 10^{-02}$	$3.03 \times 10^{-04}$
$2^6$	402	$1.26 \times 10^{-12}$	$1.07 \times 10^{-01}$	$3.95 \times 10^{-04}$
$2^7$	804	$1.22 \times 10^{-12}$	$2.22 \times 10^{-01}$	$6.88 \times 10^{-03}$
$2^8$	1608	$2.17 \times 10^{-12}$	$4.51 \times 10^{-01}$	$2.45 \times 10^{-02}$
$2^9$	3216	$4.72 \times 10^{-12}$	$9.47 \times 10^{-01}$	$1.06 \times 10^{-02}$
$2^{10}$	6433	$8.54 \times 10^{-12}$	$1.98 \times 10^{+00}$	$8.97 \times 10^{-02}$
$2^{11}$	12867	$(1.85 \times 10^{-11})$	$4.19 \times 10^{+00}$	$2.27 \times 10^{-01}$
$2^{12}$	25735	$(6.13 \times 10^{-11})$	$9.02 \times 10^{+00}$	$3.29 \times 10^{-02}$
$2^{13}$	51471	$(2.05 \times 10^{-10})$	$1.91 \times 10^{+01}$	$1.79 \times 10^{-01}$
$2^{14}$	102943	$(1.51 \times 10^{-09})$	$4.16 \times 10^{+01}$	$1.82 \times 10^{-01}$
$2^{15}$	205887	$(3.85 \times 10^{-09})$	$8.89 \times 10^{+01}$	$2.46 \times 10^{-01}$
$2^{16}$	411774	$(2.16 \times 10^{-08})$	$1.92 \times 10^{+02}$	$2.40 \times 10^{+00}$
$2^{17}$	823549	$(1.01 \times 10^{-07})$	$4.14 \times 10^{+02}$	$1.44 \times 10^{+00}$

Table 1: The results of the experiments of Section 5.3. Each row of the table corresponds to one wavenumber  $k$  and gives the number  $m$  of Fourier modes used to represent the incident wave, the maximum observed absolute error in the obtained solution (in cases in which this could be measured), and the time taken by each phase of our solver. The absolute maximum errors are calculated via comparison with solutions generated using extended precision arithmetic. Parentheses are used to indicate cases in which the wavenumber was too large for the the extended precision solution to be verified via a spectral method.

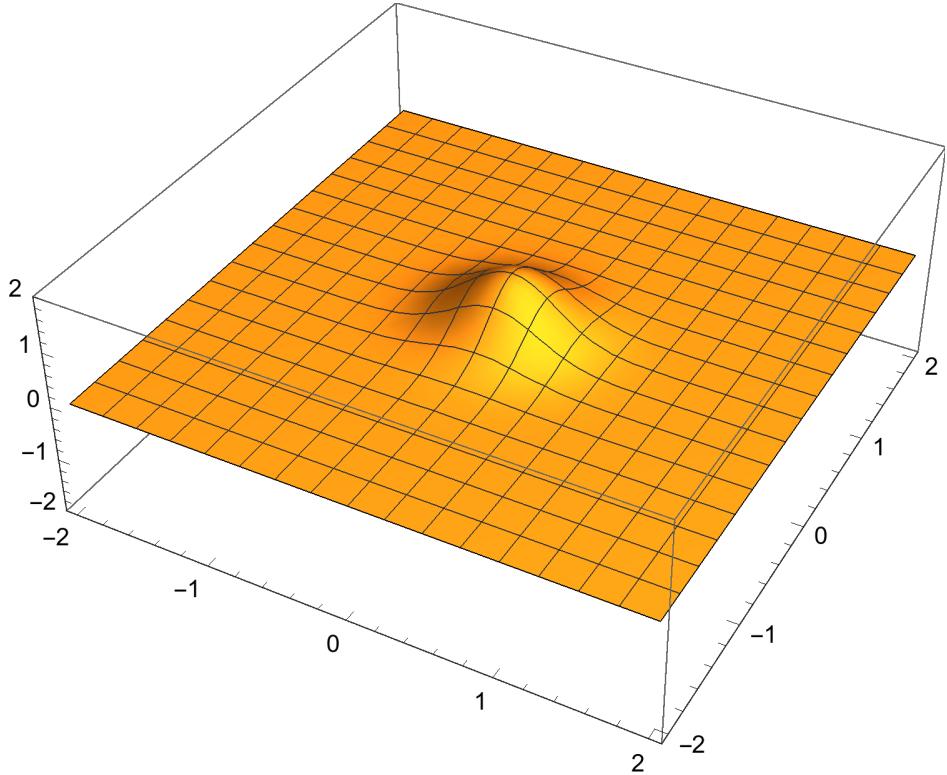


Figure 6: A plot of the scattering potential used in the experiment of Section 5.3.

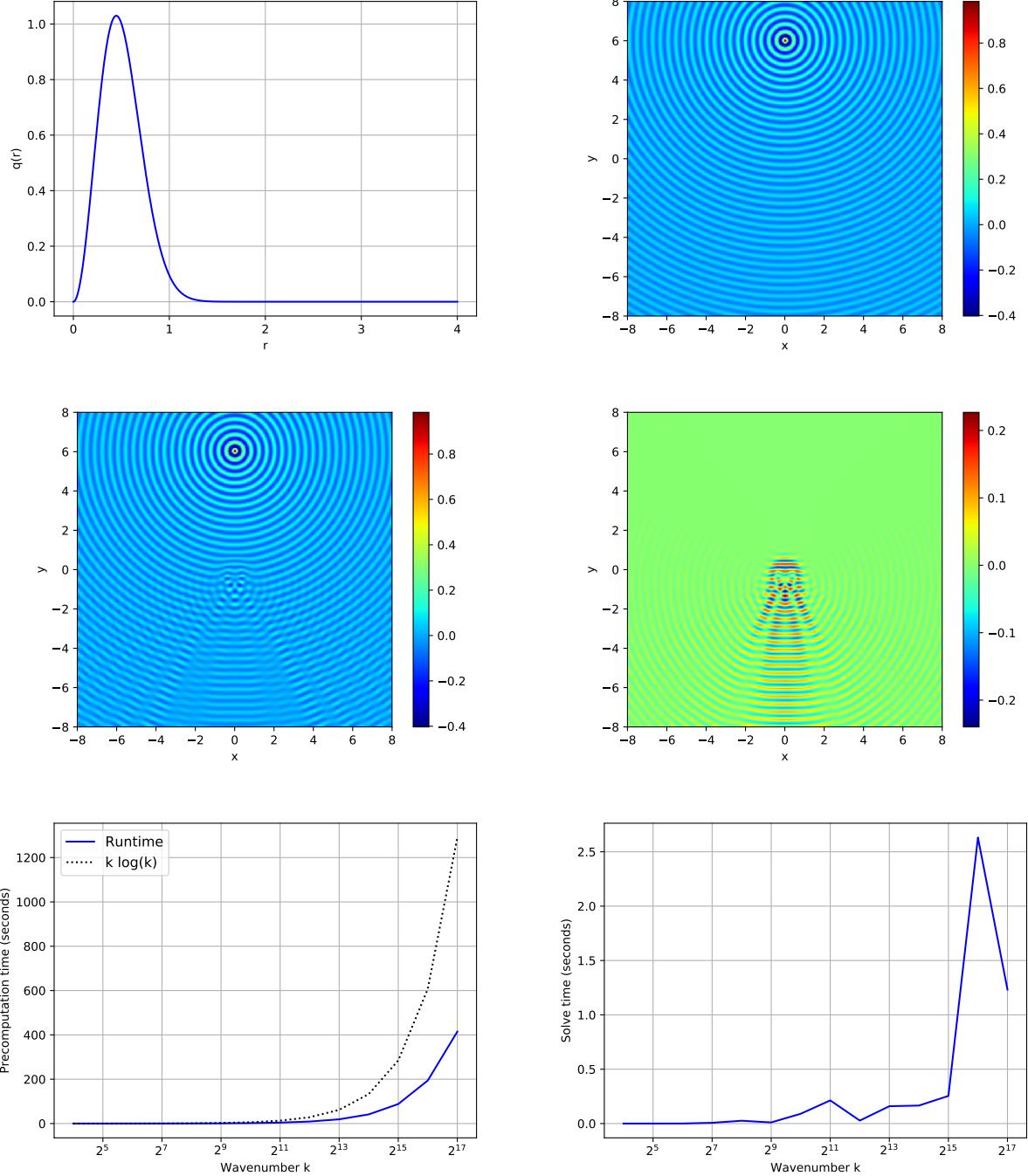


Figure 7: Figures related to the experiment of Section 5.4. In the upper left is a plot of the function  $q(r)$  and in the upper right is an image of the incident wave when  $k = 16$ . The image at middle left is of the real part of the total field  $k = 16$ . At middle right is an image of the real part of the scattered field when  $k = 16$ . At bottom left, the running time of the precomputation phase is plotted as a function of  $k$  and at bottom right, the running time of the solution phase is plotted as a function of  $k$ .

$k$	$m$	Maximum absolute error	Precomp time (in seconds)	Solve time (in seconds)
$2^4$	100	$9.81 \times 10^{-15}$	$3.70 \times 10^{-02}$	$1.95 \times 10^{-04}$
$2^5$	201	$6.61 \times 10^{-14}$	$5.39 \times 10^{-02}$	$2.84 \times 10^{-04}$
$2^6$	402	$5.96 \times 10^{-14}$	$1.08 \times 10^{-01}$	$4.45 \times 10^{-04}$
$2^7$	804	$2.94 \times 10^{-14}$	$2.25 \times 10^{-01}$	$6.99 \times 10^{-03}$
$2^8$	1608	$4.85 \times 10^{-14}$	$4.55 \times 10^{-01}$	$2.58 \times 10^{-02}$
$2^9$	3216	$6.01 \times 10^{-14}$	$9.43 \times 10^{-01}$	$1.04 \times 10^{-02}$
$2^{10}$	6433	$9.30 \times 10^{-14}$	$1.99 \times 10^{+00}$	$9.01 \times 10^{-02}$
$2^{11}$	12867	$(1.32 \times 10^{-13})$	$4.23 \times 10^{+00}$	$2.12 \times 10^{-01}$
$2^{12}$	25735	$(5.22 \times 10^{-13})$	$9.06 \times 10^{+00}$	$2.73 \times 10^{-02}$
$2^{13}$	51471	$(1.32 \times 10^{-12})$	$1.91 \times 10^{+01}$	$1.59 \times 10^{-01}$
$2^{14}$	102943	$(3.43 \times 10^{-12})$	$4.12 \times 10^{+01}$	$1.65 \times 10^{-01}$
$2^{15}$	205887	$(1.34 \times 10^{-11})$	$8.79 \times 10^{+01}$	$2.54 \times 10^{-01}$
$2^{16}$	411774	$(5.44 \times 10^{-11})$	$1.94 \times 10^{+02}$	$2.63 \times 10^{+00}$
$2^{17}$	823549	$(1.09 \times 10^{-10})$	$4.14 \times 10^{+02}$	$1.23 \times 10^{+00}$

Table 2: The results of the experiments of Section 5.4. Each row of the table corresponds to one wavenumber  $k$  and gives the number  $m$  of Fourier modes used to represent the incident wave, the maximum observed absolute error in the obtained solution (in cases in which this could be measured), and the time taken by each phase of our solver. The absolute maximum errors are calculated via comparison with solutions generated using extended precision arithmetic. Parentheses are used to indicate cases in which the wavenumber was too large for the the extended precision solution to be verified via a spectral method.

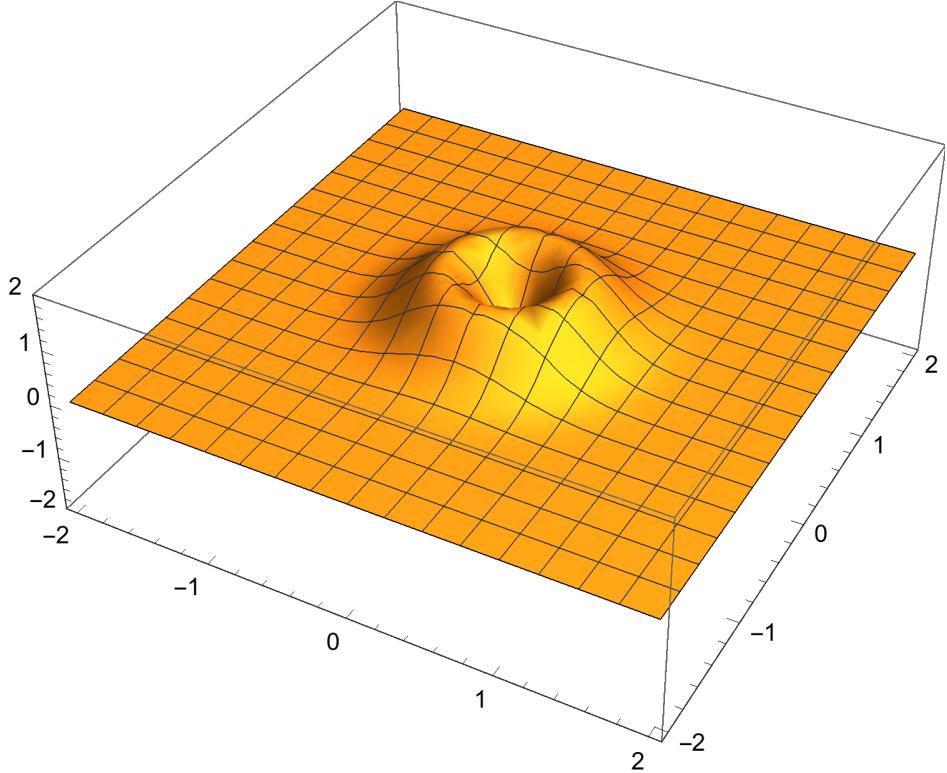


Figure 8: A plot of the volcano-shaped scattering potential used in the experiment of Section 5.4.

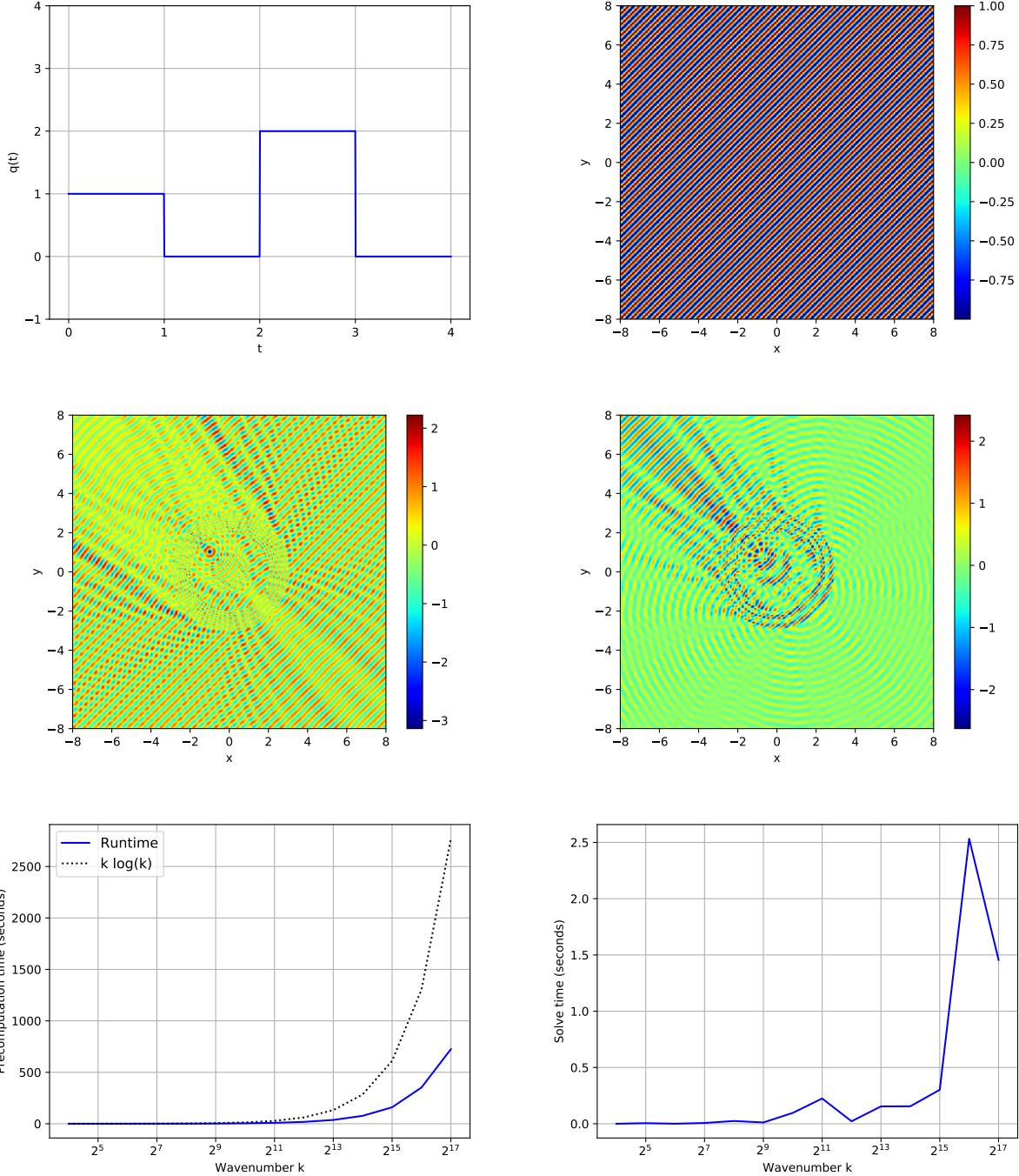


Figure 9: Figures related to the experiment of Section 5.5. In the upper left is a plot of the function  $q(r)$  and in the upper right is an image of the incident wave when  $k = 16$ . The image at middle left is of the real part of the total field  $k = 16$ . At middle right is an image of the real part of the scattered field when  $k = 16$ . At bottom left, the running time of the precomputation phase is plotted as a function of  $k$  and at bottom right, the running time of the solution phase is plotted as a function of  $k$ .

$k$	$m$	Maximum absolute error	Precomp time (in seconds)	Solve time (in seconds)
$2^4$	100	$8.03 \times 10^{-14}$	$7.95 \times 10^{-02}$	$1.81 \times 10^{-04}$
$2^5$	201	$1.20 \times 10^{-13}$	$1.57 \times 10^{-01}$	$5.22 \times 10^{-03}$
$2^6$	402	$3.15 \times 10^{-13}$	$3.23 \times 10^{-01}$	$4.14 \times 10^{-04}$
$2^7$	804	$8.20 \times 10^{-12}$	$6.48 \times 10^{-01}$	$6.87 \times 10^{-03}$
$2^8$	1608	$3.46 \times 10^{-12}$	$9.84 \times 10^{-01}$	$2.45 \times 10^{-02}$
$2^9$	3216	$8.46 \times 10^{-12}$	$2.05 \times 10^{+00}$	$1.13 \times 10^{-02}$
$2^{10}$	6433	$4.01 \times 10^{-11}$	$4.24 \times 10^{+00}$	$9.62 \times 10^{-02}$
$2^{11}$	12867	$(1.30 \times 10^{-10})$	$8.79 \times 10^{+00}$	$2.24 \times 10^{-01}$
$2^{12}$	25735	$(3.80 \times 10^{-10})$	$1.79 \times 10^{+01}$	$2.18 \times 10^{-02}$
$2^{13}$	51471	$(1.99 \times 10^{-09})$	$3.67 \times 10^{+01}$	$1.54 \times 10^{-01}$
$2^{14}$	102943	$(5.87 \times 10^{-09})$	$7.68 \times 10^{+01}$	$1.55 \times 10^{-01}$
$2^{15}$	205887	$(4.07 \times 10^{-08})$	$1.59 \times 10^{+02}$	$3.01 \times 10^{-01}$
$2^{16}$	411774	$(1.00 \times 10^{-07})$	$3.52 \times 10^{+02}$	$2.53 \times 10^{+00}$
$2^{17}$	823549	$(3.64 \times 10^{-07})$	$7.24 \times 10^{+02}$	$1.45 \times 10^{+00}$

Table 3: The results of the experiments of Section 5.5. Each row of the table corresponds to one wavenumber  $k$  and gives the number  $m$  of Fourier modes used to represent the incident wave, the maximum observed absolute error in the obtained solution (in cases in which this could be measured), and the time taken by each phase of our solver. The absolute maximum errors are calculated via comparison with solutions generated using extended precision arithmetic. Parentheses are used to indicate cases in which the wavenumber was too large for the the extended precision solution to be verified via a spectral method.

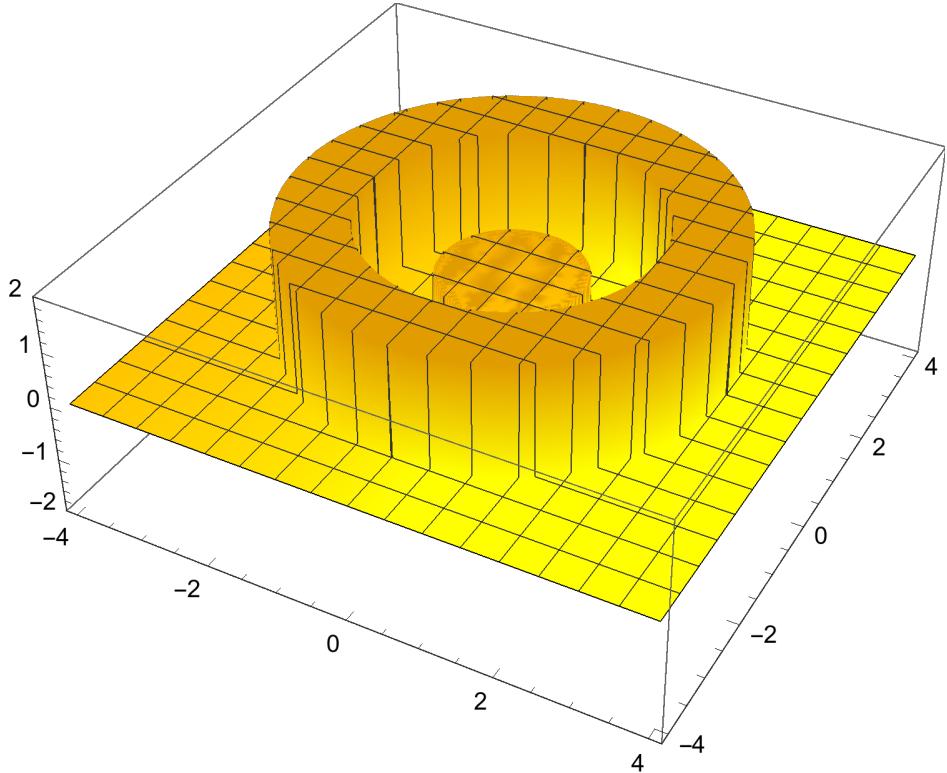


Figure 10: A plot of the discontinuous scattering potential used in the experiment of Section 5.5.

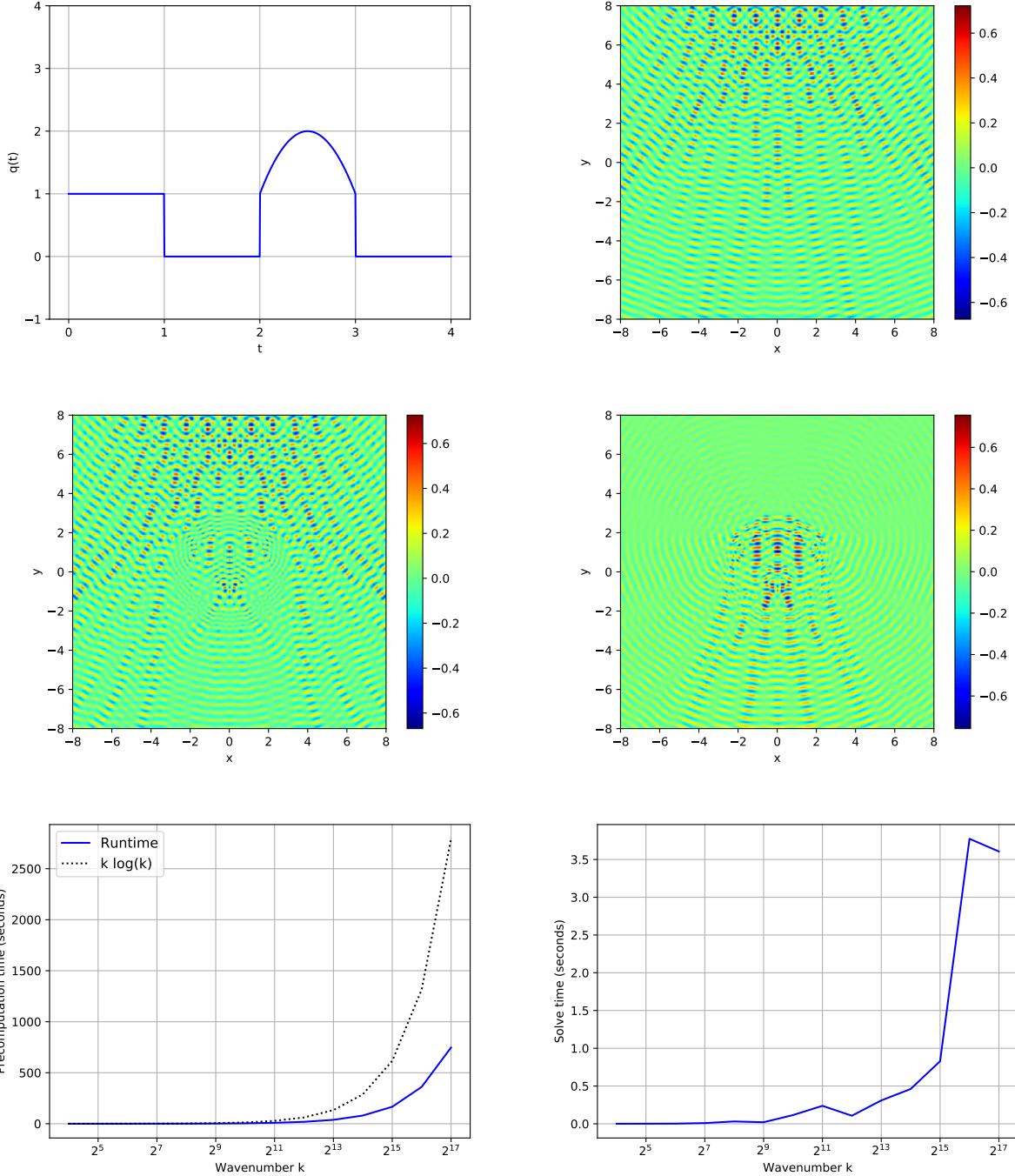


Figure 11: Figures related to the experiment of Section 5.6. In the upper left is a plot of the function  $q(r)$  and in the upper right is an image of the incident wave when  $k = 16$ . The image at middle left is of the real part of the total field  $k = 16$ . At middle right is an image of the real part of the scattered field when  $k = 16$ . At bottom left, the running time of the precomputation phase is plotted as a function of  $k$  and at bottom right, the running time of the solution phase is plotted as a function of  $k$ .

$k$	$m$	Maximum absolute error	Precomp time (in seconds)	Solve time (in seconds)
$2^4$	100	$2.71 \times 10^{-14}$	$8.02 \times 10^{-02}$	$6.33 \times 10^{-04}$
$2^5$	201	$1.71 \times 10^{-14}$	$1.57 \times 10^{-01}$	$1.18 \times 10^{-03}$
$2^6$	402	$2.90 \times 10^{-14}$	$3.33 \times 10^{-01}$	$2.35 \times 10^{-03}$
$2^7$	804	$7.41 \times 10^{-13}$	$6.55 \times 10^{-01}$	$9.38 \times 10^{-03}$
$2^8$	1608	$1.18 \times 10^{-13}$	$1.03 \times 10^{+00}$	$3.08 \times 10^{-02}$
$2^9$	3216	$4.27 \times 10^{-13}$	$2.16 \times 10^{+00}$	$2.04 \times 10^{-02}$
$2^{10}$	6433	$1.12 \times 10^{-12}$	$4.46 \times 10^{+00}$	$1.15 \times 10^{-01}$
$2^{11}$	12867	$(2.41 \times 10^{-12})$	$9.41 \times 10^{+00}$	$2.37 \times 10^{-01}$
$2^{12}$	25735	$(1.00 \times 10^{-11})$	$1.87 \times 10^{+01}$	$1.06 \times 10^{-01}$
$2^{13}$	51471	$(1.93 \times 10^{-11})$	$3.82 \times 10^{+01}$	$3.09 \times 10^{-01}$
$2^{14}$	102943	$(5.21 \times 10^{-11})$	$8.02 \times 10^{+01}$	$4.60 \times 10^{-01}$
$2^{15}$	205887	$(1.42 \times 10^{-10})$	$1.66 \times 10^{+02}$	$8.28 \times 10^{-01}$
$2^{16}$	411774	$(4.83 \times 10^{-10})$	$3.60 \times 10^{+02}$	$3.77 \times 10^{+00}$
$2^{17}$	823549	$(1.16 \times 10^{-09})$	$7.47 \times 10^{+02}$	$3.60 \times 10^{+00}$

Table 4: The results of the experiments of Section 5.6. Each row of the table corresponds to one wavenumber  $k$  and gives the number  $m$  of Fourier modes used to represent the incident wave, the maximum observed absolute error in the obtained solution (in cases in which this could be measured), and the time taken by each phase of our solver. The absolute maximum errors are calculated via comparison with solutions generated using extended precision arithmetic. Parentheses are used to indicate cases in which the wavenumber was too large for the the extended precision solution to be verified via a spectral method.

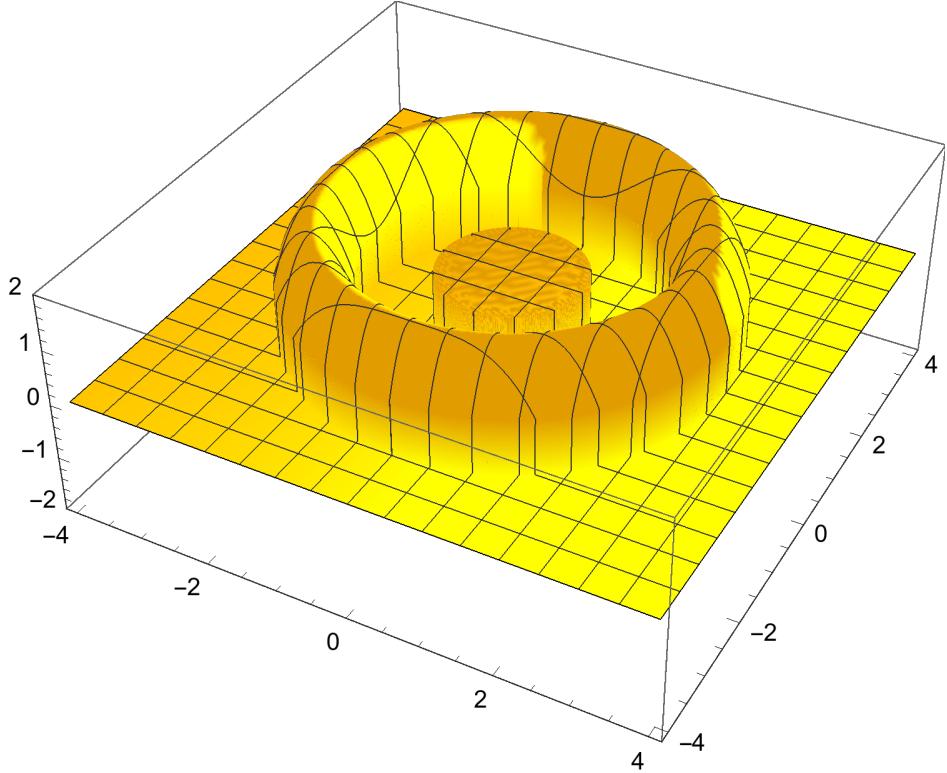


Figure 12: A plot of the discontinuous scattering potential used in the experiment of Section 5.6.